

Katastrophenrisiken und Extremwerttheorie

Dietmar Pfeifer

Einleitung und Überblick

Die jährlichen Schäden aus Naturkatastrophen wie Sturm (inklusive Hurrikane, Tornados und Taifune), Hagel, Überschwemmungen, Erdbeben, Erdbeben oder Lawinen belasten weltweit die Portfolios der Versicherungswirtschaft in erheblichem Maße. Finanziell am stärksten betroffen sind hiervon allerdings die international operierenden großen Rückversicherer, die ihre anteiligen Schadenlasten vor allem durch die globale Diversifikation des Geschäfts kontrollierbar halten.

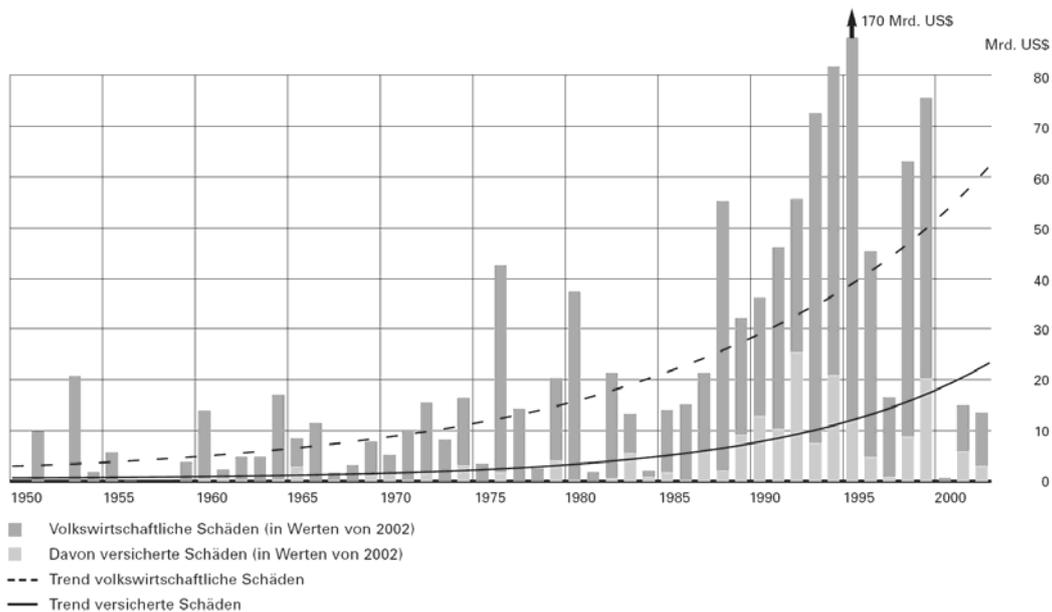


Abbildung 1: zeitliche Entwicklung der ökonomischen und versicherten Jahresschäden aus großen¹ Naturkatastrophen mit Trends; Quelle: Münchener Rück (Berz et al., 2003)

Die beobachteten zeitlichen Zunahmen sowohl der Schadenhöhen wie auch der Schadenfrequenzen sind dabei nur zum Teil durch die aktuell diskutierten möglichen klimatischen Veränderungen bedingt, sie gehen aber wesentlich auf ökonomische Faktoren wie Inflation oder zunehmende Wertkonzentrationen in für Naturgefahren exponierten Lagen zurück (z.B. Bebauung in küsten- und flussnahen Gebieten oder im Gebirge).

Die besondere Bedeutung versicherter Naturgefahren und anderer Katastrophenrisiken wird künftig auch ihren Niederschlag in den neuen europäischen Eigenkapitalanforderungen für Versicherungsunternehmen (Solvency II) finden, wo eine separate Berechnung des zugehörigen SCR (Solvency Capital Requirement) unter vereinfachter Berücksichtigung der vorhandenen Rückversicherung vorgesehen ist. In diesem Zusammenhang wird die im Rahmen von Solvency II grundsätzlich eingeräumte Möglichkeit zur Entwicklung so genannter „Interner Modelle“ interessant (Diers, 2007 und Brüske et al., 2010), mit denen eine gegenüber dem Standardansatz verbesserte individuelle Gefährdungseinschätzung des (Rück-)Versicherungsunternehmens erzielt werden kann.

¹ Als „groß“ werden nach Definition der Münchener Rück Naturkatastrophen bezeichnet, wenn sie die Selbsthilfefähigkeit der betroffenen Regionen deutlich übersteigen und überregionale oder internationale Hilfe erforderlich machen.

Eine möglichst genaue versicherungsmathematische Analyse und Modellierung von Katastrophenrisiken ist daher aus der Perspektive eines effizienten Risikomanagements für alle betroffenen Unternehmen (Erst- und Rückversicherer) von vitalem Interesse.

In der Praxis haben sich dabei zwei prinzipiell verschiedene Lösungsansätze herauskristallisiert, die in den nachfolgenden Abschnitten beide behandelt werden sollen. Zum einen sind dies Analysen auf der Basis so genannter *geophysikalischer Modelle*, bei denen die jährlichen Schadenfrequenzen und Einzelschadenhöhen im Rahmen eines geeigneten *kollektiven Modells der Risikotheorie* auf der Basis geophysikalischer oder meteorologischer Daten aus historischen Ereignissen erfasst werden. Zum anderen sind dies stochastische Ansätze, die im Rahmen der *Extremwertstatistik* auf ökonomisch geeignet bereinigten aggregierten Jahreschadendaten beruhen. Während die geophysikalischen Modelle in der Regel nur als *Vendor Models* von kommerziellen Anbietern ohne explizite Offenlegung der detaillierten Funktionsweise zur Verfügung stehen² (und damit meist auch nur für größere Unternehmen finanziell tragbar sind), können extremwertstatistische Analysen leicht mit gängigen Statistikpaketen oder den üblichen Tabellenkalkulationsprogrammen durchgeführt werden; vgl. auch das in dem Buch von Reiss/Thomas (2007) besprochene Software-Tool X_{TRE}MES. Wenn möglich, sollten für eine abschließende Risikoeinschätzung beide Verfahren zur Anwendung kommen.

Grundlagen geophysikalischer Modelle

Insbesondere im Bereich der Elementargefahren Sturm, Hochwasser und Erdbeben nutzen große Erst- und Rückversicherer, aber zunehmend auch Rückversicherungsmakler und andere Consulting-Firmen professionelle Software-Werkzeuge, die möglichst realitätsgerecht eine Vielzahl verschiedener Szenarien generieren. Diese beruhen zwar in erster Linie auf historischen Datenbanken (*Historic Event Sets*); durch stochastische Perturbationen der hinterlegten geophysikalischen Parameter wie Windrichtung und Zugbahn, Windgeschwindigkeit, Sturmdauer, Lage der Epizentren von Erdbeben, Art und Stärke der Schockwellenausbreitung usw. lassen sich hieraus aber leicht repräsentative virtuelle Szenarien generieren (*Stochastic Event Sets*), die mühelos 50000 und mehr Einträge umfassen können. Wählt man unter diesen Szenarien diejenigen aus, die ein bestimmtes Versicherungsportfolio tangieren, lassen sich auf diese Weise mit dem Computer viele Hundert oder Tausend Schadenjahre simulieren und so wichtige Informationen über die Einzel- wie Gesamtschadenverteilung sammeln.

Zum besseren Verständnis dieser Diskussion sollen hier zunächst die üblichen mathematischen Grundlagen der geophysikalischen Simulationsmodelle vorgestellt werden (in Anlehnung an Pfeifer, 2004), vor allem auch darum, weil detaillierte Dokumentationen solcher Produkte nicht leicht zugänglich sind (vgl. etwa Weimin Dong, 2001). Die hier besprochenen Modelle beruhen alle auf einem klassischen Grundmodell, dem so genannten *kollektiven Modell der Risiko-Theorie*, das folgendermaßen charakterisiert werden kann:

- Die Anzahl N der Schäden innerhalb einer bestimmten Periode (in der Regel ein Jahr) ist eine nicht-negative, ganzzahlige Zufallsvariable. Die Einzelwahrscheinlichkeiten seien mit den Größen $p_n = P(N = n)$, $n = 0, 1, \dots$ bezeichnet.
- Die in dieser Periode anfallenden Einzelschadenhöhen X_1, X_2, \dots sind stochastisch unabhängige, identisch verteilte, positive Zufallsvariablen, die auch von N stochastisch unabhängig sind. Die zugehörige Verteilungsfunktion sei mit F bezeichnet.

² Dies ist u.a. ein Grund dafür, warum die europäischen nationalen Aufsichtsbehörden augenblicklich noch keine automatisierte Zertifizierung solcher Software als Interne Modelle zulassen.

Unter dem *Gesamtschaden* (für die betrachtete Periode) versteht man dann die Zufallsvariable $S := \sum_{k=1}^N X_k$ mit der Konvention, dass die *leere Summe* als Null verstanden wird. Für die Verteilungsfunktion F_S des Gesamtschadens gilt dann:

$$P(S \leq z) = F_S(z) = p_0 + \sum_{n=1}^{\infty} p_n F^{n*}(z), \quad z \in \mathbb{R}.$$

Dabei bezeichnet F^{n*} die n -fache Faltung³ von F , der Verteilungsfunktion der Einzelschadenhöhen. Ein Beweis dieser Beziehung ergibt sich aus

$$\begin{aligned} P(S \leq z) = F_S(z) &= P\left(\bigcup_{n=0}^{\infty} \{N = n\} \cap \left\{\sum_{k=1}^n X_k \leq z\right\}\right) = p_0 + \sum_{n=1}^{\infty} p_n P\left(\sum_{k=1}^n X_k \leq z\right) \\ &= p_0 + \sum_{n=1}^{\infty} p_n F^{n*}(z), \quad z \in \mathbb{R}. \end{aligned}$$

In den geophysikalischen Simulationsmodellen bildet nun jedes einzelne Szenario i ein solches kollektives Modell. Wenn dabei n die Anzahl der für das Portfolio relevanten Szenarien bezeichnet und wir mit N_1, N_2, \dots, N_n die als voneinander unabhängig angenommenen Ereignisfrequenzen sowie mit X_{ij} , $1 \leq i \leq n$, $j \in \mathbb{N}$ die ebenfalls als unabhängig angenommenen positiven Einzelschadenhöhen bezeichnen, wobei alle $X_{i\cdot}$ dieselbe Verteilung Q_i (die von i abhängen dürfen) besitzen mögen, so erhalten wir folgende Summenschäden:

$$S_i := \sum_{j=1}^{N_i} X_{ij}, \quad i = 1, \dots, n \quad (\text{Ereignis-Gesamtschaden, Scenario Loss})$$

$$S := \sum_{i=1}^n S_i = \sum_{i=1}^n \sum_{j=1}^{N_i} X_{ij} \quad (\text{Jahres-Gesamtschaden, Aggregate Loss}).$$

Eine typische Grundannahme in solchen Modellen ist, dass die Frequenzen durch Poisson-verteilte Zufallsvariablen mit Parametern $\lambda_1, \lambda_2, \dots, \lambda_n > 0$ modelliert werden können. Der wesentliche Grund hierfür liegt darin, dass aufgrund eines wohlbekannten Satzes (vgl. Kaas et al., 2001, Abschnitt 3.4) der Jahres-Gesamtschaden dann aufgefasst werden kann als Summenschaden aus unabhängigen, identisch wie \tilde{X} verteilten Zufallsvariablen, die der Mischverteilung $\tilde{Q} = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} Q_i$ mit $\tilde{\lambda} = \sum_{i=1}^n \lambda_i$ genügen, mit einer einzigen Frequenzvariablen N , die

ihrerseits wieder Poisson-verteilt ist mit Parameter $\tilde{\lambda}$. Es ist damit möglich, das ursprünglich sehr viel komplexere Modell durch ein einfacheres, klassisches kollektives Modell der Risiko-Theorie darzustellen.

Neben dem jeweiligen Ereignis-Gesamtschaden ist in den geophysikalischen Modellen auch der *Ereignis-Maximalschaden* $M_i := \max\{X_{ij} \mid 1 \leq j \leq N_i\}$ von Interesse. Seine Verteilung

³ Die Faltung $G * H$ zweier Verteilungsfunktionen G und H entspricht der Verteilungsfunktion der Summe $U + V$ zweier stochastisch unabhängiger Zufallsvariablen U und V , die G und H als Verteilungsfunktion besitzen.

kann in einem Poisson-Modell ebenfalls recht einfach berechnet werden. Allgemein gilt hier folgendes:

Im kollektiven Modell der Risikotheorie sei mit $M := \max \{X_i \mid 1 \leq i \leq N\}$ der *Maximalschaden* bezeichnet. Dann gilt:

$$P(M \leq z) = F_M(z) = \sum_{n=0}^{\infty} p_n F^n(z) = \varphi_N(F(z)), \quad z \in \mathbb{R}.$$

Hierbei bezeichnet die Potenzreihe $\varphi_N(t) = E(t^N) = \sum_{n=0}^{\infty} p_n t^n$ die so genannte *wahrscheinlichkeitserzeugende Funktion* von N .

Ein einfacher Beweis dieser Beziehung ergibt sich durch

$$\begin{aligned} P(M \leq z) &= F_M(z) = P\left(\bigcup_{n=0}^{\infty} \{N = n\} \cap \{\max \{X_k \mid 1 \leq k \leq n\} \leq z\}\right) \\ &= p_0 + \sum_{n=1}^{\infty} p_n P\left(\bigcap_{k=1}^n \{X_k \leq z\}\right) = p_0 + \sum_{n=1}^{\infty} p_n F^n(z) = \sum_{n=0}^{\infty} p_n F^n(z) = \varphi_N(F(z)). \end{aligned}$$

Der Unterschied zwischen der Verteilung des Gesamt- und des Maximalschadens ist also lediglich der, dass im ersten Fall *Faltungspotenzen*, im zweiten Fall *gewöhnliche* Potenzen der Verteilungsfunktionen auftreten.

Für das Poisson-Modell, wo die Schadenfrequenz N Poisson-verteilt ist mit Parameter $\lambda > 0$, bedeutet dies gerade:

$$P(M \leq z) = F_M(z) = \sum_{n=0}^{\infty} p_n F^n(z) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} F^n(z) = e^{-\lambda} \exp\{\lambda F(z)\} = e^{-\lambda\{1-F(z)\}}, \quad z \in \mathbb{R}.$$

Übertragen auf die geophysikalischen Simulationsmodelle ergibt sich damit, dass der maximale Ereignisschaden M_i die folgende Verteilungsfunktion besitzt:

$$P(M_i \leq z) = F_{M_i}(z) = e^{-\lambda_i\{1-F_i(z)\}}, \quad z \in \mathbb{R}.$$

Hiermit lässt sich wegen der angenommenen Unabhängigkeit aller Einzelschäden sofort auch die Verteilungsfunktion des *Jahres-Maximalschadens* $M = \max \{M_i \mid 1 \leq i \leq n\}$ (*Occurrence Loss*) angeben:

$$\begin{aligned} P(M \leq z) &= P\left(\bigcap_{i=1}^n \{M_i \leq z\}\right) = \prod_{i=1}^n P(M_i \leq z) = \prod_{i=1}^n e^{-\lambda_i\{1-F_i(z)\}} \\ &= \exp\left\{-\sum_{i=1}^n [\lambda_i \{1-F_i(z)\}]\right\} = \exp\{-\tilde{\lambda}[1-\tilde{F}(z)]\}, \end{aligned}$$

wobei $\tilde{F}(z) = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} F_i(z)$, $z \in \mathbb{R}$ die Verteilungsfunktion zur Mischverteilung \tilde{Q} bezeichnet.

In der geophysikalischen Modellierungswelt werden nun insbesondere die Überschreitungswahrscheinlichkeiten des *Occurrence Loss* und des *Aggregate Loss* betrachtet, also die Funktionen

$$P(M > z) = 1 - \exp\left\{-\tilde{\lambda}\left[1 - \tilde{F}(z)\right]\right\}, \quad z \in \mathbb{R} \quad (\text{Occurrence Loss Exceeding Probability, OEP})$$

$$P(S > z) = 1 - e^{-\tilde{\lambda}} - e^{-\tilde{\lambda}} \sum_{n=1}^{\infty} \frac{\tilde{\lambda}^n}{k!} \tilde{F}^{n*}(z), \quad z \in \mathbb{R} \quad (\text{Aggregate Loss Exceeding Probability, AEP})$$

(vgl. Weimin Dong, 2001, S. 14ff und 46ff sowie Grossi/Kunreuther, 2005, Chapter 2 und 6). Für die Berechnung der Faltungspotenzen in der *AEP* gibt es geeignete Algorithmen, die meist auf einer geeigneten Diskretisierung der Schäden beruhen. Populär sind beispielsweise Rekursionsverfahren vom Panjer-Typ oder die diskrete Fourier-Transformation (vgl. Klugman et al., 2004, Chapter 6 und Brüske et al., 2010). Mit Hilfe moderner Computer-Algebra-Systeme kann man aber auch die zugehörigen wahrscheinlichkeitserzeugenden Funktionen direkt in Potenzreihen mit „vielen“ Termen entwickeln, aus denen man die Verteilung des Gesamtschadens sofort ablesen kann. Bezeichnet nämlich X_{Δ} den diskretisierten individuellen Schaden in einem kollektiven Modell (mit Schrittweite $\Delta > 0$), so ist durch $\varphi_S(t) = \varphi_N(\varphi_{X_{\Delta}}(t))$ die wahrscheinlichkeitserzeugende Funktion des Gesamtschadens gegeben, weil die Faltungspotenzen in Produkte der erzeugenden Funktionen übergehen (vgl. Klugman et al., 2004, Chapter 6).

Alternativ ist auch eine statistische Bestimmung der *AEP* mittels Monte-Carlo-Simulation möglich, indem eine große Zahl von zufälligen Beobachtungen aus dem Modell gezogen und die zugehörige *empirische Verteilung* bestimmt wird. Allerdings sind hier je nach Parameter-Situation bis zu einigen 100000 Stichproben notwendig.

Eine weitere Vereinfachung der Berechnungen ergibt sich für die *OEP* in der *Basisversion* aller geophysikalischen Simulationsmodelle, in der angenommen wird, dass die vorkommenden Schadenhöhen *deterministisch* sind, d.h. dass die Verteilungen Q_i Einpunktverteilungen mit Massen $\varpi_i \in \mathbb{R}$, $i = 1, \dots, n$ darstellen. Für die nachfolgenden Überlegungen können wir dabei annehmen, dass diese strikt der Größe nach geordnet sind, d.h. dass gilt

$$\varpi_1 < \varpi_2 < \dots < \varpi_n.$$

Es ist dann $F_i(\varpi_k) = 1$ für $1 \leq i \leq k$ und $F_i(\varpi_k) = 0$ für $i > k$, so dass hieraus

$$P(\tilde{X} \leq \varpi_k) = \tilde{F}(\varpi_k) = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} F_i(\varpi_k) = \sum_{i=1}^k \frac{\lambda_i}{\tilde{\lambda}}, \quad k = 1, \dots, n$$

folgt. Für die *OEP* bedeutet dies:

$$P(M > \varpi_k) = 1 - \exp\left\{-\tilde{\lambda}\left[1 - \tilde{F}(\varpi_k)\right]\right\} = 1 - \exp\left\{-\sum_{i=k+1}^n \lambda_i\right\}, \quad k = 1, \dots, n.$$

Die folgende Graphik, die mit einem Computer-Algebra-System erzeugt wurde, zeigt die *OEP* und *AEP* für ein virtuelles Beispiel mit $n = 300$ Szenarien und $\tilde{\lambda} = 2,465$. Der maximale Einzelschaden beträgt hier $\varpi_{300} = 489909$. Für die Berechnung der *AEP* wurde eine Diskretisierung der Schäden mit Schrittweite $\Delta = 2500$ gewählt. Man beachte, dass beide Kurven mit einem Wert von $1 - e^{-\tilde{\lambda}} = 0,915$ im Nullpunkt beginnen, was bedeutet, dass mit Wahrscheinlichkeit $e^{-\tilde{\lambda}} = 0,085$ innerhalb eines Jahres keine Schäden auftreten.

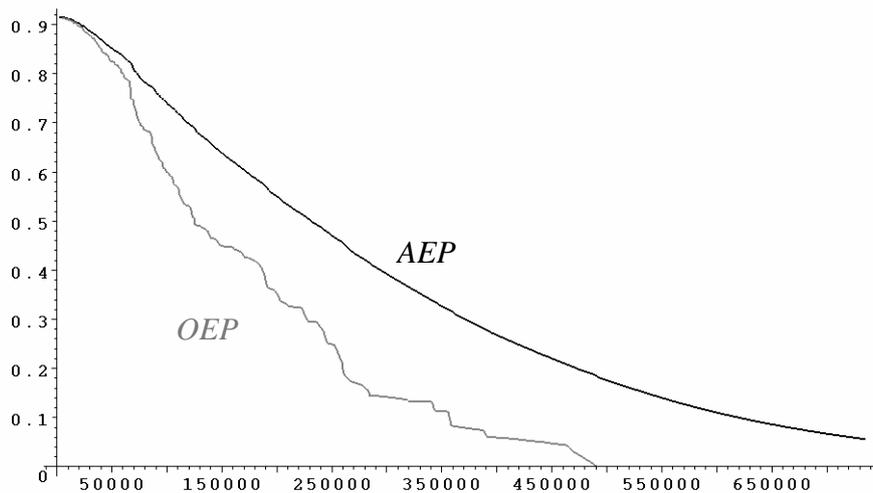


Abbildung 2: Vergleich der AEP und OEP für ein virtuelles Beispiel

In den neueren Versionen der *Vendor Models* werden neben der Poisson-Verteilung als Frequenzverteilung auch andere Verteilungen, z.B. die häufig auf historische Datensätze besser passende *negative Binomialverteilung*, betrachtet. Für die Schadenhöhenverteilungen sind neben deterministischen insbesondere Lognormalverteilungen und Beta-Verteilungen (mit der jeweiligen individuellen Versicherungssumme als Obergrenze) üblich. In solchen Fällen lassen sich allerdings im Allgemeinen keine geschlossenen Formeln mehr für die OEP und AEP angeben, auch nicht im Basismodell, so dass diese Kurven nur mittels umfangreicher Monte-Carlo-Simulationen bestimmt werden können.

Neben den hier besprochenen mathematisch-statistischen Grundlagen zeichnen sich alle *Vendor Models* durch folgenden modularen Aufbau aus (vgl. Grossi/Kunreuther, 2005, Chapter 2):

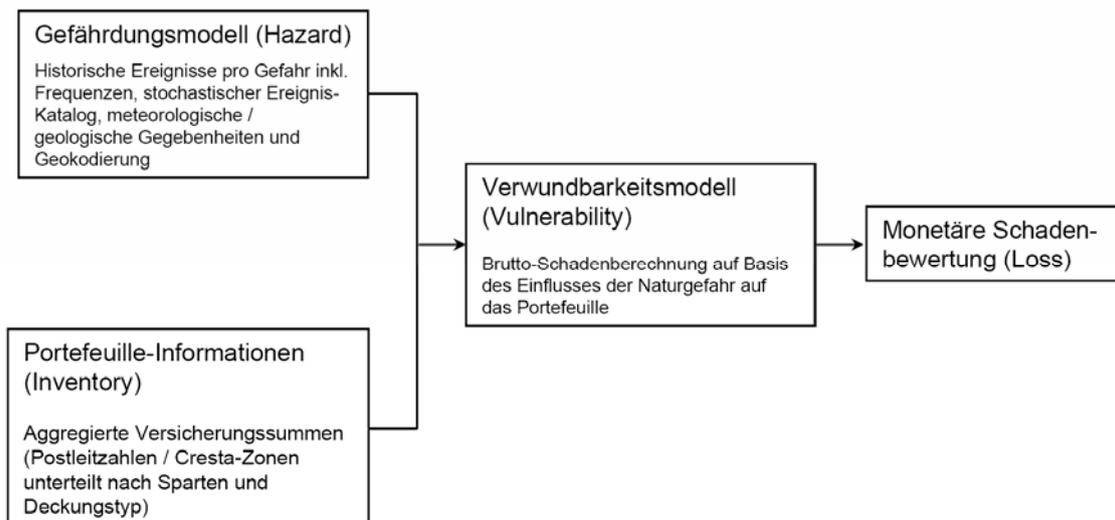


Abbildung 3: modularer Aufbau geophysikalischer Modelle

Die verschiedenen geophysikalischen Modelle unterscheiden sich hauptsächlich im Design der (im Allgemeinen nicht offen gelegten) *Hazard*- und *Vulnerability*-Module, die stark von dahinter stehenden ingenieurwissenschaftlichen Ansätzen (Meteorologie, Geologie) geprägt sind.

Grundzüge der Extremwertstatistik

Historisch geht die Extremwertstatistik auf Untersuchungen des Grenzverhaltens geeignet normalisierter extremer Werte einer Datenreihe, etwa des laufenden Maximums, zurück (Satz von Fisher und Tippett (1928), siehe Beirlant et al., 2004, Chapter 2.1, Embrechts et al., 2001, Chapter 3.2, oder Reiss/Thomas, 2007, S. 18). Sind X_1, \dots, X_n stochastisch unabhängige Zufallsvariablen (hier typischerweise Risiken) und bezeichnet $X_{(n)} = \max(X_1, \dots, X_n)$ für $n \in \mathbb{N}$, so gilt ähnlich der obigen Rechnung für den Jahresmaximalschaden

$$P(X_{(n)} \leq x) = P\left(\bigcap_{k=1}^n \{X_k \leq x\}\right) = F^n(x), \quad x \in \mathbb{R},$$

wobei wieder F die Verteilungsfunktion der X_k bezeichne (vgl. Embrechts et al., 2001, Proposition 4.1.2). Mit der Notation $\mu_{(n)} = E(X_{(n)})$, $\sigma_{(n)}^2 = \text{Var}(X_{(n)})$ kann man nun versuchen, ähnlich zum Zentralen Grenzwertsatz einen Grenzwert für die normalisierten Maxima $Z_n = \frac{X_{(n)} - \mu_{(n)}}{\sigma_{(n)}}$ zu finden⁴. Wir zeigen hier exemplarisch am Beispiel der stetigen Gleichverteilung über dem Intervall $[0,1]$, dass dies in gewissen Fällen funktioniert. Wegen $F(x) = x$, $0 \leq x \leq 1$ ist $f_{(n)}(x) = n x^{n-1}$, $0 \leq x \leq 1$ eine Dichte von $X_{(n)}$, woraus

$$\mu_{(n)} = \int_0^1 x \cdot f_{(n)}(x) dx = \frac{n}{n+1},$$

$$\sigma_{(n)}^2 = E(X_{(n)}^2) - \mu_{(n)}^2 = \int_0^1 x^2 \cdot f_{(n)}(x) dx - \mu_{(n)}^2 = \frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2 = \frac{n}{(n+1)^2(n+2)}$$

folgt. Dies führt zu

$$P(Z_n \leq x) = P(X_{(n)} \leq \sigma_{(n)}x + \mu_{(n)}) = (\sigma_{(n)}x + \mu_{(n)})^n = \left(1 - \frac{1 - \sqrt{\frac{n}{n+2}}x}{n+1}\right)^n$$

$$\sim \left(1 - \frac{1-x}{n}\right)^n \xrightarrow{n \rightarrow \infty} \exp(-(1-x)), \quad x \leq 1.$$

Die Grenzverteilung ist hier also eine gespiegelte und verschobene Exponentialverteilung.

Allgemeiner untersucht man in der Extremwertstatistik die Existenz von Grenzverteilungen für normalisierte Maxima der Form

$$Z_n = \frac{X_{(n)} - b_n}{a_n},$$

⁴ Die Z_n besitzen also alle den Erwartungswert 0 und die Varianz 1.

wobei $\{a_n\}_{n \in \mathbb{N}}$ und $\{b_n\}_{n \in \mathbb{N}}$ geeignete (positive bzw. reelle) Zahlenfolgen bezeichnen, die nicht notwendigerweise den Standardabweichungen oder Erwartungswerten der Maxima⁵ entsprechen müssen. Der Satz von Fisher und Tippett besagt hier, dass im Falle der Konvergenz die nicht-degenerierten Grenzverteilungsfunktionen bis auf positiv-lineare Transformationen von genau einem der drei folgenden Typen sind (vgl. Embrechts et al., 2001, Theorem 3.2.3):

$$\text{Fréchet – Klasse: } \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad (\alpha > 0)$$

$$\text{Weibull – Klasse: } \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad (\alpha > 0)$$

$$\text{Gumbel – Klasse: } \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathbb{R}.$$

Die Grenzverteilung der normalisierten Maxima im Fall der stetigen Gleichverteilung gehört damit zur Klasse der Weibull-Verteilungen mit $\alpha = 1$ (mit Verschiebung um 1 nach rechts). Man kann die drei Klassen von Verteilungsfunktionen auch durch eine einzige Verteilungsfunktionsklasse (mit nur einem Parameter) charakterisieren, und zwar folgendermaßen:

$$G_\gamma(x) = \exp\left(-(1 + \gamma x)^{-1/\gamma}\right), \quad 1 + \gamma x > 0 \quad (\gamma \in \mathbb{R}),$$

wobei der Sonderfall $\gamma = 0$ als Grenzfall aufzufassen ist (vgl. Beirlant et al., 2004, S. 47). Für $\gamma > 0$ ergibt sich die Fréchet-Klasse, für $\gamma < 0$ die Weibull-Klasse (in beiden Fällen mit $\alpha = 1/\gamma$) und für $\gamma = 0$ die Gumbel-Klasse. Der Parameter $\gamma \in \mathbb{R}$ wird auch als *extremaler Index* bezeichnet.

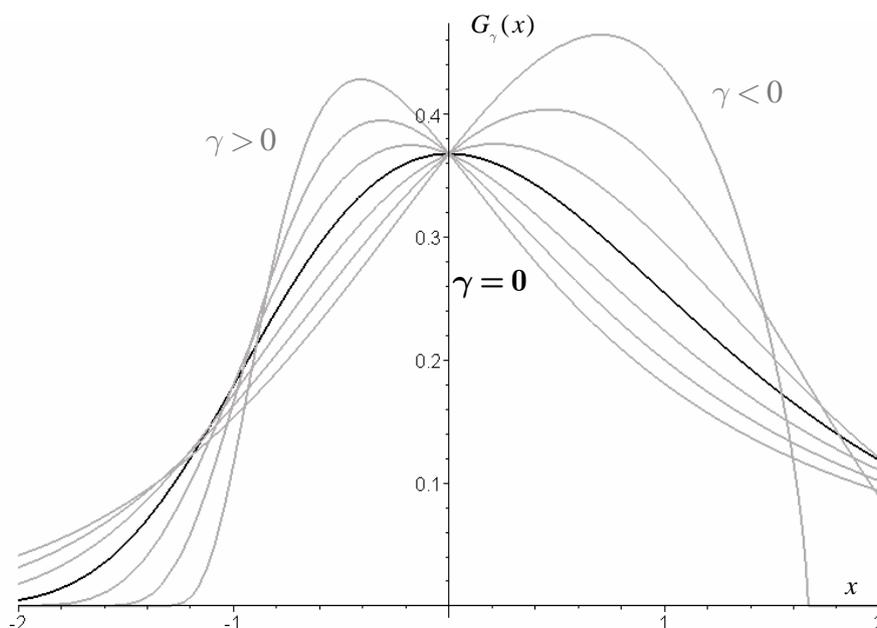


Abbildung 4: Dichten von G_γ für verschiedene Werte von γ

⁵ Solche Folgen kann es z.B. auch dann geben, wenn Erwartungswerte oder Varianzen der Maxima gar nicht existieren.

Auf Grund der herausragenden Bedeutung des *extremalen Index* als charakteristisches Merkmal für die (Grenz-)Verteilungsklasse sind geeignete Schätzer auf der Basis der größten Beobachtungen in der Extremwertstatistik von besonderer Wichtigkeit (vgl. etwa Beirlant et al., 2004, Chapter 4). Ein sehr populärer Schätzer für γ ist der *Hill-Schätzer* (genauer gesagt handelt es sich um eine Klasse von Schätzern), der nur von den k größten Beobachtungen abhängt und folgendermaßen definiert ist:

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln \left(\frac{X_{(n-i+1)}}{X_{(n-k)}} \right) \quad \text{für } k \in \{1, \dots, n\}.$$

Hierbei bezeichnet allgemeiner $X_{(j)}$ den j -größten Wert unter den X_1, \dots, X_n (so genannte *j-te Ordnungsstatistik*). Leider schwankt der Hill-Schätzer mit variierendem k teilweise sehr beträchtlich; es gibt daher sowohl heuristische als auch theoretisch fundierte Empfehlungen, wie k sinnvollerweise zu wählen ist (siehe etwa Reiss/Thomas, 2007, S. 137 oder Beirlant et al., 2004, Chapter 4.7; siehe auch Brüske et al., 2010, Abschnitt 2.1.2).

Für die Modellierung von Naturkatastrophenrisiken, insbesondere im Sturm-Segment, hat sich die Klasse der Fréchet-Verteilungen aus Sicht der Praxis als besonders geeignet herausgestellt (vgl. etwa Pfeifer, 2003). Dies liegt u.a. daran, dass für kleine Werte von α (typischerweise unter 2) der maximale Jahresschaden und der Jahres-Gesamtschaden stochastisch von ähnlicher Größenordnung sind, d.h. der größte Jahresschaden ist so groß, dass alle weiteren Ereignis-Schäden im Verhältnis dazu deutlich abfallen. Extremwertstatistische Analysen historischer Daten auf der Basis solcher Verteilungsmodelle bilden daher schon allein zu Vergleichszwecken eine gute Ergänzung zu den geophysikalischen Modellen.

Ein Anwendungsbeispiel

Wir betrachten hier eine Analyse des Sturmrisikos aus der Sparte VGV für den gesamten deutschen Markt, die in ähnlicher Form auch Grundlage des im Dezember 2005 vom Gesamtverband der deutschen Versicherungswirtschaft (GDV) vorgeschlagenen Solvency II-Standardmodells war (Schubert/Grießmann, 2005). Die ökonomische Bereinigung der Rohdaten erfolgte hier durch Indizierung mit der Vertragszahl und einem Baukosten-Index, um alle Daten auf der Basis des Jahres 2001 vergleichbar zu machen.

Jahr	Zahl der Verträge VGV	Schadenaufwand VGV-Sturm (Mio. €)	Baukosten-Index	Aufwand vertragsindiziert	Aufwand gesamtindiziert	Log Aufwand	u_k	q_k	Log Aufwand sortiert
1975	4.355.103	29,790	922	124,880	273,501	5,611	0,036	-1,198	5,399
1976	4.968.777	243,520	965	894,762	1.872,910	7,535	0,073	-0,964	5,561
1977	5.176.286	61,470	1010	216,804	433,733	6,072	0,109	-0,796	5,611
1978	5.356.209	51,910	1054	176,936	338,974	5,826	0,145	-0,656	5,727
1979	5.608.034	65,221	1108	212,324	387,089	5,959	0,182	-0,533	5,794
1980	5.855.568	65,865	1226	205,356	338,269	5,824	0,218	-0,420	5,824
1981	6.048.049	121,707	1298	367,386	571,697	6,349	0,255	-0,314	5,826
1982	6.222.791	134,757	1336	395,356	597,993	6,394	0,291	-0,211	5,959
1983	6.433.187	269,902	1364	765,953	1.134,579	7,034	0,327	-0,111	5,993
1984	6.599.008	461,018	1397	1.275,444	1.843,708	7,520	0,364	-0,012	5,993
1985	6.790.903	103,567	1403	278,430	400,790	5,993	0,400	0,087	6,072
1986	7.005.770	321,934	1423	838,944	1.191,247	7,083	0,436	0,187	6,136
1987	7.163.400	138,809	1450	353,770	492,974	6,200	0,473	0,289	6,200
1988	7.304.886	135,607	1481	338,915	462,417	6,136	0,509	0,393	6,209
1989	7.439.555	101,664	1535	249,484	328,418	5,794	0,545	0,501	6,222
1990	7.771.017	1.383,590	1633	3.250,512	4.019,857	8,299	0,582	0,613	6,345
1991	8.161.856	100,594	1747	225,012	260,189	5,561	0,618	0,732	6,349
1992	8.654.930	443,899	1859	936,360	1.017,618	6,925	0,655	0,858	6,394
1993	9.147.955	589,156	1950	1.175,786	1.217,744	7,105	0,691	0,995	6,492
1994	12.455.069	477,333	1997	699,677	707,700	6,562	0,727	1,144	6,562
1995	17.125.964	380,342	2044	405,454	400,693	5,993	0,764	1,311	6,925
1996	17.060.594	208,739	2041	223,373	221,129	5,399	0,800	1,500	7,034
1997	17.300.353	478,316	2025	504,757	503,461	6,222	0,836	1,722	7,083
1998	17.661.989	550,315	2018	568,845	569,409	6,345	0,873	1,994	7,105
1999	17.921.302	644,670	2011	656,735	659,740	6,492	0,909	2,351	7,520
2000	18.085.493	492,033	2017	496,691	497,429	6,209	0,945	2,881	7,535
2001	18.256.696	306,958	2020	306,958	306,958	5,727	0,982	3,998	8,299

Abbildung 5: Datengrundlage für die Einschätzung der Sturmgefahr in Deutschland; Quelle: GDV

Die $n = 27$ indizierten Daten x_1, \dots, x_{27} wurden mittels eines Q-Q-Plots an die Fréchet-Verteilung angepasst (vgl. Reiss/Thomas, 2007, Chapter 2 und Beirlant et al., 2004, Chapter 1); dazu wurden die Daten zunächst logarithmiert, wodurch die Fréchet-Verteilungsfamilie in eine Lage-Skalenfamilie vom Gumbel-Typ übergeht. Für die Skalierung der Achsen wurde $u_k = \frac{\text{Jahr}_k - 1974}{27,5}$ und $q_k = -\ln(-\ln u_k)$ für $k = 1, \dots, 27$ gewählt. Der Q-Q-Plot zeigt die Wertepaare $(q_k, \ln(x_{(k)}))$ sowie die durch lineare Regression ermittelte Ausgleichsgerade der Form $y = ax + b$, wobei $x_{(1)} < x_{(2)} < \dots < x_{(27)}$ die der Größe nach geordnete Datenreihe bezeichnet. Die Anpassungsgüte ist hier erstaunlich hoch.

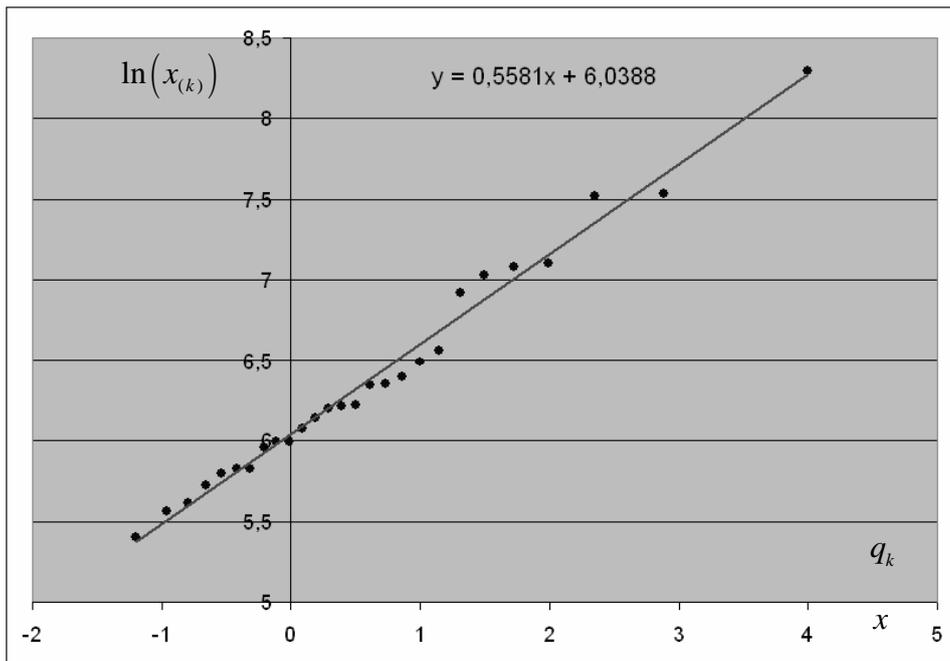


Abbildung 6: Q-Q-Plot der logarithmisch transformierten Jahresschäden des Sturmrisikos Deutschland

Der Schätzer $a = 0,5581$ für die Steigung der Regressionsgeraden ist zugleich ein Schätzer für den *extremalen Index*, wenn auch nicht der bestmögliche (vgl. Beirlant et al., 2004, Chapter 4.3). Er führt zu der Schätzung $\hat{\alpha} = 1/a = 1,7918 < 2$, was zumindest eine gewisse „Gefährlichkeit“ des Sturmrisikos signalisiert. Eine graphische Darstellung der Hill-Schätzer mit variablem k ergibt hier als Kontrast folgendes Bild:

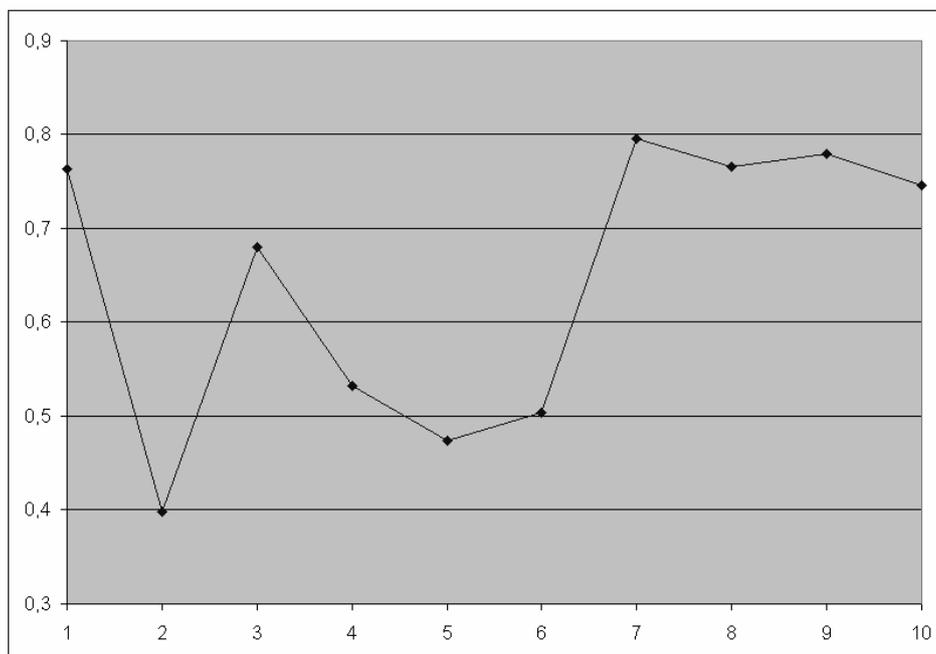


Abbildung 7: Plot der Hill-Schätzer für $k = 1, \dots, 10$

In tabellarischer Form ergeben sich dafür die folgenden Werte:

k	1	2	3	4	5	6	7	8	9	10
$H_{k,n}$	0,764	0,398	0,680	0,532	0,474	0,504	0,795	0,766	0,779	0,746

Der hervorgehobene Wert $H_{k,n} = 0,504$ für $k = 6$ ergibt sich auf Grund einer Variante des minimalen Varianzkriteriums als empfohlener Hill-Schätzer für den *extremalen Index* (Reiss/Thomas, 2007, S. 137). Er liegt etwas unterhalb des Schätzers für den extremalen Index aus dem Q-Q-Plot. Interessant ist hier die Beobachtung, dass der Mittelwert der ersten 6 Hill-Schätzwerte $\frac{1}{6} \sum_{k=1}^6 H_{k,n} = 0,5586$ fast genau mit dem Schätzer aus dem Q-Q-Plot übereinstimmt.

Abschließend soll das Ergebnis der extremwertstatistischen Analyse noch mit drei Berechnungen aus kommerziellen geophysikalischen Modellen verglichen werden. Die aus dem Q-Q-Plot ermittelten Parameter a und b können dazu wie folgt in die *AEP* des stochastischen Modells umgerechnet werden (Fréchet-Typ):

$$AEP_{\text{stochastisch}} = 1 - \exp\left(-e^{b/a} x^{-1/a}\right) = 1 - \exp\left(-\left(\frac{419,39}{x}\right)^{1,792}\right)$$

(x in Mio. €). Für $x = 8000$ (entsprechend 8 Mrd. €) erhält man hieraus eine *AEP* von $\alpha = 0,005$ (Solvency II-Standard) entsprechend einer Wiederkehrperiode von $1/\alpha = 200$ Jahren. [Der GDV hatte in seinem Modell vom Dezember 2005 hierfür einen Wert von 8,2 Mrd. € ausgewiesen.] Die nachfolgende Graphik zeigt diese $AEP_{\text{stochastisch}}$ im Vergleich zu drei *AEP* aus kommerziellen geophysikalischen Modellen.

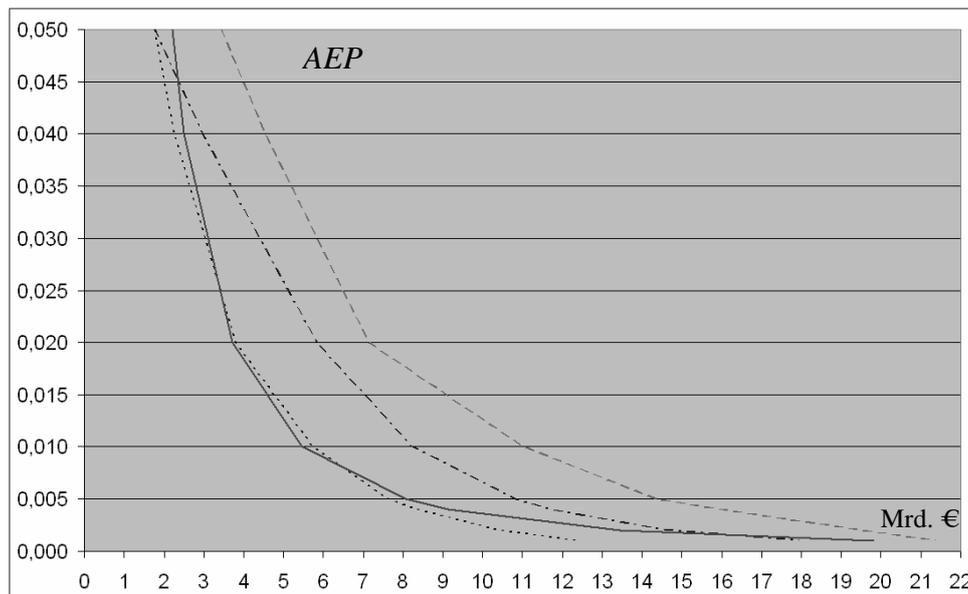


Abbildung 8: Vergleich der *AEP* für die Einschätzung des Sturmrisikos Deutschland nach drei geophysikalischen Modellen (gestrichelt) und dem stochastischen Ansatz (durchgezogene Linie); Quelle: AON Re Services, Hamburg (persönliche Kommunikation)

Die Abweichungen der kommerziellen Modelle untereinander sind doch beträchtlich; der 200-Jahres-Schaden (Solvency II-Standard) schwankt hier von knapp unter 8 Mrd. € bis über 14 Mrd. € Nur eines der kommerziellen Modelle liegt mit der *AEP* annähernd im Bereich der

$AEP_{\text{stochastisch}}$ · Damit dürfte die auf dem stochastischen Ansatz beruhende Einschätzung der Sturmgefahr Deutschland trotz aller Vorbehalte einigermaßen realistisch sein.

Extremwerttheorie und operationale Risiken

Extrem große Schäden treten natürlicherweise nicht nur im Bereich der Naturgefahren auf, sondern in jüngerer Zeit verstärkt auch im Finanzsektor, wo sie im Wesentlichen auf doloses Handeln einzelner Personen zurückgehen (z.B. Nick Leeson und der Zusammenbruch der Barings Bank auf Grund von Spekulationsgeschäften mit Derivaten auf den Nikkei 225 im Jahr 1995 mit einem Verlust von 825 Mio. £, Jérôme Kerviel und die Krise der Société Générale im Jahr 2008 auf Grund von nicht autorisierten Spekulationsgeschäften mit einem Verlust von 4,82 Mrd. € Kweku Adoboli und die Finanzkrise der schweizerischen UBS im Jahr 2011 auf Grund von Spekulationsgeschäften mit Aktienindexfutures mit einem Verlust von 2,3 Mrd. US-\$).⁶ Diese und andere nicht-technische exponierte Risiken werden unter dem Begriff „operationale Risiken“ zusammengefasst, welche insbesondere im Rahmen der aktuellen europäischen Bankenaufsicht (Basel II und III) eine Rolle spielen. Auf Grund ihres Charakters als extrem seltene, aber substanzgefährdende Risiken bieten sich auch hier Analysemethoden auf der Grundlage der statistischen Extremwerttheorie an (King, 2001; Panjer, 2006). Weitergehende Aspekte der Extremwerttheorie im Zusammenhang mit operationalen Risiken findet man z.B. auch in Nešlehová et al., 2006.

Literatur

- Beirlant, J./Goegebeur, Y./
Seger, J./Teugels, J. *Statistics of Extremes. Theory and Applications*. Wiley, Chichester 2004.
- Berz, G. et al. *topics: Jahresrückblick Naturkatastrophen 2002*. 10. Jahrgang, Münchener Rückversicherungsgesellschaft, München 2003.
- Brüske, S./Cottin, C./Hiebing,
A./Hille, B. *Die stochastische Modellierung von Großschäden für den Einsatz in internen Risikomodellen der Schadenversicherung*. Zeitschrift für die gesamte Versicherungswissenschaft, Band 99, Heft 2, 2010, S. 133 - 154.
- Diers, D. *Interne Unternehmensmodelle in der Schaden- und Unfallversicherung*. Entwicklung eines stochastischen internen Modells für die wert- und risikoorientierte Unternehmenssteuerung und für die Anwendung im Rahmen von Solvency II. ifa-Schriftenreihe, Institut für Finanz- und Aktuarwissenschaften, Ulm 2007.
- Embrechts, P./Klüppelberg,
C./Mikosch, T. *Modelling Extremal Events for Insurance and Finance*. Springer Verlag, Berlin 1997.
- Grossi, P./Kunreuther, H.
(Eds.) *Catastrophe Modelling: A New Approach to Managing Risk*. Springer Verlag, New York 2005.
- Kaas, R./Goovaerts, M./
Dhaene, J./Denuit, M. *Modern Actuarial Risk Theory*. Kluwer Academic Publishers, Boston 2001.
- King, J.L. *Operational Risk: Measurement and Modelling*. Wiley, Chichester 2001.
- Klugman, S.A./Panjer, H.H./
Willmot, G.E. *Loss Models. From Data to Decisions*. 2nd Ed., Wiley, Hoboken 2004.
- Nešlehová, J./Embrechts,
P./Chavez-Demoulin, V. *Infinite-mean models and the LDA for operational risk*. Journal of Operational Risk, Vol. I, 2006, S. 3 - 25.

⁶ Quelle: Wikipedia; weitere Beispiele findet man z.B. auch in King, 2001.

- Pfeifer, D. *Solvency II: neue Herausforderungen an Schadenmodellierung und Risikomanagement?* In: *Risikoforschung und Versicherung - Festschrift für Elmar Helten zum 65. Geburtstag*, Verlag Versicherungswirtschaft, Karlsruhe 2004, S. 467 - 481.
- Pfeifer, D. *Möglichkeiten und Grenzen der mathematischen Schadenmodellierung*. Zeitschrift für die gesamte Versicherungswissenschaft, Band 92, Heft 4, 2003, S. 667 - 696.
- Reiss, R.-D./Thomas, M. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3. Aufl., Birkhäuser, Basel 2007.
- Schubert, T./Grießmann, G. *Diskussionsbeitrag für einen Solvency II kompatiblen Standardansatz (Säule I). Modellbeschreibung* (Version 1.0 vom 1.12.2005). Gesamtverband der Deutschen Versicherungswirtschaft e.V., Betriebswirtschaftliches Institut, Berlin 2005
- Weimin Dong *Building a More Profitable Portfolio: Modern Portfolio Theory with Application to Catastrophe Insurance*. Reactions Publishing Group, London 2001.