Modeling, Measuring and Managing Extreme Risk in a Multi-Period Environment

Anna Timonina-Farkas, PhD

Ecole Polytechnique Fédérale de Lausanne (EPFL-RAO, Lausanne, Switzerland)
Schrödinger Fellowship of the Austrian Science Fund (FWF)
Extreme Risks

Under extreme risks we understand uncertain financial, social, environmental etc. outcomes, caused by events with low probability of occurrence, but high consequences (i.e. extreme events).

One would like to model, measure and manage such risks, but faces the following problems:

1. **Lack of data**: there are events that have not yet happened or have happened only very rarely, so relevant data are scarce. Hence, standard statistical methods are generally inapplicable;

2. **Inaccuracy and Inefficiency** of numerical methods for decision-making under uncertainty.

In this presentation, we consider flood risk in Europe as the example of extreme risk.¹

---

Flood Risk in Europe

1. European regions and countries are very much subject to the flood risk;
2. The flood events have direct impact on the financial strength of affected countries.

When we talk about modeling and measuring of flood risk, we do not mean estimation of the average loss caused by a disaster, but the estimation of a probability loss distribution giving us information about size of losses and probabilities not to exceed them.

**Figure:** Loss distribution.
River basins in Europe

There are 796 river basins in Europe subject to flood risk. Solving a decision-making problem under uncertainty at this stage, would require 796 random variables with known distribution functions. The question is, if we can avoid such high-dimensional decision-making problems via studying interdependencies between risks?

Figure: River basins in Europe.
Modeling and Measuring Risk
Extreme Value Theory (EVT)

Let $L_1, \ldots, L_n$ be a sequence of i.i.d. random variables (e.g. flood losses) with unknown distribution function $F$ and let $M_n = \max\{L_1, L_2, \ldots, L_n\}$ denote the maximum.
Then,

$$P(M_n < z) = P(L_1 < z, \ldots, L_n < z) = P(L_1 < z) \cdots P(L_n < z) = (F(z))^n.$$ 

According to the Fisher–Tippett–Gnedenko theorem, the following asymptotic result holds for the distribution function $F$ with $\zeta$ dependent on the tail shape of the distribution:

$$F(z) \sim \exp \left[-(1 + \zeta z)^{-1/\zeta} \right],$$

if there exist sequences of constants $a_n > 0$ and $b_n \in \mathbb{R}$, such that

$$P \left( \frac{M_n - b_n}{a_n} \leq z \right) \rightarrow F(z),$$
as $n \rightarrow \infty$. 

Anna Timonina-Farkas, PhD
Why cannot we just use the EVT to estimate loss distributions on any levels (local, national, regional, world)?

1. Because for this we would need to guarantee that losses $L_1, \ldots, L_n$ are independent, identically distributed random variables!

2. The i.i.d. condition, though, does not hold for any level, except very local one.

To model and measure flood risk for all levels in as much details as possible, one needs to take interdependencies between regions into account.\(^2\)

Regional dependencies for flood events

Hazards typically spread over wider areas. Hence, risk assessment must take into account interrelations between regions. Neglecting such dependencies can lead to a severe underestimation of potential losses, especially for extreme events.

**Figure:** Floods, influencing different regions.
Basic principle in addressing regional dependencies

To estimate total losses after flood events, it is necessary to account for regional interdependencies (i.e. not to neglect influence of an event in one region on other regions). To do so, we use coupling techniques to upscale losses from local to national level.

**Figure:** Coupling losses in a pair of basins.

**Figure:** Coupling of coupled regional losses.
Pair-wise coupling

Consider for simplicity two regions with losses $L_1$ and $L_2$ correspondingly. The total loss in these regions together is, clearly, $L = L_1 + L_2$. Suppose, that marginal densities for the regions are $f_1$ and $f_2$ respectively.

1 **Independent regions:**
   In case the regions and, hence, losses, are independent, the density of the total loss $L_1 + L_2$ can be received by the convolution of marginals:
   
   $$ f(x) = f_1 * f_2 = \int f_1(x - y)f_2(y)dy. $$

2 **Dependent regions:**
   In case the regions and, hence, losses, are dependent, the density of the total loss $L_1 + L_2$ can be received by the convolution over the copula $C(\cdot)$ (with density $c(\cdot)$):
   
   $$ f(x) = f_1 * C f_2 = \int c(F_1(x - y), F_2(y)) f_1(x - y)f_2(y)dy. $$
In our analysis we introduce the Flipped Clayton Copula, which distribution function for parameter \( \theta > 0 \) can be written as

\[
C_{\theta}(u, v) = u + v - 1 + \left[ (1 - u)^{-\theta} + (1 - v)^{-\theta} - 1 \right]^{-\frac{1}{\theta}}
\]

and that satisfies all the necessary properties (1)-(3) and well describes the flood loss behavior.

**Figure:** Flipped Clayton CDF.
Coupling multiple dependent loss distributions

If we consider a set of basins \( I \in \{1, 2, \ldots, N\} \) and estimate the Flipped Clayton Copula between every pair of basins, we receive the following matrix

\[
\Theta = \begin{pmatrix}
\theta_{11} & \cdots & \theta_{1N} \\
\vdots & \ddots & \vdots \\
\theta_{N1} & \cdots & \theta_{NN}
\end{pmatrix} = \begin{pmatrix}
1 & \cdots & \theta_{1N} \\
\vdots & \ddots & \vdots \\
\theta_{N1} & \cdots & 1
\end{pmatrix}.
\]

where pair-wise copula parameters \( \theta_{i,j} \) are estimated via maximization of the log-likelihood function

\[
\ln L(\theta) = \sum_{i=1}^{N} \ln c_{\theta}(u_i, v_i).
\]

Figure: Multivariate Flipped Clayton Copula for some basins in Romania.
Vine Copulas

1. **Multidimensional coupling**: needs estimation of 796! pair-wise copulas;
2. **Hierarchical coupling**: can be conducted via the use of geographical river structure;
3. **Ordered coupling**: minimax approach.

**Figure**: River structure (tree structure).
Hierarchical coupling

For hierarchical coupling we need to estimate copula \( \tilde{C} \) suitable for coupling of copulas \( C_{1,2} \) and \( C_{3,4} \):

\[
C(x_1, x_2, x_3, x_4) = \tilde{C}(C_{1,2}(x_1, x_2), C_{3,4}(x_3, x_4)).
\]

In this case, \( C_{1,2} \), \( C_{3,4} \) and \( \tilde{C} \) may be copulas of different types. But we focus on Flipped Clayton Copulas.

Hierarchical copulas follow a tree structure, but are difficult to estimate, especially if the topology of the tree has to be estimated as well.
Ordered coupling

Suppose, that \( L_1 \) influences \( L_2 \), that influences \( L_3 \), that influences \( L_4 \). Hence, the tree structure is the following:

By ordered coupling, one could estimate 2-dimensional copulas \( c_{1,2}, c_{2,3} \) and \( c_{3,4} \) and combine them in the following way to the 4-dimensional copula density:

\[
c(x_1, x_2, x_3, x_4) = c_{1,2}(x_2|x_1) \cdot c_{2,3}(x_3|x_2) \cdot c_{3,4}(x_4|x_3).
\]
Algorithm for ordering vector $\theta$ out of matrix $\Theta$
Minimax v.s. River structure

To answer the question, if the minimax technique adequately represents the flood losses, we compare two loss curves: one, received based on the minimax approach, and the other, received, based on the geographical tree structure of rivers in Romania.

Figure: Two groups of basins in Romania and their loss curves.
Coupling methods demonstration for the case of Romania

In order to demonstrate flood risk in Romania, we consider

1. Coupling method with Flipped Clayton Copula;
2. Independent convolution;
3. Fully dependent case.

Figure: Comparison of different ordering techniques for Romania.
Why geographical distance is a bad dependency measure? (Romania example)

a. Pair of basins along the Danube.

b. Copula for basins along the Donube.

**Figure:** The dependency is high and is equal to $\theta = 9.9999$, though the distance between the midpoints of the basins is large and equals $d = 242km$. 
Why geographical distance is a bad dependency measure? (Romania example)

- a. Pair of basins in Romania not along the same river.
- b. Copula for basins belonging to different rivers.

**Figure:** The dependency is low and is equal to $\theta = 1.5684$, though the distance between the midpoints of the basins is small and equals $d = 64km$. 
Managing Risk
River basins in Austria

39 basins in Austria; Marginal loss distributions are available for each of these 39 basins (corresponding losses in case of 2-, 5-, 10-,..., 1000- year events); 16 basins with known locations/connections in Austria.

Figure: Basins in EU.

Figure: Basins in Austria.
Risk estimates for Austria for 2030

Probability loss distribution estimates are obtained via coupling method described before, as well as theoretical distributions fit (Fréchet, Weibull and Gumbel).

Figure: Probability loss distribution for Austria for 2030.
Consider a government, which may lose a part of its capital $S_t$ at any future times $t = 1, ..., T$ (i.e. years numbered) because of random natural hazard events with uncertain relative economic loss $\xi_t$ (i.e. $\xi_t \in [0, 1]$). As a result of this loss, the country would face a drop in GDP in the end of the year. Suppose, that under the uncertainty about the amount of loss the decision-maker should decide how much of the available budget $B_{t-1}$ to spend on investment $x_{t-1}$ and on consumption $c_{t-1}$ in absolute terms $\forall t = 1, ..., T$. Suppose also, that an insurance scheme against natural disasters is available for this country and that the decision about the amount of insurance $z_{t-1}$, $\forall t = 1, ..., T$ needs to be made.
Mathematical formulation

Country, exposed to extreme events, may want to solve the following decision-making problem under uncertainty:

$$\max_{x_t, c_t, z_t} \mathbb{E} \left[ (1 - \beta) \sum_{t=0}^{T-1} \rho^{-t} u(c_t) + \beta \rho^{-T} u(S_T) \right]$$

subject to \( S_0 \) (fixed), \( x_t, c_t, z_t \in \mathcal{F}_t, \ t = 0, \ldots, T - 1, \)

\( S_{t+1} = [(1 - \delta) S_t + x_t](1 - \xi_{t+1}) + z_t \xi_{t+1}, \ t = 0, \ldots, T - 1, \)

\( B_t = \alpha S_t = x_t + c_t + \pi(\mathbb{E}(\xi_{t+1})) z_t, \ t = 0, \ldots, T - 1, \)

\( x_t, z_t, c_t \geq 0, \ t = 0, \ldots, T - 1, \)

where \( \alpha, \beta \) and \( \rho \) are given constants; \( \pi(\mathbb{E}(\xi_{t+1})) \) is the insurance premium.

---

Solution methods

There are multiple methods for the solution of multi-stage stochastic optimization problems:

1. **Theoretical solution** (not always possible);
2. **Numerical solution:**
   - via generation of scenario trees\(^4\);
   - via dynamic programming\(^5\).

---


Filtration as a tree

As the stochastic process $\xi$ is given by its continuous probability distribution only, we need to approximate this distribution by discrete one, i.e. to generate points from this distribution. Hence, our aim is

1. To generate points from the given distribution (using quantization algorithms);
2. To solve the multi-stage stochastic optimization program using the generated points.

For these purposes we represent stochastic process $\tilde{\xi} = (\tilde{\xi}_1, ..., \tilde{\xi}_T)$ as a finitely valued tree.
Consider a finitely valued stochastic process $\tilde{\xi}$ that is represented by the tree with the same number of successors $b_t$ for each node at the stage $t$, $\forall t = 1, \ldots, T$. The vector $\text{bush}=(b_1, \ldots, b_T)$ is a bushiness vector of the tree.

**Figure:** Tree with bushiness $b = 2$. 

**Figure:** Tree with bushiness $b = 3$. 

Anna Timonina-Farkas, PhD
Extreme Risk
Kantorovich distance between measures

The Kantorovich distance between measures is defined as

\[ d_{KA}(P, \tilde{P}) = \inf_{\pi} \left\{ \int_{\Omega \times \tilde{\Omega}} d(w, \tilde{w}) \pi[dw, d\tilde{w}] \right\}, \]

\[ \pi[\cdot \times \tilde{\Omega}] = P(\cdot), \]

\[ \pi[\Omega \times \cdot] = \tilde{P}(\cdot). \]

Wasserstein distance: \( d_{WA,r}(P, \tilde{P}) = \inf_{\pi} \left\{ \int_{\Omega \times \tilde{\Omega}} d(w, \tilde{w})^r \pi[dw, d\tilde{w}] \right\}^{\frac{1}{r}} \) under the same constraints.
Nested distance between trees

The multistage distance\(^6\) of two structures \(\mathbb{P} \sim (\Omega, \mathcal{F}, P, \xi)\) and \(\tilde{\mathbb{P}} \sim (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}, \tilde{\xi})\) is

\[
dl_r(\mathbb{P}, \tilde{\mathbb{P}}) = \inf_{\pi} \left( \int d(w, \tilde{w})^r \pi(dw, d\tilde{w}) \right)^{\frac{1}{r}},
\]

\[
dl_1(\mathbb{P}, \tilde{\mathbb{P}}) := dl(\mathbb{P}, \tilde{\mathbb{P}}),
\]

where \(\mathbb{P}, \tilde{\mathbb{P}}\) contain information about filtration, values and probabilities.

\[
\pi[A \times \tilde{\Omega} | \mathcal{F}_t \otimes \tilde{\mathcal{F}}_t](w, \tilde{w}) = P(A | \mathcal{F}_t)(w), \quad (A \in \mathcal{F}_T, 1 \leq t \leq T),
\]

\[
\pi[\Omega \times B | \mathcal{F}_t \otimes \tilde{\mathcal{F}}_t](w, \tilde{w}) = \tilde{P}(B | \tilde{\mathcal{F}}_t)(\tilde{w}), \quad (B \in \tilde{\mathcal{F}}_T, 1 \leq t \leq T).
\]

Nested distance approximation

The fundamental result of G. Pflug and A. Pichler says that

\[ |v(P) - v(\tilde{P})| \leq L_\beta dl(P, \tilde{P}) \leq \sum_{t=1}^{T} d_{KA}(P_t, \tilde{P}_t) \prod_{s=t+1}^{T} (K_s + 1), \]

if the Lipschitz property holds with constants \( K_t \ \forall t = 1, ..., T \).

In order to approximate the distance \( d(P, \tilde{P}) \) between stochastic process and tree \( \tilde{P} \) by the distance \( d(\tilde{P}, P^*) \), we should guarantee that the distance between stochastic process \( \xi \) and its approximation (tree \( P^* \)) is small enough:

\[ dl(P, P^*) \leq \varepsilon. \]

We can guarantee it if we are increasing the bushiness of the tree \( P^* \), because in this case \( dl(P, P^*) \rightarrow 0 \) and \( dl(P^*, \tilde{P}) \rightarrow dl(P, \tilde{P}). \)

\[ ^7 \text{Timonina, Anna. (2013). "Multi-stage Stochastic Optimization: the Distance between Stochastic Scenario Processes." Springer-Verlag Berlin Heidelberg.} \]
Optimal quantization means:

1. to find optimal supporting points $z_i, \ i = 1, \ldots, N$ ($z_1 \leq z_2 \leq \ldots \leq z_N$):

$$\min_{z=(z_1, \ldots, z_N)} \int \min_s d(x, z_s)^r \, dP(x)$$

2. given the supporting points $z_i$, to find the probabilities $p_i$, such that

$$\min_{\tilde{P}} d_{KA}(P, \tilde{P})$$
Stage-wise optimal quantization on a treestructure

The stage-wise optimal tree approximation of the stochastic process \( \xi = (\xi_1, \ldots, \xi_T) \) solves the minimization problem

\[
dl(\mathbb{P}, \hat{\mathbb{P}}) \leq \sum_{t=1}^{T} \min_{\hat{P}_t} d_{KA}(P_t, \hat{P}_t) \prod_{s=t+1}^{T} (K_s + 1)
\]

1. First stage of the tree has \( N_1 \) nodes. We generate \( N_1 \) values of \( \xi_1 \) according to the unconditional probability distribution of \( \xi_1 \);

2. For each of the following stages \( t = 2, \ldots, T \) we generate \( \xi_t \) according to the conditional distribution of \( \xi_t \) given the historical values of the random variables \( \xi^{t-1} \).
Flood risk representation

In the multi-stage case, the risk can be represented in form of scenario trees (where we assumed absence of climate change in this case): Scenarios can be generated either randomly (Monte-Carlo sampling) or optimally (via optimal quantization).

**Figure:** Monte-Carlo sampling.

**Figure:** Optimal quantization.
Optimal value

Optimal value obtained via quantization of scenario trees is:

Figure: Optimal solution for the optimization problem obtained by the Monte-Carlo and stage-wise optimal scenario generation on scenario trees.
Optimal decision

Optimal decision obtained via quantization of scenario trees is:

Figure: Optimal decision of the optimization problem dependent on the insurance load $V$. 
Test of accuracy and efficiency
Inventory Control Problem

A company needs to decide about order quantities \( x = (x_0, x_1, ..., x_{T-1}) \) (where \( x_t \in \mathbb{R}_J^{J} \{0, +\} \), \( \forall t = 0, ..., T - 1 \)) for \( J \) products to satisfy random future demands \( \xi = (\xi_1, \xi_2, ..., \xi_T) \) (where \( \xi_t \in \mathbb{R}_J^{J} \{0, +\} \), \( \forall t = 1, ..., T \)).

The cost for ordering one piece of the good may change over time and is denoted by \( c_{jt-1} \forall t = 1, ..., T, j = 1, ..., J \). Unsold goods may be stored in the inventory with a storage loss \( 1 - l_{jt} \forall t, j \). If the demand exceeds the inventory plus the newly arriving order, the demand has to be fulfilled by rapid orders (delivered immediately), for the price of \( u_{jt} > c_{jt-1} \forall t, j \) per piece. The selling price of the good \( j \) is \( s_{jt} \) (\( s_{jt} > c_{jt-1} \forall t, j \)).
Mathematical formulation

The optimization problem is to maximize the expected cumulative profit:

$$\max \ x_{jt} \geq 0, \ \forall j, t \sum_{t=1}^{T} \sum_{j=1}^{J} \left[ -c_{jt-1}x_{jt-1} - u_{jt}M_{jt} \right] + \sum_{j=1}^{J} l_{jT}K_{jT},$$

subject to

$$l_{jt-1}K_{jt-1} + x_{jt-1} - \xi_{jt} = K_{jt} - M_{jt}, \ \forall t = 1, \ldots, T, \ \forall j = 1, \ldots, J,$$

where index \( j \) corresponds to the good \( j \).

The optimal solution \( x^*_{jt} \) can be computed explicitly and is equal to

$$x^*_{jt} = F^{-1}_{jt} \left( \frac{u_{jt} - c_{jt-1}}{u_{jt} - l_{jt}} \right) - l_{jt-1}K_{jt-1}, \ \forall j = 1, \ldots, J, \ \forall t = 1, \ldots, T,$$

where \( F_{jt}(d) = P_{jt}(\xi_{jt} \leq d) \) is the marginal probability distribution of \( \xi_{jt} \).
Solution methods

To test the quality of the approximation and the efficiency of the algorithm, we use numerical solution methods and compare the approximate results with the theoretical solution of the multi-stage Inventory Control Problem with normally distributed demand.

The approximate solutions are estimated by the use of the following numerical algorithms:

1. Monte-Carlo sampling on scenario trees;
2. Optimal quantization on scenario trees;
3. Dynamic programming with optimal quantizers.
Dynamic programming

According to the Bellman’s principle of optimality, the multi-stage stochastic optimization problems aiming to minimize the expected cumulative loss can be formulated and solved backwards $\forall t = T, \ldots, 0$.

$$V_t(s_t, \xi_t) := \min_{x_t} \left\{ f_t(s_t, x_t, \xi_t) + \mathbb{E}_{P_{t+1}} \left[ V_{t+1}(s_{t+1}, \xi_{t+1}) \mid \xi_t \right] \right\},$$

subject to

- $x_t \in X_t$
- $x_t \triangleq F_t$
- $s_{t+1} = g_t(s_t, x_t, \xi_{t+1})$,

where the endogenous state $s_t$ captures all decision-dependent information. This problem can be solved by the method proposed by Grani Hanasusanto and Daniel Kuhn\(^8\) using historical data directly.

Accuracy

The optimal solution of all numerical algorithms converges to the true solution in probability:

a) Case: 1 product; 2 stages.

b) Case: 3 products; 2 stages.

Figure: Accuracy comparison of numerical algorithms for solution of multi-stage Inventory Control Problem.
Efficiency

The efficiency of numerical algorithms depends on the number of iterations necessary for the convergence and on the dimension of the problem:

**Figure**: Efficiency comparison of numerical algorithms for solution of multi-stage Inventory Control Problem.
Thank you for your attention!


