International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

The Design of Optimal Insurance Decisions in the Presence of Catastrophic Risks

Ermolieva, T.Y.

IIASA Interim Report
October 1997

IIASA
International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-97-068/October

The Design of Optimal Insurance Decisions in the Presence of Catastrophic Risks

Tatiana Ermolieva (ermol@iiasa.ac.at)

Approved by

Joanne Linnerooth-Bayer (bayer@iiasa.ac.at)
Co-Leader, *Risk, Modeling and Policy*

Contents

Abstract

This paper deals with the development of decision making tools for managing catastrophic (low probability – high consequences) risks. Catastrophes produce rare and highly correlated claims, which depend on various decision variables, i.e. coverages at different locations, mitigation measures and reinsurance agreements. Joint probability distributions of these claims depicting their complex spatial and temporal interactions and effects of decision variables are analytically intractable. Spatial stochastic models of catastrophes can bypass these difficulties. Catastrophic models combine the simulation of realistic and geographically explicit catastrophic events with the differentiation of property values and insurance coverages in different locations of the region. Catastrophic models can be combined with stochastic optimization techniques to aid decision making on the spatial diversification of contracts, insurance premiums, reinsurance requirements, effects of mitigation measures, and the use of other financial mechanisms. The aim of this paper is to extend a two-stage spatial catastrophic model to dynamic cases reflecting dependencies of risk accumulation processes in time. This extension is important since it can be used for the analysis of decisions under changing frequencies of events and values of properties. It is also possible to incorporate catastrophes caused by the clustering in time of such events as rains and droughts due to persistence in climate. The model can be used by individual insurers, pools of insurers or regulatory authorities.

Key words: Catastrophes, Insurance, Decisions under uncertainty, Risk, Stochastic optimization, Adaptive Monte Carlo method.

Acknowledgments

The author would like to thank Professor Georg Pflug from IIASA's Risk, Modeling and Policy (RMP) project for his academic guidance during the YSSP studies. Special thanks go to the Director of IIASA, Professor Gordon MacDonald, and the staff members of the RMP project for their recommendations, scientific advice and help. I am thankful to Aniello Amendola, Richard Cumpston and Joanne Linnerooth-Bayer for numerous comments on the draft of this paper. The study was supported by the Austrian Academy of Sciences.

About the Author

Tatiana Y. Ermolieva is a Ph.D. student at the Vienna University, Department of Economic Sciences, Institute of Statistics and Operation Research. She was a participant in IIASA’s Young Scientists Summer Program 1997. Together with Yuri Ermoliev and Vladimir Norkin, she shared the Kjell Gunnarson's Risk Management Prize of the Swedish Insurance Society, June 18, 1997, for a paper titled "Spatial Stochastic Model for Optimization Capacity of Insurance Networks under dependent Catastrophic Risks".

The Design of Optimal Insurance Decisions in the Presence of Catastrophic Risks

Tatiana Ermolieva

1. Introduction

The concentration of property values and population in certain regions combined with the introduction of new technologies in different sectors of the economy imposes risks to the public and environment. Possible climate changes [29] may also increase the exposure of society to human-made and natural disasters [25]. Natural hazards alone cost in 1995 about $150 billion.

Hurricane Andrew in the U.S., for instance, is estimated to have caused $20 billion of insured loss, and is the most costly natural disaster in the history of the insurance industry. Insurers such as State Farm and Allstate suffered losses from Andrew of $3.5 and $2.5 billions respectively [18]. This summer also showed that such events as rain clustered in time in the same region may produce high losses.

Human-made catastrophes [25] are also of great concern. The meltdown of the atomic power plant in Chernobyl, the explosion of a chemical tank in Bhopal, and oil spills from tanker crashes, as well as other technological catastrophes may have cost even more then natural catastrophes.

Insurance is a mechanism for the financial protection against different kinds of disasters. Insurers are currently concerned with the possibility of claims even higher than already experienced [6]. Traditional insurance operates with well-defined cases. For example, automobile and life insurance are types of insurance where decisions on premiums, estimates of insolvency and possible losses are calculated using rich data bases collected over long periods.

The principal problem in insuring catastrophic risks is insufficient historical data for predicting events at any particular location, although rich data may exist on their occurrence and magnitude on an aggregated (say regional) level. Potential damages in a particular location may be unlike anything that has been experienced in the past. Catastrophes produce highly correlated damages and claims, which depend on the region of occurrence, coverages at different locations, mitigation measures, reinsurance agreements and so on.

The lack of data and the complex spatial and dynamic interdependencies make it dangerous to use purely adaptive "trial-and-error" approaches. For this reason, models can be useful for specifying the implicit dependencies and for predicting possible damages and losses. Models can be used to study company solvency, decisions on insurance premiums, reinsurance requirements, effects of mitigation measures, and the diversification of contracts (see [12], [14]). The occurrence of various catastrophic events in a region can be simulated on a computer in the same way as it might happen in reality. For tracking dependencies between all possible damages the model has to be geographically explicit, allowing for geographical representation of catastrophic patterns in space and time, distribution of property values and insurance contracts.

The aim of this paper is to further the development of spatial stochastic catastrophic models. The dynamic version of a two-stage model (Ermolieva, Ermoliev, Norkin [12]) is introduced together with stochastic optimization procedures for improving the geographical diversification of insurance contracts, stabilizing the insurance business, increasing insurance profits, and providing financial protection of the population. In the general case, dependencies between possible claims have a complex character defined by spatial patterns of events and feasible policy variables. The spatial dynamic stochastic model tracks these dependencies with Monte-Carlo simulations, and a stochastic optimization procedure sequentially adjusts the decision variables without exact evaluation of all the risks associated with the infinite combinations of feasible policy variables. Section 2 overviews the classical risk models, premium calculation and estimation of insolvency. Section 3 briefly discusses Borch's classical model for the optimal diversification of risks. The crucial limitation of the model is the assumption on the substitutability of risks. In Section 4, a dynamic spatial stochastic model is proposed for the optimal diversification of dependent nonsubstitutable risks. Section 5 describes the implemented adaptive Monte Carlo methods based on stochastic optimization techniques. Numerical experiments in Section 6 demonstrate how adaptive Monte Carlo methods may easily "learn" about dependencies among damages and "propose" that insurers either reduce risks in some locations or (and) take more catastrophic risks from other locations to stabilize their business. Section 7 presents some concluding remarks related to the development and use of the catastrophic model.

2. Classical Insurance Model, Insurability of Risk

Insurance, a mechanism for reducing financial risk and spreading financial loss, is a major social institution that is essential to the functioning of many industrialized economies. Historically, insurance dates back at least as far as the Romans, whose burial clubs financed funeral expenses and made payments to families of the deceased. In the United States, where one active company dates from before the Revolutionary War, some 6000 insurance companies collect well in excess of $200 billion in annual premiums, employ more than 2 million people, and hold assets valued at close to $800 billion.

Traditionally insurance companies deal only with what is called "pure risk", which has to satisfy certain conditions [7]:

1) The risk must be predictable. That means there should exist sufficient data to permit actuaries to predict the number and average size of insured losses for a given period.

2) Each risk must be measurable.

3) The premium charged on the risk must be low enough to attract a sufficient number of insured people, yet high enough to support the numbers of probable losses.

4) The risk must be free of any potential catastrophe that could produce loss in excess of the ability of the insurer to respond.

5) Homogeneous units must be independently exposed to loss. That is, a loss of one should not lead to a loss of another.

The existing insurance risk theory gives reliable results for dealing with such risks. Though the theory is not perfect it deepens the intuition and helps to understand insurance as a complex dynamic process [2].

2.1. Risk Reserves

For each insurance company the main variable of concern is its risk reserve at time t or in other words the money which a company has at its disposal. In general form, risk reserve is calculated as

$$R(t) = R_0 + P(t) - S(t), \quad t > 0\,, \tag{1.1}$$

where $P(t)$ is aggregated premiums on $[0,t)$, $S(t)$ is aggregated claims, and R_0 is the initial risk reserve. A trajectory of a risk reserve process is shown in Fig.1 (section 2.2). At time moments τ_i , $i = 1,2,...$ claims pushes it down, whereas premiums push it up.

Aggregated claims $S(t)$ are also called in insurance risk theory the aggregated claim size process. It depends on the number of claims and their sizes. Claim number process is usually characterized by a probability $p(k,t) = prob(N_k(t) = k)$ that the number of claims $N_k(t)$ up to time t is equal to k. A very often proposed model for $p(k,t)$ is the Poisson law

$$p(k,t) = e^{-\rho t}\frac{(\rho t)^k}{k!}\,,$$

where ρ is a parameter indicating the average number of claims in a time unit. Claim size up to time t is

$$S(t) = \sum_{i=1}^{N_k(t)} S_i\,,$$

where $N_k(t)$ is a random variable of claim numbers up to time t, and S_i is the claim size at time i. If we assume that $F_1(X) = prob(S_1 \leq X)$ is a distribution function (d.f.) of a single claim size, $F_k(X) = prob(S_k \leq X)$ is a d.f. of the sum of exactly k mutually independent claims each of them distributed according to $F_1(X) = prob(S_1 \leq X)$. Then d.f. of the sum of k claims $S_k(t)$ is

$$F_k(X) = \int_0^x F_{k-1}(X - Z)dF_1(Z) = F_1^{k*}(X),$$

which is called a k-th convolution of the d.f. $F_1(X)$. Therefore the distribution function of the aggregated claim $S(t)$

$$F_t(X) = \sum_{k=0}^{\infty} p_k F_k(X) = \sum_{k=0}^{\infty} p(k,t) F_1^{k*}(X),$$

where $p(k,t)$ is the probability of k claims up to time t. The distribution function $F_t(X)$ is called a compound distribution function. This simple formula shows the difficulties of deriving tractable analytical formulas for the distribution, $F_t(X)$, even for simple cases with only one insurer. In more general cases, the distribution of claim size, S_k, $k = 1,2,...$ depends on reinsurance policy variables and applicable mitigation measures, leading to additional difficulties. In these cases the development of computational approaches is crucially important for the practical applications of mathematical models. One approach is concerned with analytical approximation of complex probability distributions. The most important approach is based on the use of Monte Carlo Methods [15].

The choice of distributions approximating claim sizes with possible catastrophic volumes is approached in the following way. Large claims are rare events, having a low probability of occurrence concentrated in the tails of distributions. It is important not to underestimate these tails, but to consider them separately from the main part of the distribution. The distribution of the claim size may be a composition of two or more weighted distribution functions, each taken for a particular interval. For example, a distribution function $G_1(X)$ may represent volumes of claims below or equal to some predefined level Z_0, and $G_2(X)$ is a distribution function of catastrophic claims with claim size exceeding Z_0. Distribution $G_2(X)$ is often approximated by the Pareto law

$$\Pr ob(S \le X) = G_2(X) = 1 - (Z_0 / X)^{\alpha}, \ X \ge Z_0, \ \alpha > 0,$$

where Z_0 is the smallest claim considered as catastrophic. If the risk index α is less then 2, the distribution can be characterized as heavy tailed. The Pareto distribution has the following property that is convenient in modeling large claims

$$\Pr ob(S \ge X \mid S \ge Z_0) = \frac{\Pr ob(S \ge X)}{\Pr ob(S \ge Z_0)} = \left(\frac{Z_0}{X}\right)^{\alpha}, \quad X > Z_0.$$

The two parametric Pareto distribution function

$$\Pr ob(S < X) = 1 - b\left[1 + \left(\frac{X - Z_0}{Z_0}\right)^{\beta}\right]^{-\alpha}, \quad X \ge Z_0,$$

where α and β are positive parameters, Z_0 is the limit for the tail for which the formula is fitted, b indicates the weight of probability mass concentrated in the tail area $S \ge Z_0$, that is $b = 1 - \Pr ob(S < Z_0)$. Often $G_2(X)$ is also represented by Weibull distribution

$$\Pr ob(S < X) = 1 - \exp\left\{-\left[(S - Z_0)/a\right]^{b}\right\},$$

where a, b are distribution parameters.

Another existing approach to model catastrophic claim size is to use extreme value distributions connected with the modeling of maximum-magnitude events ([4]), i.e. when a catastrophe is considered as one rare event with a high consequence.

Catastrophes produce dependent damages at different locations. If the insurer has coverages in these locations, then the distribution of aggregated claims depends on existing coverages of insurer and the geographical pattern of catastrophic events. The use of joint distributions may be rather cumbersome for this task since the catastrophes may have rather complicated patterns. To bypass these difficulties it is possible to use claim-generating stochastic processes (1.1) instead of the probability distribution $F(X)$, which leads to Monte Carlo methods.

The aggregated premium $P(t)$ significantly influences the whole insurance business and should strictly reflect the distribution of claim size. It is clear that the distribution of damaged values without insurance should not be better (in a sense) than the distribution of damaged values plus the difference between coverage by insurance damages and paid premiums. The meaning "better distribution" is discussed in sections 4, 5. In general we can say that the volume of premium depends on the distribution $F(X)$ of accumulated claims. If $F(X)$ is a distribution function of accumulated claims from a single risk, then $P(F(\cdot))$ is called a premium; if $F(X)$ is a d.f. of collective risks, then $P(F(\cdot))$ is called a collective premium.

Actuaries use known basic principles for the calculation of premiums [7], [23]. According to the *equivalence principle* premiums are usually calculated relying on the mean value of aggregated claims increased by the so-called safety loading. For the *expected value principle*

$$P(F(\cdot)) = (1+\lambda)\mathrm{E}L,$$

where $\lambda > 0$ is the safety loading, reflecting possible fluctuations of the risk process and uncertainties in the loss distribution. In practice expected value $\mathrm{E}L$ of losses L according to the law of large numbers is substituted by observable average loss

$$\overline{L}_N = \frac{1}{N}\sum_{k=1}^{N} L_k .$$

For the *standard deviation principle*

$$P(F(\cdot)) = \mathrm{E}L + \alpha\sigma(X),$$

where $\alpha > 0$. The *variance principle* requires

$$P(F(\cdot)) = \mathrm{E}L + \beta\sigma^2(X),$$

where $\sigma^2(X) = E((X - E(X))^2)$, and $\beta > 0$.

The expected value principle is almost always used in life insurance, and in contrast, it is only seldom used in property and casualty insurance. The standard deviation principle is probably the most frequently used approach in property and casualty insurance. The variance principle is not so popular as the standard deviation principle. In the case of heavy tailed distributions, premiums may be calculated not only on the bases of the first moments, but also higher moments of distribution F may be required.

The choice of α, β, λ depends individually on each type of risk and each particular company. The levels of α, β, λ and, therefore, the levels of premiums in the case of large losses should ensure the desired probability of survival for each insurer [2], [7]. In the case of catastrophic losses it becomes extremely difficult to make decisions on premiums. They often may not suffice to cover losses of insurers, and the need for reinsurance and other financial mechanisms and regulations becomes obvious.

2.2. Long Term Stability of Insurers

The long term stability of the insurer depends on the type of coverages, the distribution of claims, the volumes of premiums, reinsurance contracts, and the mitigation measures. It is defined by the risk reserve $R(t)$, a complex jumping stochastic process. A random trajectory of this process is shown in Fig.1 for $P(t) = ct,\ c > 0$.

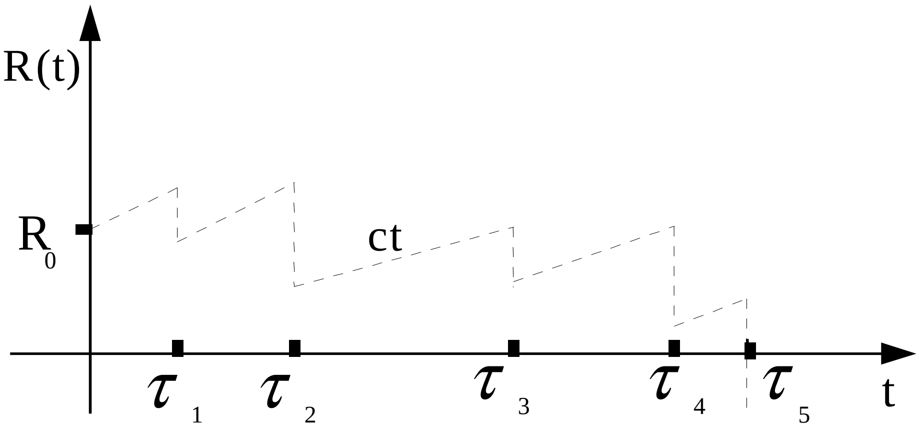

Figure1. A sample trajectory

As we can see the timing of claims and their sizes cause the ruin at τ_5. The long-term stability of $R(t)$ can be characterized by the probability of ruin (insolvency)

$$q(R_0, c) = \Pr\{R(t) \le 0 \ \text{for some}\, t > 0\}. \qquad (2.1)$$

An important problem of an optimal insurance policy is the choice of premium c and initial risk reserve R_0 which guarantee a given level of insolvency ($q(R_0, c) \le \textit{given level}$) and maximize profit within the feasible demand for insurance. In Sections 4, 5 we discuss the general problems on the optimal choice of contracts by carefully selected coverages from different geographical locations. Let us outline here the general methodological challenges.

An analytical formula for q is available only in the simple cases, for simple distributions of claim processes and claim sizes (see [2], [7]). The Monte Carlo methods were developed for the study of complex, stochastic processes where analytical approaches fail. It is important that these methods avoid the use of integrodifferential equations governing the change of the probability distribution of the stochastic processes.

The direct computer simulation of accumulated risk reserve processes $R(t)$ can easily be done for any given decision variables such as r, λ and for a large enough interval $[0, T]$. An example of a possible random trajectory is shown in Fig.1. The straightforward estimation of q can be based on the identity:

$$q(r, \lambda) = EI(R(\tau)), \qquad (2.2)$$

where $I(y) = 1$, if $y \le 0; I(y) = 0$, if $y > 0$, and τ is the random stopping time $\tau = \inf\{t : R(t) \le 0, t \le T\}$. The function $I(R(\tau))$ indicates ruin, i.e. it is equal to 1 in the case of ruin and 0 otherwise. Unfortunately, the consistent estimation of $q(r, \lambda)$ may be time consuming, especially when low probability/high consequence events play an essential role.

The first problem is to develop fast Monte Carlo estimation procedure using importance sampling and possible analytical transformation of the model [24]. The second problem is the search for decision variables, r, λ, which guarantee a desirable performance, for example, a given level of ruin probability with minimal λ and fixed r. Large λ-s increase premiums and decrease the demand for coverages. The minimization of λ in this case implicitly takes this into account and avoids more complicated models. The straightforward application of the Monte Carlo method for each combination of desirable policy variables is impossible, since the number of such combinations is equal to infinity. Let us now demonstrate the advantages of adaptive Monte Carlo methods and fast estimation procedures.

2.3. Stochastic Optimization Procedure

Let us distinguish between two parts of the risk portfolio: “normal”, associated with ordinary, independent claims, and “catastrophic”, associated with catastrophic risks. Consider a discrete time interval, $t = 0,1,...,T-1$, and assume that at time $t \ge 0$

the “normal” part is characterized by a random variable, M_t, accumulated premiums from catastrophic risks are xt, where x is a desirable policy variable. If the probability of a catastrophic event at t is $\underline{p} \le p \le \overline{p}$, then the probability of ruin is defined as the expectation

$$q(x) = E\sum_{t=1}^{T} p(1-p)^{t-1} I(M_t + xt - S_t < 0),$$

where S_t is the catastrophic claim generated at time t. Assume that the probability distribution $V_t(z) = \Pr[M_t < z]$ can be evaluated. Then it is possible to reduce the variance of this estimator by taking the conditional expectation with respect to M_t:

$$q(x) = E\sum_{t=1}^{T} p(1-p)^{t-1} V_t(S_t - xt). \qquad (2.3)$$

This simple formula provides faster estimates of $q(x)$ than formula (2.2). Assume that the goal is to choose an x that guarantees a given level of stability:

$$q(x) = \gamma,\ \gamma > 0,$$

which also can be achieved by maximizing the function

$$F(x) = \int_0^x q(\alpha)d\alpha - \gamma x.$$

The stochastic optimization procedure starts with a given initial combination of policy variables. In this case it is only the value of premium x^0. Let us denote x^k as the value of the premium after k simulations. Step $k+1$: choose t_k with probability

$1/T$ from the set $\{1,2,...,T\}$, generate $p_k \in [\underline{p}, \overline{p}]$ and the claim $S_{t_k}^k$. Adjust the current value x^k according to the feedback:

$$x^{k+1} = \max\left\{0, x^k + \frac{\rho}{k+1}[Tp(1-p)^{t_k-1}V_{t_k}(S_{t_k}^k - x^k t^k) - \gamma]\right\},$$

where ρ is a positive constant. The value x^k converges to the desired value of premium such that $q(x) = \gamma$. This follows from the fact that the term $Tp(1-p)^{t_k-1}V_{t_k}(S_{t_k}^k - x^k t^k)$ is an estimate of $q(x)$ given by (2.3). We develop this type of approach for the general problem with many insurers and complex dependent claim processes in sections 4, 5.

3. Optimal Diversification

The surplus of the insurance industry is potentially enough to pay for losses from catastrophic events. However (see Cummins, Doherty [6]) in practice the available capacity of reinsurers is very limited and, depending on the spread of coverages, many insurers could become insolvent in the case of large catastrophes. Cummins and Doherty analyzed the capacity of the insurance industry to respond to catastrophic events assuming that the industry acts as a single firm. This analyses critically rests on the results of Borch [3] for optimal arrangements of a reinsurance market, which are valid only for substitutable risks. The analysis is, therefore, not applicable to the general problems of sections 4, 5.

3.1. Borch's Model: Substitutable Risks

The model deals with optimal redistribution of risks which companies have accepted by their direct underwriting. In the initial situation company i ($i = 1,2,...,n$) is committed to pay x_i, the total amount of claims which occur in its own portfolio. The company also has the initial reserve of R_i^0, which is available to pay the commitment. Thus the initial risk situation of company, i, is characterized by the random variable, $R_i^0 - x_i$. Reinsurance contracts redistribute the initial commitments, x_i, and change the probability distribution of the risk reserves. The new commitments of companies can be characterized by a set of nonnegative functions $y_i(x_1, x_2,..., x_n)$, $i = 1,2,...n$, where $y_i(x_1, x_2,..., x_n)$ is the amount company i has to pay if claims in the respective portfolios amounts to $x_1, x_2,..., x_n$. It is assumed that companies act as a single company and all risks (claims) are substitutable, i.e. new commitments are constrained only by the aggregated claim $\sum_{i=1}^{n} x_i$:

$$\sum_{i=1}^{n} y_i(x_1, x_2,..., x_n) = \sum_{i=1}^{n} x_i \qquad (3.1)$$

and new risks of companies are characterized by $R_i^0 - y_i(x)$ with the same R_i^0.

Thus, reinsurance contracts $y_i\,(x)$ change the initial risk reserve of company i from the random variable $R_i^0 - x_i$ to $R_i^0 - y_i(x)$. What is the optimal redistribution of $y(x)$? How can we compare random outcomes (variables)? In the general case random outcomes are characterized by probability distributions and other indicators such as average costs, profits, moments of (probability, cost, profit) distributions. An ordering among random variables can be achieved in a variety of ways depending on the problem at hand.

Assume that company i attaches an expected utility

$$U_i(y) = \int_{R_+^n} u_i(R_i^0 - y_i(x))dH(x) = E_x u_i(R_i^0 - y_i(x)) \qquad (3.2)$$

to the risk situation $R_i^0 - y_i(x)$, where $u_i(\cdot)$ is continuous function with decreasing positive derivatives, $H(x)$ is the joint distribution of $x = (x_1, x_2,..., x_n)$, and R_+^n stands for the positive orthant in the n-dimensional Euclidean space. A Pareto optimal set of redistributions $y_i\,(x)$, $i = 1,2,..., n$ is achieved when there is no other set of contracts $\overline{y}_i(x)$ such that $u_i(y_i) \le u_i(\overline{y}_i)$, $i = 1,...,n$, with at least one strict inequality.

Borch proved that *for any Pareto optimal set of redistributions* $\{y_i(x)\}$ *the amount* $y_i(x)$ *which company* i *has to pay will depend only on the total amount of claims* $\sum_{i=1}^{n} x_i$ *made against the insurance industry and functions;* $y_i(x)$, $i = 1,2,...,n$ *satisfy the relations*

$$u_i'(R_i^0 - y_i(x)) = k_i u_1'(R_1^0 - y_1(x)) \qquad (3.3)$$

where $k_1, k_2, ..., k_n$ *are positive arbitrary constants.*

A rigorous statement of this proof is lengthy and rather tedious. The elementary proof is derived from the following construction, which is used further. Any Pareto optimal vector $y(x) = (y_1(x),..., y_n(x))$ is achieved by maximizing

$$\sum_{i=1}^{n} v_i E_x u_i (R_i^0 - y_i(x))$$

with positive weights v_i. Since $y_i(x)$ is an arbitrary function of x, then

$$\max\left\{\sum_{i=1}^{n} v_i E_x u_i (R_i^0 - y_i(x)) : \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i\right\} =$$
$$= E_x\left\{\max\left\{\sum_{i=1}^{n} v_i u_i (R_i^0 - y_i(x)) : \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i\right\}\right\},$$

i.e. for each given $x = (x_1, x_2, ..., x_n)$ a Pareto optimal redistribution $y(x)$ is an optimal solution of the simple problem:

$$\text{maximize } \sum_{i=1}^{n} v_i U_i (R_i^0 - y_i(x)), \text{ subject to } \sum_{i=1}^{n} y_i(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i .$$

Hence, a Pareto optimal solution $y_i(x)$ depends only on the $\sum_{i=1}^{n} x_i$ and there exists a constant λ such that

$$v_i u_i'(R_i^0 - y_i(x)) = \lambda ,\ i = 1,2,...,n ,$$

or

$$v_i u_i'(R_i^0 - y_i(x)) = v_1 u_1'(R_1^0 - y_1(x))$$

for $i = 1,2,...,n$, which is equivalent (3.3) for $k_i = v_1 / v_i$, $i = 1,2,...,n$.

3.2. Measuring the Capacity of an Insurance Market

Cummins and Doherty [6] use Borch's results for measuring the capacity of an insurance market. Consider an insurance market with insurers $i = 1,2,...,n$. The risk reserve of a company i can be represented in a simple two-stage model as $R_i = \max\{0, R_i^0 + P_i - x_i\}$, where x_i is a total amount of claims, P_i is the premium income from x_i and R_i^0 is the initial reserve or the fund. The industry's surplus after a catastrophe x_i is defined as

$$\sum_{i=1}^{n} R_i = \sum_{i=1}^{n} \max\{0, R_i^0 + P_i - x_i\}.$$

The problem is to maximize the average industry surplus

$$F(x) = \sum_{i=1}^{n} E \max\{0, R_i^0 + P_i - y_i(x)\} \qquad (3.4)$$

subject to constraints

$$\sum_{i=1}^{n} y_i(x_1, x_2,..., x_n) = \sum_{i=1}^{n} x_i .$$

Let us note that (3.4) can be written in the form of (3.2) with a convex utility function $u_i(y) = \max\{0, R_i^0 + P_i - y_i\}$. Assuming that Borch's results are valid and using the assumption that x_i has a normal distribution, Cummins and Doherty analyze the case when the optimal $y_i(x)$ is necessarily proportional to the aggregated industry losses $\sum_{i=1}^{n} x_i$.

These results crucially depend on the assumption (3.1), that different claims associated with different companies are not distinguishable. It is assumed that companies behave as a single company, i.e. claims x_i of all companies are mixed up in one aggregated claim $\sum_{i=1}^{n} x_i$. A key assumption of Borch's model is also that the aggregated claim $\sum_{i=1}^{n} x_i$ is redistributed between companies without redistributing the initial fund R_i^0. Thus a company dealing with risky contracts and receiving high premiums may have less risky new commitments with the same high incomes. The following example illustrates the limitations of these assumptions in the case of more realistic problems.

Assume that catastrophes may occur independently in locations $l = 1,2$ with probabilities p_1, p_2. In the initial state company 1 covers $x_1 = 4$ units of property from the location 1; company 2 covers $x_2 = 2$ units of property from location 2. Premiums $\pi_1 = 1/2$, $\pi_2 = 1/3$ per unit of coverage; $R_1^0 = 4$, $R_2^0 = 4$. Assume also that the catastrophes entirely damage the property at the locations. In this case the aggregated claim

$$\sum_1^2 x_i = \begin{cases} x_1 = 4 & \text{with prob.} = p_1(1 - p_2), \\ x_2 = 2 & \text{with prob.} = p_2(1 - p_1), \\ x_1 + x_2 = 6 & \text{with prob.} = p_1 p_2, \\ 0 & \text{with prob.} = (1 - \mathrm{p}_1)(1 - p_2). \end{cases}$$

If catastrophes occur in both locations then the industry's surplus is

$$(R_1^0 + \pi_1 x_1 - x_1) + (R_2^0 + \pi_2 x_2 - x_2),$$

and the aggregated claim $x_1 + x_2 = 6$. Since $\pi_1 > \pi_2$, the optimal redistribution $y = (y_1, y_2)$, $y_1 + y_2 = 6$, $y_1 \geq 0$, $y_2 \geq 0$ is achieved according to (3.3) for $y_1 = 6$ and $y_2 = 0$. This conclusion is not correct, since there is only 4 units of risk with premium π_1. Thus claims x_1, x_2 cannot be aggregated, i.e. a constraint on the aggregated claim $y_1 + y_2 = 6$ must be substituted by two constraints on available amount of claims from each location:

$$y_{11} + y_{21} = 4, \; y_{12} + y_{22} = 2,$$

where y_{ij} is the coverage of company i in location j. In the next section we propose this type of model to deal with the more general case.

4. Spatial Dynamic Model of Stochastic Optimization

The models of sections 2, 3 have a rather simplified illustrative character. In reality damages and claims depend on geographical patterns of catastrophes, clustering of property values in the region, available mitigation measures and regulations, and the spread of insurance coverages among different locations. Catastrophes produce highly correlated claims from different locations affected by the same event. For all these reasons, the model should be geographically explicit (see [12]) for the description of property values and insurance contracts in different parts of the region, and for explicit modeling of catastrophes.

Although still limited in its use, catastrophic modeling (see [14]) is becoming increasingly important to insurance companies for making decisions on the allocation and values of contracts, premiums, reinsurance arrangements, and effects of mitigation measures. For any given combination of an insurer's decision variables it is possible to simulate different patterns of catastrophes in a region as they may happen in reality and analyze their impacts on the stability of the companies or the industry. Such models compensate for the lack of historical data on the occurrence of catastrophes in locations where the effects of catastrophes may have never been experienced in the past. Different catastrophic scenarios lead to (in general) different "optimal" decision strategies. The important question is how we can find a decision strategy, which is the "best" against all possible catastrophes. In paper [12] it was shown that the search of "robust" optimal decisions can be done by incorporating stochastic optimization techniques into catastrophic modeling. By using this approach it is possible to take into account complex interdependencies between damages at different locations, available decisions and resulting losses and claims. In this section the spatial two-stage model [12] is extended to dynamic cases.

4.1. Flows and Stocks of Risk Reserves

Similar to [12] the study region is subdivided into subregions (compartments) or locations $j = 1,2,...,m$. Locations may correspond to a collection of households, a zone with similar seismic activity, to a watershed, etc. They may also be identified with the collection of grid cells for meaningful representation of the simulated patterns of events in space and time. We assume that for each location j there exists an estimation W_j of the property value or "wealth" of this location, that includes values of houses, lands, factories, etc.

Suppose that n insurance companies $i = 1,...,n$ have contracts in all locations and partially cover their losses. Each company i has initial funds or a risk reserve R_i^0, which in general is characterized by a random variable dependent on catastrophic events. Assume that time span consists of $t = 0,1...,T-1$ time intervals. In general the risk reserve R_i^t of the company i is calculated according to the following formula for $t = 0,1,...,T-1$:

$$R_i^{t+1} = R_i^t + M_i^t + \sum_{j=1}^{m} \left[\pi_{ij}^t (q^t) - c_{ij}^t (q^t) \right] - \sum_{j \in \varepsilon_t(\omega_t)} L_j^t (\omega_t) q_{ij}^t \, , \qquad (4.1)$$

where $i = 1,2,...,n$, M_i^t is the "normal" part of risk reserves (see section 2.3), R_i^0, M_i^0 are initial risk reserves, $q^t = \left\{ q_{ij}^t, \ i = \overline{1,n}, \ j = \overline{1,m} \right\}$, q_{ij}^t is the coverage of a company

i in location j at time t, $\pi_{ij}^t(q^t)$ is the premium from contracts characterized by coverages $\{q_{ij}^t\}$. Full coverages of losses correspond to $q_{ij}^t = 1$. Assume that $c_{ij}^t(q^t)$ is the transaction cost due to administrative or other expenses, $L_j^t(\omega_t)$ is the loss (damage) at j caused by the simulated catastrophic event ω_t at time t. The index t in π_{ij}^t, c_{ij}^t, L_j^t reflects in particular discount rates. Random events $\omega = (\omega_0,...,\omega_{T-1})$ may have random directions of propagation through the region, and they affect a random number of locations $j = 1,2,...,n$. In general, a catastrophic event at time t is modeled by a random subset $\varepsilon_t(\omega)$ of locations j and its strength in each j. The value $L_j^t(\omega_t)$ depends on the event ω_t, mitigation measures, and type of properties in j. The losses of each location may be covered partially by all companies, i.e. variables q_{ij}^t satisfy constraints:

$$\sum_{i=1}^{n} q_{ij}^t \le 1,\ q_{ij}^t \ge 0,$$

where $j = 1,2,...m$, $t = 0,1,...,T-1$.

Variables q_{ij}^t allow us to characterize differences in risks from different locations. It is assumed that all companies operate in the direct market with locations and may cover different fractions of catastrophic losses from the same location. The dependence of functions $\pi_{ij}^t(q^t)$, $c_{ij}^t(q^t)$ on i and q^t implicitly incorporate a possibility for some companies (reinsurers) to transact with the insured parties only through other companies (insurers) with additional administrative costs, premiums, etc. Thus $\pi_{ij}^t(q^t)$, $c_{ij}^t(q^t)$ reflect in a sense the best possibilities for i to transact with j. Variables q_{ij}^t interconnect processes R_i^t, $i = 1,2,...,n$ with each other. Inflows of premiums push their trajectories up, whereas claims and transactions costs push them down.

The analytical structure of the probability distribution of the random vector $R^t = (R_1^t, ..., R_n^t)$ is intractable, although, it is possible in special cases to partially evaluate its analytical parts. This information is used in the design of an adaptive Monte Carlo procedure similar to that described in Subsection 2.3.

4.2. Simulation of Catastrophic Events

There are two possibilities to analyze dependent risk processes $R^t = (R_1^t, ..., R_n^t)$: either through analytical evaluation of their probability distributions or directly through underlying stochastic processes, in particular by the Monte Carlo method.

An essential issue for designing a fast adaptive Monte Carlo procedure is the existence of a submodel for catastrophic events enabling fast simulation of losses for any given combination of decision variables. As pointed out by Hammersley and Handscomb [15] and Pugh [24], all Monte Carlo computations may be regarded as estimating the value of an integral

$$\int f d\mu , \tag{4.2}$$

where μ is a measure on a Euclidean space and f is some measurable (sample performance) function. The measure μ is often not known explicitly but only in terms of other explicitly known measures. The function (2.3) is an example of such an integral, where f and the implicitly given measure μ depend (in contrast to the standard Monte Carlo method) on decision variables which must be sequentially adjusted by sampling trajectories of R^t for different combinations of decision variables.

In the case of general processes R^t stochastic spatial patterns of catastrophic event are simulated as a path dependent random field, with different probabilities of moving to adjacent locations. Spatial random trajectories of wind storms are modeled by random lines or as an asymmetric random walk, characterized by a random length, random strength, and random decay at each step. After each simulation of an event, we calculate damages in each location, thus after a sufficient number of simulations we are able (if needed) to obtain a histogram of damages for each location. The histograms of claims depend on decisions and can also be computed for any given combination of decision variables.

Initial property values of different parts of a region can be represented as a "landscape" on Fig. 2.

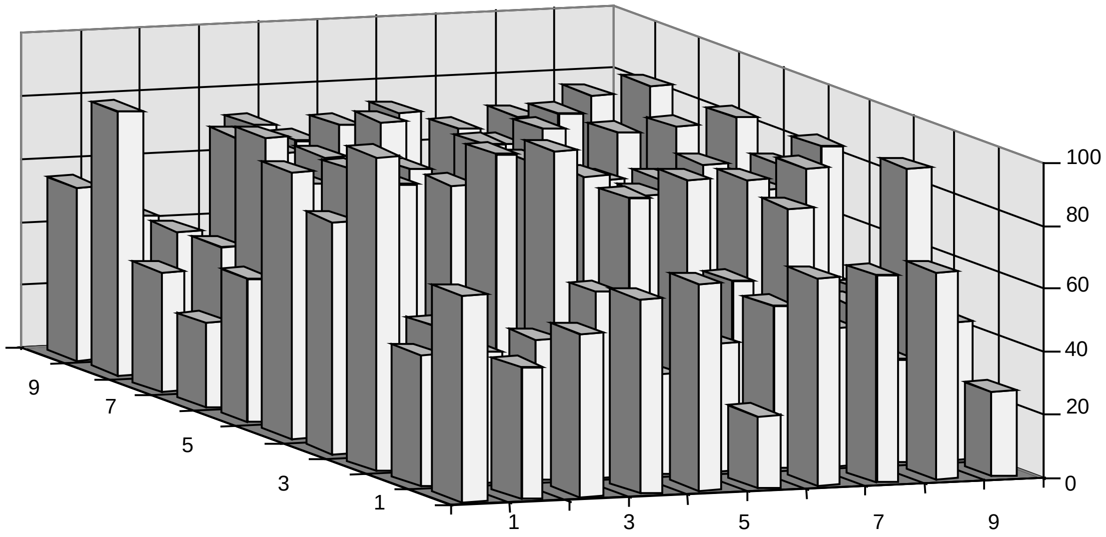

Figure 2. A 'landscape' of initial properties

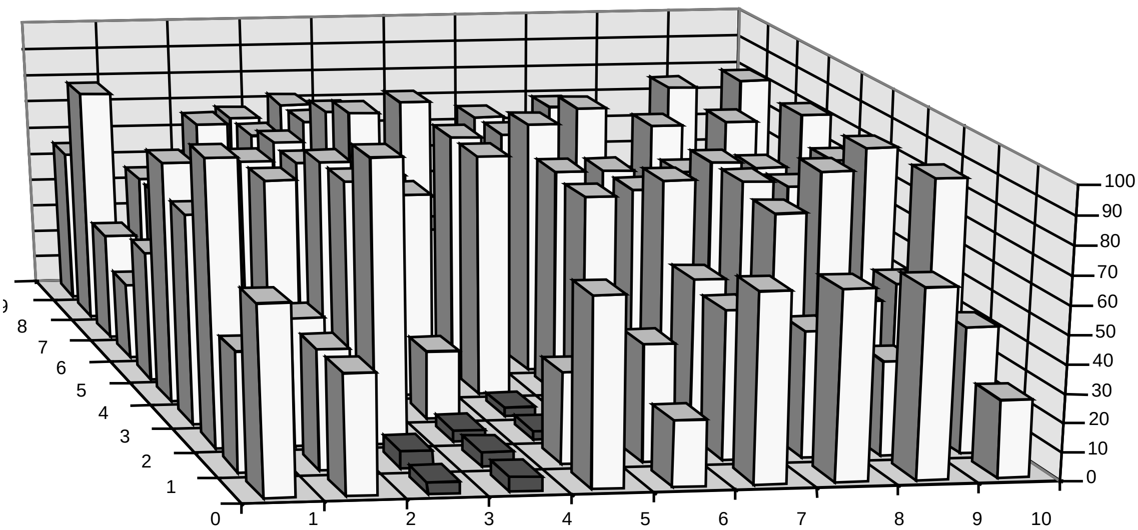

Figure 3. A landscape of damaged property values

A simulated pattern of an event causes damages and may modify the 'landscape' in the way seen on Fig.3.

4.3. General Description of the Model

Without insurance a location j faces losses (damages) L_j^t. Individuals from this location receive compensation $L_j^t q_{ij}^t$ from company j when such a loss occurs. If W_j^0 is the initial wealth (property value), then locations j initial wealth at time $t+1$ is

$$W_j^{t+1} = W_j^t + \sum_{i=1}^{n} (L_j^t q_{ij}^t - \pi_{ij}^t (q^t)) - L_j^t \ . \qquad (4.3)$$

Individuals maximize their wealth, which depend on

$$v_j^t = \sum_{k=0}^{t-1} \left(L_j^k \sum_{i=1}^{n} q_{ij}^k - \sum_{i=1}^{n} \pi_{ij}^k (q^k) \right).$$

Therefore assume that coverages q_{ij}^t are chosen from the maximization of the expectation function

$$F_j(q) = E\left\lfloor v_j^{\tau_j - 1} + \gamma_j \min\{0, W_j^{\tau_j}\} \right\rfloor \qquad (4.4)$$

subject to

$$\sum_{i=1}^{n} q_{ij}^t \le 1 \, , \ j = \overline{1, m}, \ t = 0,1..., T-1 \, , \qquad (4.5)$$

where γ_j is a substitution coefficient or risk coefficient between possible wealth and the risk, τ_j is the time of ruin not exceeding T (stopping time) for location j:

$$\tau_j = \min\left\{ t : W_j^t \le 0, t \le T \right\}.$$

In general case (4.4) is substituted by an evaluation

$$F_j(q) = Ef_j(W_j^t, 0 \le t \le \tau_j)$$

for some function $f_j(\cdot)$.

Similarly, R_i^t describes the wealth (risk reserves) of insurer i at time t. The insurer maximizes (by choosing coverages q_{ij}^t) his expected wealth

$$r_i^t = \sum_{t=0}^{t-1} \left\{ \sum_{j=1}^{m} \left[\pi_{ij}^k(q^t) - c_{ij}^k(q^t) \right] - \sum_{j \in \mathcal{E}_t(\omega_t)} L_j^t(\omega_t) q_{ij}^t \right\}$$

taking into account the risk of insolvency ($R_i^t < 0$). Coverages q_{ij}^t are chosen from maximization of expectation function

$$G_i(q) = E\left[r_i^{\varphi_i - 1} + \delta_i \min\{0, R_i^{\varphi_i}\} \right], \qquad (4.6)$$

subject to (4.5), where δ_i is a substitution coefficient between profit and the risk of insolvency, φ_i is the stopping time

$$\varphi_i = \min\left\{ t : R_i^t \le 0, t \le T \right\}.$$

In general case it is possible again to use an evaluation

$$G_i(q) = Eg_i(R_i^t, 0 \le t \le \varphi_i)$$

for some function $g_i(\cdot)$.

Note that the maximization of $Er_i^{\varphi_i}$ is equivalent to the maximization of the expected profit whereas the maximization $E\min\{0, R_i^{\varphi_i}\}$ eliminates the risk of insolvency of company i.

Remark 4.1. *It can be shown [13] that if the risk coefficients* γ_j, δ_i *become large enough, then the maximization (4.4) and (4.5) is equivalent to the maximization of expected wealth subject to the so-called stability constraints requiring that the probability of insolvency for each insured and insurer does not exceed a given level of "survival".*

The maximization of (4.4) and (4.6) generates the insurance-demand functions $q_{ij}^{D_i}(\pi)$ and the insurance-supply functions $q_{ij}^{S_i}(\pi)$ depending on the premiums $\pi = \{\pi_{ij}^t\}$. The choice of premiums must reflect a certain balances between insurance demand and supply, otherwise higher premiums may decrease profits. In this paper we do not analyze the choice of premiums from this general perspective in contrast to actuarial approaches outlined in Section 2. The main goal is to develop computational approaches that enable the analysis of the choice of optimal coverages improving public benefits, profits of insurers and their solvency for analytically intractable problems. Using the same basic framework as outlined in Section 3 we analyze the choice of insurance contracts for dependent risks subject to additional constraints on the class of feasible contracts.

4.4. Pareto Optimal Coverages

A Pareto optimal improvement of the initial risk situation for insured and insurers with respect to the goal function $F_j(q)$, $G_i(q)$, $q = \{q_{ij}^t, i = \overline{1,n}, j = \overline{1,m}, t = \overline{0,T-1}\}$ can be achieved by maximizing the function

$$W(q) = \sum_{j=1}^{m} \alpha_j F_j(q) + \sum_{i=1}^{n} \beta_i G_i(q) \qquad (4.7)$$

subject to

$$\sum_{i=1}^{t} q_{ij}^t \le 1, q_{ij}^t \ge 0, \; j = 1,2,...,m, \; t = 1,2,...,T, \qquad (4.8)$$

where $\alpha_j \geq 0$, $\beta_i \geq 0$, $\sum_{j=1}^{m} \alpha_j + \sum_{i=1}^{n} \beta_i = 1$.

The Pareto optimality is achieved with respect to the set of goal functions F_j, G_i,

where $\alpha_j > 0$ and $\beta_i > 0$. If we introduce the function

$$W(q,\omega) = \sum_{j=1}^{m} \alpha_j f_j^{\tau_j}(q,\omega) + \sum_{i=1}^{n} \beta_i g_i^{\varphi_i}(q,\omega), \qquad (4.9)$$

where

$$f_j^t(q,\omega) = v_j^t + \gamma_j \min\{0, W_j^t\},$$

$$(4.10)$$

$$g_i^t(q,\omega) = r_i^t + \delta_i \min\{0, R_i^t\},$$

then $W(q)$ can be written as

$$W(q) = EW(q,\omega) \qquad (4.11)$$

We may call $W(q)$ a performance or welfare function and $W(q,\omega)$ a sample performance or sample welfare function. Functions $W(q)$, $W(q,\omega)$ have a complex analytical structure and nonsmooth character. The complexity stems from the complexity of underlying stochastic spatial processes (random fields) defined by simulated patterns of catastrophes. Consistent evaluation of $F_j(q)$, $G_i(q)$ for any feasible strategy q may be time consuming. Since the number of feasible combinations of q is infinite, then the straightforward "trial-and-error" approach to the choice of desirable coverages q is impossible. The nonsmooth character of the functions $F_j(q)$, $G_i(q)$ is also a methodological challenge. It is due to the presence of operations min, max, and stopping times τ_j, φ_i in the definition of $W(q,\omega)$.

The above model can be modified for analyzing the capacity of the insurance "industry" as well as for making decisions by individual companies and pools of companies. In the model described below the emphasis is on the most damaging (extreme) catastrophic events consistent with the existing knowledge of their spatial patterns and occurrence. This stochastic maximin model is a tradeoff between a conservative worst-case approach (all catastrophes are clustered at once in the most "valuable" locations) and the above model. All uncertainties with sufficient historical data are characterized by random variables and other uncertainties are considered from the worst case perspective. For example, the occurrence of events in the region and their magnitudes can be characterized by a given probability distribution (Poisson, Pareto), whereas geographical location and their patterns can be chosen from the worst case.

Let us denote by $\omega = (\omega_1, \omega_2, ..., \omega_{T-1})$ random uncertain variables. For any particular realization of ω_t there exists a set $\mathcal{E}_t(\omega_t)$ of other uncertain variables, say

patterns of catastrophes at time t. Then the guaranteed stochastic risk reserves are

$$R_i^{t+1} = R_i^t + \sum_{j=1}^{m} \left[\pi_{ij}^t(q_{ij}^t) - c_{ij}^t(q_{ij}^t)\right] - \max_{e \in \mathcal{E}_t(\omega_t)} \sum_{j \in e} L_j^t(\omega_t) q_{ij}^t , \qquad (4.12)$$

where $i = 1,2,...,n$, $t = 0,1,...,T-1$. By using R_i^t as (4.12) we can again define functions $F_j(q)$, $G_i(q)$ as (4.4), (4.6), and formulate the problem (4.7)-(4.8).

In the problem (4.7)-(4.8) the risk indicators v_j^t, r_i^t are chosen to guarantee the concavity of the expectations Ev_j^t, Er_j^t. The use of stopping time arguments destroys the concavity of expectation $W(q)$, despite the concavity of involved components. Therefore, let us consider a different model with concave $W(q)$. This model reflects the nature of catastrophes as an extreme event challenging the stability of the whole system once it occurs. Hence the dynamics of the system is modeled until the occurrence of a catastrophic event. Suppose that at each time $t = 0,1,...,T-1$ there may occur a random number of catastrophic events with different magnitudes and geographical patterns. In general it can be represented by two sets of parameters (Ω_t, U_t), where $\omega \in \Omega_t$ characterize their random features and $u \in U_t$ characterize other uncertainties. Define τ as the first moment $t = 0,1,...,T-1$ when a catastrophe occurs. Sample functions W, f_j^t, g_i^t defined by (4.9), (4.10) depend now on the triple of variables (q, ω, u).

Let

$$W(q,\omega) = \min_{u \in U_\tau} = \min_{u \in U_\tau} \left[\sum_{j=1}^{m} \alpha_j f_j^\tau (q,\omega,u) + \sum_{i=1}^{n} \beta_i g_i^\tau (q,\omega,u) \right], \qquad (4.13)$$

i.e. the extreme (worst case) catastrophe is considered with respect to uncertainties $u \in U_\tau$ consistent with other random uncertainties $\omega \in \Omega_\tau$. Instead of $W(q,\omega,u)$ we can consider a more conservative approach where the worst case situation is defined with respect to the risk reserves of each insurer $i = \overline{1,m}$ according to (4.12). It is important that the expectation $W(q)$ in these cases is a concave function. A special case occurs when $U_\tau = 0$, i.e. catastrophes are characterized completely by random patterns. This type of two-period model was considered in [12].

There is a flexibility in choosing the weights α_j, β_i, γ_j, δ_j in (4.4), (4.6), (4.7). These weights can be adjusted to satisfy additional constraints, for example, on fairness or equity. It can be proven (see [13]), that if weights δ_j become large enough, then the effect of the risk function defined by r_i^t is equivalent to the so-called stability constraints [21] requiring that the probability of solvency for each insurer must not drop below a given level of "survival". The performance function (4.7) is composed of different goal functions depending on the choice of weights α_j, β_i. For example if $\alpha_j = 0,\ j = 1,2,...,m$, contracts will take into account only the interests of insurers, with weights δ_j controlling probability of insolvency. The choice $\alpha_j > 0$ emphasizes the interests of the insured, and it can be used to define the levels of premiums depending on the frequency of events, their severity, thus minimizing the losses of insured. Changing γ_j, δ_j it is possible to find contracts satisfying different restrictions on insurance demand and supply, the level of survival for insures and insurers. In Section 6 we describe the results of simulations with different risk weights.

This type of analysis in a sense corresponds to a welfare analysis of the insurance industry as the whole. But the same analysis can be used for a single insurer or a pool of insurers.

4.5. The Role of the Insurance Industry in Managing Catastrophic Risks

The standard analysis of the demand for insurance assumes that an insurance contract is the only available asset for hedging risk (see Mayers and Smith [20]). Catastrophes are characterized by significant interdependencies of claims across different assets of an individual's portfolio, where insurance contracts are not a separable decisions of a general portfolio hedging activity. The demand for insurance exists since not all assets are marketable, i.e. capital markets are not perfect.

The demand for insurance in the presence of other assets can be modeled similar to Mayers and Smith [20]. Instead of eq.(4.3) we define beginning-of-period $t+1$ wealth as

$$W_j^{t+1} = W_j^t + \sum_{k=1}^{l} x_{jk}^t (h_k^t - p_k^t) + \sum_{i=1}^{n} \left[L_j^t(\omega_t) q_{ij}^t - \pi_{ij}^t(q^t) \right] - L_j^t(\omega_t) - r_j^t d_j^t ,$$

where x_{jk}^t is the fraction of a firms's shares held by individuals from location j at time t, h_k^t is the total monetary value of firm j - th shares, p_k^t is the current total market value of firm j -th shares, r_j^t is the riskless rate of return, and d_j^t is the net debt of location j.

The goal function (4.4) for the individuals (from location j) depends now on the decision variables $x = \{x_{jk}^t\}$. The maximization of this function $F_j(q, x)$ subject to (4.5) and additional constraints on x_{jk}^t, $0 \le x_{jk}^t \le 1$, provides a demand function for each type of insurance policy and for risky marketable assets.

The capacity of the insurance industry for managing catastrophic risks depends also on the implemented and available mitigation measures. Assume that for each j there exists a set M_j^t of available mitigation measures at time t. Mitigation measures can be taken by individuals and governments for reducing losses L_j^t. Some of these measures can be enforced by insurers through premiums. From a formal point of view it is equivalent to the assumption that the probability distribution of losses L_j^t and premium functions depend on a new decision variables $y_j^t \in M_j^t$, i.e. $L_j^t(w_t, y^t)$, $\pi_{ij}^t(q^t, y^t)$, where $y^t = \left\{ y_j^t, j = 1,2,...,m \right\}$. The wealth accumulation processes (4.3) in this case include also additional costs associated with decisions y_j^t.

5. Adaptive Monte Carlo Method

As it was mentioned in Section 4 all Monte Carlo computations may be regarded as estimating the value of an integral (4.2). The performance function (4.7) or (4.11) can be written in the same form

$$W(q) = \int_{\Omega} \left[\sum_{j=1}^{m} \alpha_j f_j\ (q,\omega) + \sum_{i=1}^{n} \beta_i g_i\ (q,\omega) \right] d\mu(\omega)\,, \qquad (5.1)$$

where the probability measure μ is defined on the set Ω of catastrophic events $\omega = (\omega_0, \omega_1, ..., \omega_{T-1})$, and $f_j = v_j^{\tau_j} + \gamma_j W_j^{\tau_j}$, $g_i = r_i^{\varphi_i} + \delta_i R_i^{\varphi_i}$ (see also eqs. (4.4), (4.6). The measure μ is not explicitly known and the analytical evaluation of $W(q)$ is practically impossible. Let us begin by fixing decision q. Standard Monte Carlo techniques can be viewed as sampling procedures providing an unbiased estimate of $W(q)$. The smaller the variance of the estimate for a given sample size, the better. By "adaptive Monte Carlo" it is usually meant [24] a technique which makes on-line use of sampling information to sequentially improve the efficiency of the sampling procedure itself. We use "adaptive Monte Carlo" in a rather broad sense when the efficiency of the sampling procedure is considered as a part of more general improvements with respect to different decisions and goals, for example, towards certain equilibriums. It is also possible to use the notion of Monte Carlo optimization but this notion emphasizes only a part of possible adjustments. It can be understood in a narrow sense by those "practitioners" who do not know that the search of equilibriums and solutions of equations can also be viewed as a special optimization problem. The function $W(q)$ depends on unknown decision variables q, and the problem concerns estimating an optimal value $W(q)$ by sampling values of functions $f_j(q,w)$, $g_i(q,w)$ for possibly different q. It is also desirable to combine this with sequential variance minimizing sampling.

In this section we develop the necessary adaptive Monte Carlo procedures by using general ideas of stochastic optimization (see, for example, [10]), which seems to be quite natural for these problems. Let us denote W^* the maximal value of $W(q)$, and rewrite $W(q)$ as

$$W(q) = \int_{\Omega} W(q,\omega) d\mu(\omega) \quad , \tag{5.2}$$

where

$$W(q,\omega) = \sum_{j=1}^{m} \alpha_j f_j(q,\omega) + \sum_{i=1}^{n} \beta_i g_i(q,\omega) \, .$$

The main question is to find a sequence $\{q^k\}$, $k \geq 1$, such that

$$P\left\{ \lim_{k \to \infty} k^{-1} \sum_{s=1}^{k} W(q^k, \omega^k) = W^* \right\} = 1 \, . \tag{5.3}$$

The main complexity in maximizing $W(q)$ concerns the lack of exact information on $W(q)$. Each sample (simulation) provides only a random value

$W(q,w)$ of $W(q)$, which should be used in the search of desirable (optimal) decisions q. There is a number of possibilities to meet this challenge.

5.1. Hypotheses Testing, Response Surface Method

The simplest possibility is to restrict attention to a finite number of feasible coverages $q^1, q^2, \ldots, q^K$. The search of the best q^{k^*} among given K alternatives such that

$$W(q^{k^*}) \geq W(q^k) := E_\omega W(q^k, \omega), k = 1,2,\ldots,K,$$

by using sample functions (for different ω) $W(q^k, \omega)$ is equivalent to a hypothesis testing.

Such an approach is possible only with a good intuition about the structure of optimal decisions. It may be difficult to have such an intuition in the case of structural changes, new policies, complex dependencies and significant effects of low probability events. In these cases we have to take into account something that may be unlike anything we have experienced in the past.

Another approach is to derive an explicit deterministic approximation for $W(q)$ and to use well known deterministic optimization techniques. A family of such procedures is known as the Response Surface Methods.

An initial approximate solution q^0 is usually a very conservative guess. Samples $W(q^k, \omega)$ are used to estimate the optimal (k+1)-step decision q^{k+1}. For this purpose $W(q)$ at $q = q^k$ is approximated by a quadratic regression function and q^{k+1} is constructed by maximizing this function in the feasible set. Such a procedure requires the estimation of a large number of coefficients of the quadratic function at each step $k = 0,1,\ldots$, which may be time consuming and restricts its applicability.

5.2. Sample Mean Approximation

In the Response Surface Methods $W(q)$ is approximated locally at each current approximate solution q^k. Another alternative is to use an explicit approximation for $W(q)$ in the whole feasible set [16]. An important approach is to use the sample mean approximation

$$W^N(q) = N^{-1}\sum_{s=1}^{N} W(q,\omega^s) \tag{5.4}$$

defined by N simulated histories of catastrophes $\omega^s = (\omega_0^s, \omega_1^s, ..., \omega_{T-1}^s)$. The use of this approach is restricted to cases when the sample functions $W(\cdot,\omega)$ have well defined analytical structures. Unfortunately for important applications $W^N(q)$ may have a large number of local optima in addition to the local optima of $W(q)$ (see [9]), which may even happen in the case of concave $W(q)$. A more critical case is when $W(\cdot,\omega)$ is not known explicitly as a function of q. This situation occurs in problems defined by (5.1) since f_j, q_i depend on stopping times τ_j, φ_i, which are implicit functions of current decisions. Therefore at each step $k = 1,2,...$ a deterministic maximization procedure would require new samples of $W(q,\omega)$ at different $q = q^1, q^2, ...$. In addition approximation (5.4) may lead to a significant increase of the dimensionality (see [12]) in contrast to the original problem defined by eqs. (4.7), (4.8).

5.3. Stochastic Quasi-Gradient Methods

These methods can be used in cases with unknown sample performance functions $W(q,\omega)$. A sequence of approximate solutions $q^0, q^1, ...$ is generated directly by using statistical estimates (stochastic quasi-gradients) of $gradW(q,\omega)$ without approximating $W(q)$ by an explicit function. The adaptive search procedure is defined as follows. Let q^0 be an initial guess, and q^k is the approximate solution after k steps. Then

$$q^{k+1} = \Pr j(q^k + \rho_k \xi^k), k = 0,1,..., \quad (5.5)$$

where $\rho_k \geq 0$ is $(q^0, q^1, ...q^k)$ – measurable random variable (“step-size” multiplier depending on $(q^0, q^1, ...q^k)$)), ξ^k is $(q^0, q^1, ...q^k)$ - measurable random vector such that

$$\left| E\left[\xi^k \mid q^0, q^1, ...q^k\right] - gradW(q^k)\right| \to 0, k \to \infty .$$

The symbol $\Pr j(y)$ defines the projection of y onto a feasible set Q defined by (4.8), i.e. it is the point from Q minimizing the distance to y

$$\Pr j(y) = \arg\min\left\{\|z - y\|^2 : z \in Q\right\}.$$

The projection of $q^k + \rho_k \xi^k$ (calculation of q^{k+1}) is a very fast operation when it starts from q^k.

Stochastic quasi-gradients ξ^k are often defined at each step $k = 0,1,...$ by using only one independent sample ω^k. Below we show how it can be applied to problem (4.7), (4.8) and give conditions that ensure the convergence of the nonstationary random process $W(q^k, \omega^k)$ in the sense of (5.3).

The sample function $W(q, \omega)$ in (5.2) is defined by $\min$ and $-\min$ operations. Such functions [11] belong to the so-called generalized differentiable functions that guarantee the convergence of (5.5) to a local optimal solution with probability 1. The class of generalized differentiable functions is closed under operations $\min$ and

$\max(-\min)$ and smooth transformations. Continuously differentiable functions belong to this class.

A stochastic quasigradient ξ^k of function (4.7) is calculated similar to formula given without proof in Ermoliev and Norkin [11] for a simple problem. Assume that

$$\Pr ob\left\{W_j^t(q^t,\omega)=0\right\}=0\ ,\ \ \Pr ob\left\{R_j^t(q^t,\omega)=0\right\}=0$$

for all $q^k, 0\le t\le T$ and j. We can always achieve this by adding some independent random noise with density to R_j^k. Then it can be proven that with probability 1, functions $W(\cdot,\omega)$, $W(q)$ are generalized differentiable with stochastic quasi-gradients ξ^k computed as follows. Let after k steps of adjustments we have a set

$$q=q^k=\left\{q_{ij}^k, i=\overline{1,n},\ j=\overline{1,m}, t=0,1,...,T-1\right\}.$$

Simulate $\omega=\left(\omega_0,\omega_1,...,\omega_{T-1}\right)$ and compute $R_i^0,R_i^1,...,R_i^{\varphi_i}$, $W_j^0,W_j^1,...,W_j^{\tau_j}$ for all i,j. The vector ξ^k consists of components $\left\{\xi_{ij}^k(t), i=\overline{1,n},\ j=\overline{1,m}, t=0,1,...,T-1\right\}$, where $\xi_{ij}^t(t)$ is the sum of four terms $\xi_{ij}^{k1}(t),\xi_{ij}^{k2}(t),\xi_{ij}^{k3}(t),\xi_{ij}^{k4}(t)$:

$$\xi_{ij}^{k1}(t)=\begin{cases}\alpha_j(L_j^{t_j}(\omega)-\partial\pi_{ij}^{t_j}/\partial q_{ij}^t)\ ,\ \text{ if }\ t<\tau_j(q,\omega),\\ 0\ ,\ \text{otherwise,}\end{cases}$$

$$\xi_{ij}^{k2}(t)=\begin{cases}\alpha_j\gamma_j(L_j^t(\omega)-\partial\pi_{ij}^t/\partial q_{ij}^t)\ ,\ t\le\tau_j(q,\omega), W_j^{\tau_j}<0,\\ 0\ ,\ \text{otherwise,}\end{cases}$$

$$\xi_{ij}^{k3}(t)=\begin{cases}\beta_i\left[\partial\pi_{ij}^{\tau_i}/\partial q_{ij}^t-\partial c_{ij}^{\tau_i}/\partial q_{ij}^t-L_j^{\tau_i}(\omega)\right],\ \text{if }\ t<\tau_i(q,\omega)\\ 0\ ,\ \text{otherwise.}\end{cases}$$

$$\xi_{ij}^{k4}(t)=\begin{cases}\beta_i\delta_i\left[\partial\pi_{ij}^t/\partial q_{ij}^t-\partial c_{ij}^t/\partial q_{ij}^t-L_j^t(\omega)\right],\ t\le\tau_i(q,\omega), R_i^{\varphi_i}<0\\ 0\ ,\ \text{otherwise,}\end{cases}$$

Thus there are simple formulas for computing a stochastic quasigradient $\xi^k = \left\{\xi_{ij}^k(t)\right\}$ $\xi_{ij}^k(t) = \xi_{ij}^{k1}(t) + \xi_{ij}^{k2}(t) + \xi_{ij}^{k3}(t) + \xi_{ij}^{k4}(t)$ after each simulation $\omega = \omega^k$ and $q = q^k$, $k = 0,1,...$. The current approximate decision variables $q^k = \{q_{ij}^{kt}\}$ are adaptively adjusted according to feedback (5.5). Since function $W(q)$ is not concave even for convex $\pi_{ij}^t(\cdot)$, $c_{ij}^t(\cdot)$, then random sequence q^k generated according to (5.5) may not converge to a global solution. The choice of step-size-multipliers ρ_t in (5.5) satisfies conditions

$$\rho_t \geq 0, \sum_{t=0}^{\infty} \rho_t = \infty, \sum_{t=0}^{\infty} \rho_t^2 < \infty .$$

For example $\rho_t = C_t / t$, where $0 \leq \underline{C}_t \leq C_t \leq \overline{C}_t < \infty$ ensures the convergence of $\left\{W(q^k)\right\}$ to a local maximum value with probability 1 in all practically important cases. If we define the set of local maximum values of $W(q)$ as W^* then instead of (5.3) it is possible to show that

$$P\left\{\lim_{k\to\infty} k^{-1} \sum_{s=1}^{K} W(q^k, \omega^k) \in W^*\right\} = 1. \qquad (5.6)$$

The random adjustment mechanism (5.5) has the ability to by-pass local solutions. Global convergence can be achieved by introducing “shocks” when the sequence q^k shows a steady-state tendency.

The important feature of the functions (4.13) is that for convex $\pi_{ij}^t(q^t)$, $c_{ij}^t(q^t)$ this function is concave despite the very complex character of the implicitly given function $W(q,\omega)$. This was achieved by a special choice of risk indicators v_j^t, r_i^t in the definition of functions f_j , g_i . Note that slightly different indicators as in

(3.4) lead to a nonconcave generalized differentiable function $W(q)$. The concavity of $W(q)$ simply follows from the concavity of functions $v_j^t(\cdot,\omega)$, $r_i^t(\cdot,\omega)$, $W_j^t(\cdot,\omega)$, $R_i^t(\cdot,\omega)$ and general properties of the expectation operator. In this case the convergence is global and (5.3) holds.

The numerical experiments (see section 6) so far have been done only for performance indicator (4.13).

5.4. Adaptive Importance Sampling

The fast simulation of rare events and the variance reduction of estimates $W(q^k), k = 0,1,...$ can be achieved in particular by the method of importance sampling. The general idea of adaptive gradient type improvement of sampling procedure was introduced by Pugh [24]. Unfortunately this itself requires the additional estimation of some involved integrals. Stochastic optimization procedure (5.5) allow us to incorporate sequential variance reduction processes without additional major computations.

Consider a probability measure ν on the domain of μ such that whenever ν is zero μ is zero. Then the derivative $d\mu / d\nu$ exists and

$$W(q) = \int W(q,\omega)d\mu(\omega) = \int W(q,\omega)\frac{d\mu}{d\nu}d\nu(\omega) := \int W'(q,\omega)d\nu(\omega)$$

with variance

$$VarW'(q,\omega) = \int W^2(q,\omega)\left(\frac{d\mu}{d\nu}\right)^2 d\nu - W^2(q).$$

The aim is to find a ν that minimizes

$$E_\omega W^2(q,\omega)\left(\frac{d\mu(\omega)}{d\nu(\omega)}\right)^2 = \int W^2\left(\frac{d\mu}{d\nu}\right)^2 d\nu. \qquad (5.7)$$

Let the family of distributions ν be indexed by the vector parameter $y = (y_1, y_2, ..., y_k)$. Thus (5.7) is a function $\Psi(y)$ of y and the direction of steepest decent of this function at y (assuming regularity conditions) is

$$-\partial\Psi / \partial y_l = -\left(\int W^2 \left(\frac{d\mu}{d\nu} \right)^2 d\nu \right)_{y_l} = -\left(\int W^2 \frac{d\mu}{d\nu} d\mu \right)_{y_l} = -\int W^2 \left(\frac{d\mu}{d\nu} \right)_{y_l} \frac{d\mu}{d\nu} d\nu \quad (5.8)$$

Together with procedure (5.5) consider a sequence of measures ν_k defined by a sequence of vectors $\{y^k\}$. Assuming that ν_k is known we seek the ν_{k+1} which decreases (5.7) for current $\nu = \nu_k$, i.e. we choose y^{k+1} defined by

$$y^{k+1} = y^k - \sigma_k W^2 (q^k, \omega^k) \frac{\partial}{\partial y_l} \left(\frac{d\mu}{d\nu} \right)_{y=y^k} \left(\frac{d\mu}{d\nu} \right), \qquad (5.9)$$

where ω^k is a sample from ν_k, and $\sigma_k > 0$ is a positive $(q^0, y^0, q^1, y^1, ..., q^k, y^k)$ measurable random variable satisfying some natural joint requirements with ρ_k. The procedure (5.9) requires exact values

$$\frac{d}{dy_l} \left(\frac{d\mu}{d\nu} \right)_{y=y^k}, \left(\frac{d\mu}{d\nu} \right)_{y=y^k},$$

which are not known explicitly because μ is also not explicitly given. These values can be substituted by statistical estimates, which is discussed in a forthcoming paper for some important special cases. Section 2.3 illustrates in a sense such a possibility. The convergence of the resulting processes easily follows because $W(q)$ does not depend on y.

6. Numerical Experiments

We consider a fictitious region subdivided into 10×10 grids. An example of a geographical representation of the property values in a "landscape" is shown in Fig.2. The time span is $T = 1000$. The occurrence of catastrophes in the region is modeled according to a given distribution of interoccurrence times. We also assumed that at each time interval $t = 0,1,...,T-1$ only one catastrophe may occur. Numerical experiments so far have been done with concave version (4.13) of the dynamic model. Catastrophes are assumed to be random events, i.e. $U_t = 0$.

A catastrophic event starts at random from a grid and propagates through the region in the form of a random walk having a random magnitude and rate of decay. The transition probability to an adjacent grid depends on some characteristics of the grid. In particular sample trajectories may have the form of random lines starting at random grids and having random direction and random length. An example of the damage caused by catastrophic events is shown in Fig.3. The initial geographical diversification of contracts for three companies is shown in Fig.4.

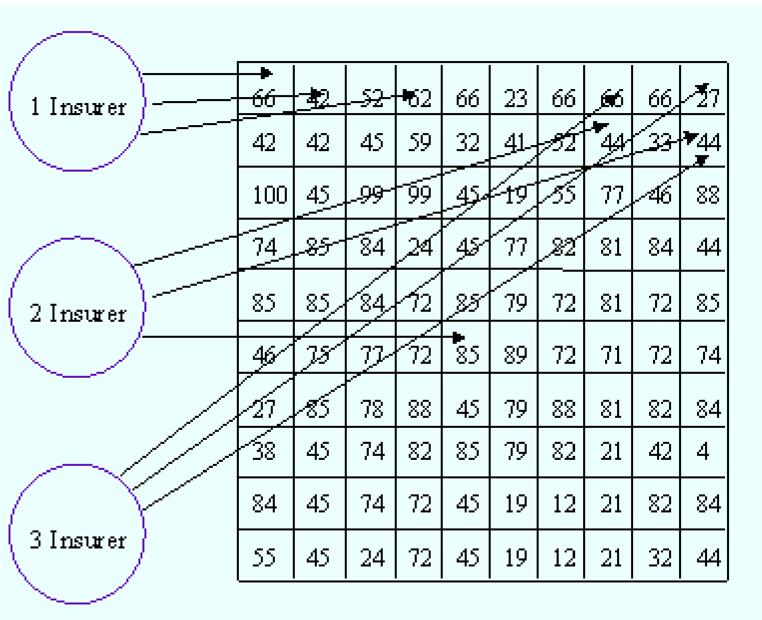

Figure 4. Initial allocations of contracts

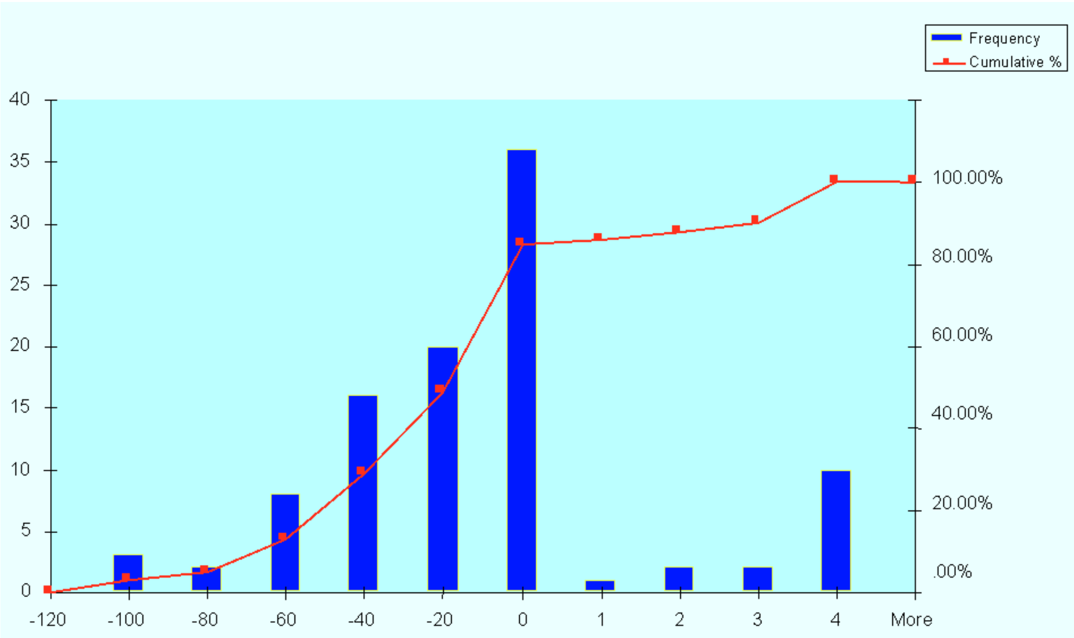

Figure 5. Histogram of the risk reserve (insurer 1) at initial contracts

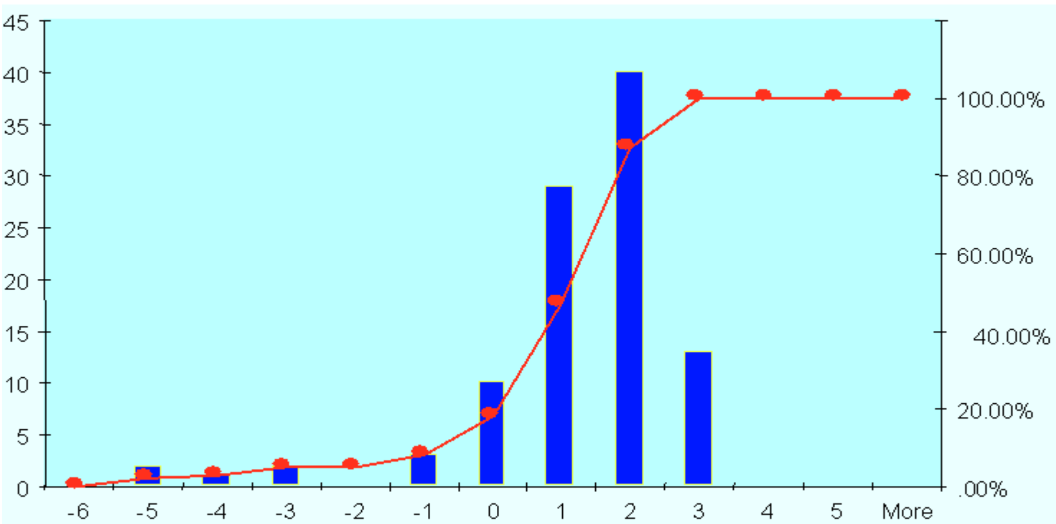

Figure 6. Histogram of the risk reserve (insurer 1) at improved contracts

After simulating a sufficient number of events it is possible (if needed) to obtain histograms of risk reserves for the initial contracts as shown in Fig.3.

In this case the insolvency of the insurer within the time span happens rather often.

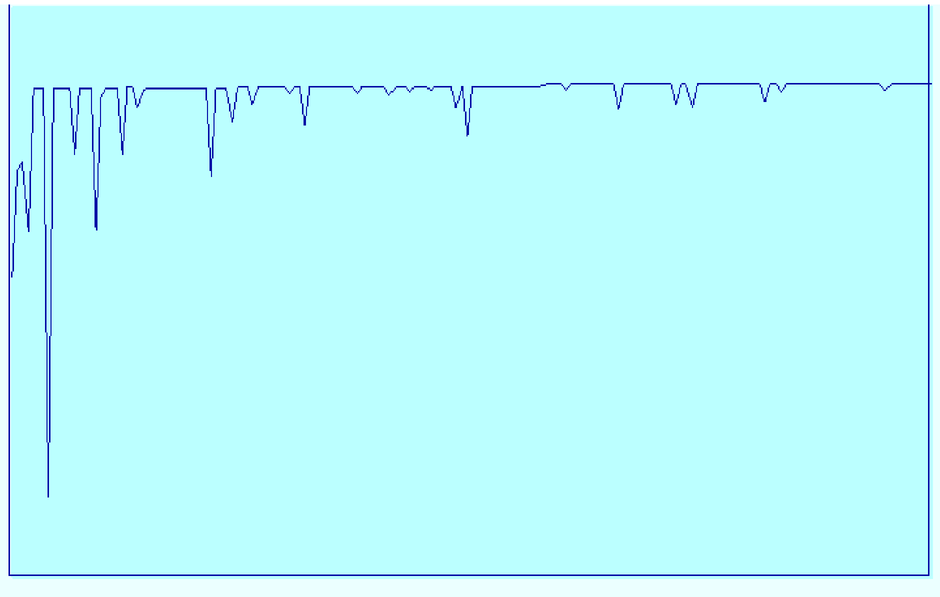

Figure 7. Dynamics of improvements for performance function

Optimal geographical diversification of coverages improves the insolvency of insures (Fig.6), although different levels of insolvency may still occur with some probabilities. It can be changed by choosing different weights δ_i. The dynamics of improvements for the performance function (4.13) is shown in Fig.7. It shows slow improvements of the performance function with considerable elimination of ruins by choosing better coverages.

The performance function is stabilized rather fast, but variances exist and even last simulations eliminate influences of rare events. The difference between initial and final histograms (Fig.5, 6, 9) is remarkable.

The new allocations of contracts, shown in the Fig.8, are diversified over the territory with respect to simulated events. More deep analysis shows that insurers tend to deal with locations where damages are almost mutually exclusive. Perhaps, insurers allocate contracts in ’safe’ regions, where events may occur with possibly minimal probabilities.

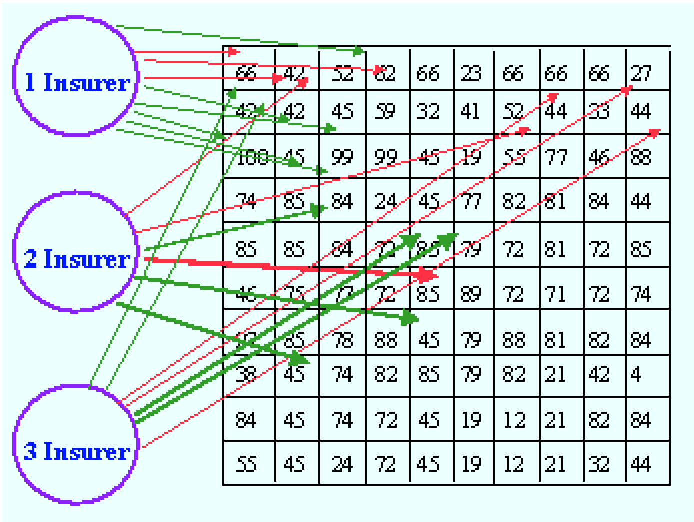

Figure 8. Improved allocations of contracts

All three insurers shown in Fig.4 differ in their initial allocations and their restrictions on the possibility of new contracts. The first insurer may obtain new contracts only in the most risky upper left corner. The second can deal with "safer" clients. The third insurer may have less "safe" new clients than the second, but from locations where catastrophes can be regarded as almost mutually exclusive.

From the final spread of optimal contracts we can see that insurer 1 improves her situation getting more additional risks from locations where she can operate. Therefore, this insurer does not become "afraid" of catastrophes, keeps operating in the region, and therefore provides support to the population. The "safest" insurer 2, preserving small contracts with "risky" regions, makes business mainly with remote clients where catastrophes are very rare. He makes his business as profitable as possible and protects himself from the risk of insolvency as much as possible. The third insurer may be regarded as the most socially oriented one providing additional help to the suffering in catastrophic locations. For protection against insolvency it takes new contracts in locations where catastrophes are often almost mutually exclusive.

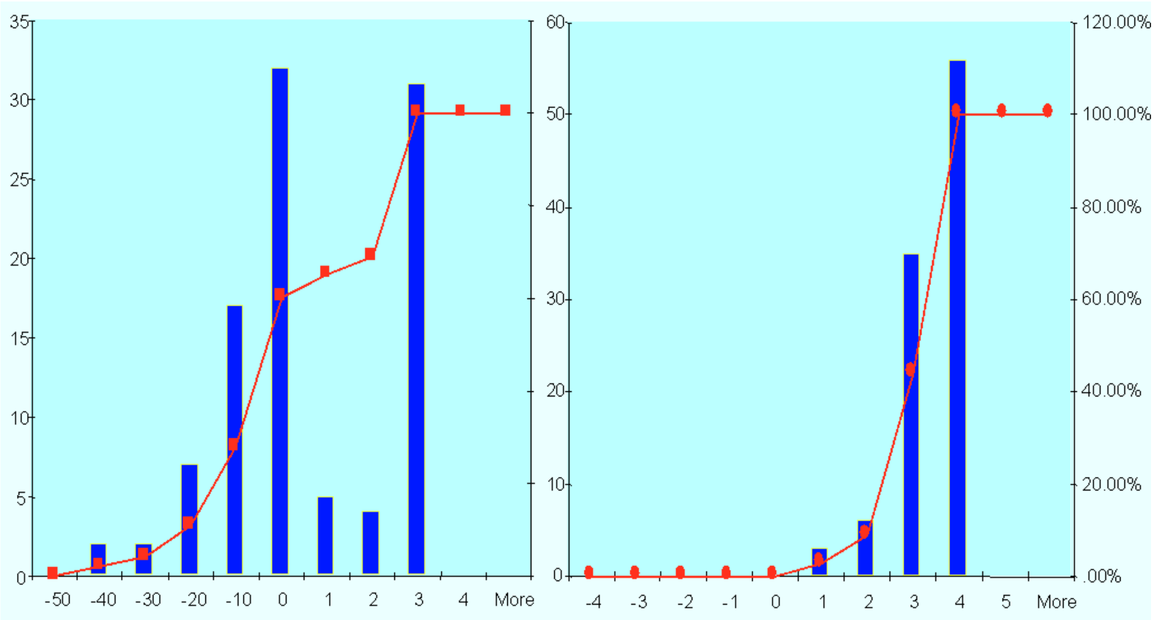

Figure 9. Histograms of the risk reserve (insurer 2) at initial contracts and improved contracts with risk weight 100

Varying risk weights in the performance function it is possible to satisfy different conditions on the companies solvency. Computational results show that increasing the risk weights may decrease the risk of insolvency of different companies to some predefined levels. Thus Fig.9 shows histograms of risk reserve at initial contracts for insurer 2 and its improved contracts for the risk weight equal to 100. All risk constants equal to 1000 lead to optimal contracts providing for the absolutely safe business of insurers. This case eliminates coverages in locations where catastrophes are rather often and can not be perfectly diversified. For insurer 2 the histogram of risk reserve at the improved contracts with risk weight 1000 are shown in Fig.10.

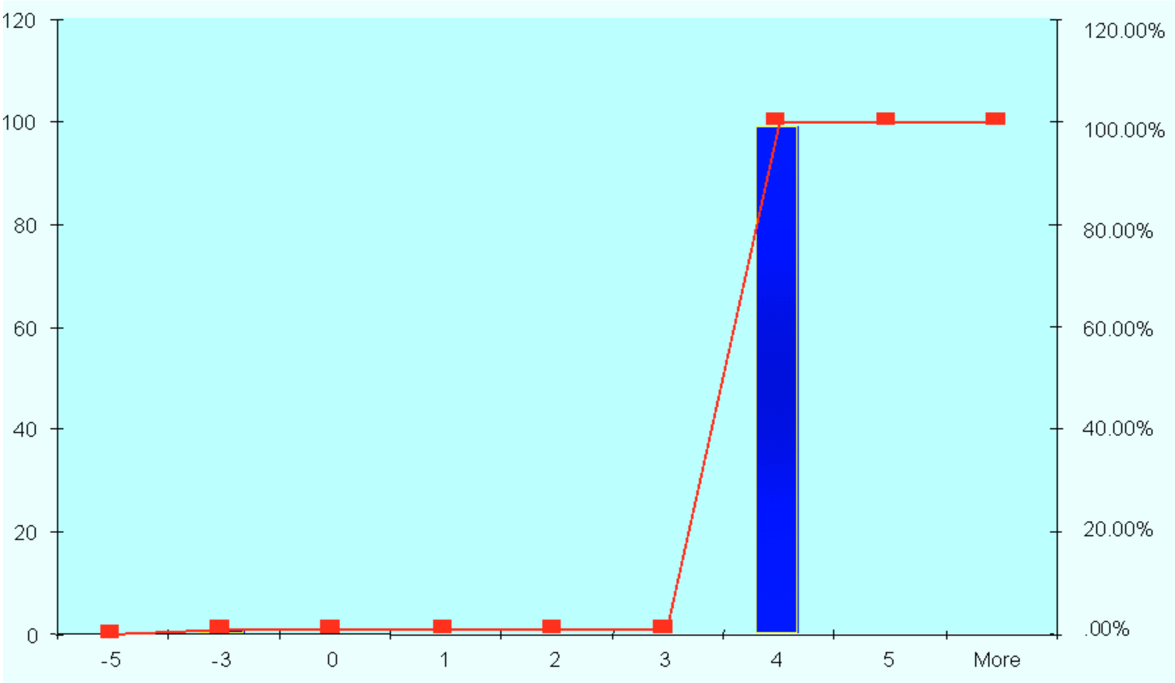

Figure 10. Histogram of the risk reserve (insurer 2) at improved contracts with risk weight 1000.

7. Concluding Remarks

In this paper I have presented computational approaches for designing optimal insurance strategies in the presence of the dependent catastrophic risks. The developed spatial dynamic model of stochastic optimization can be used either by a single insurer (n=1), a pool of insurers (n>1) or regulatory authorities. The model can also be used for analyzing the capacity of the "insurance industry" in dealing with catastrophes. In this case the model requires a detailed representation of other types of hedging decisions, which have been outlined in the paper.

The model tracks the dependencies of catastrophic claims by explicit representation of the special characteristics of the property values and the spatial patterns of possible catastrophic events. It enables one to bypass some serious limitations of Borch's model concerning the substitutability of risks. For this purpose, constraints on risks from each location instead of the single constraint on the “total” risk have been imposed. The model also allows one to introduce transactions costs for dealing with different locations (remote clients). Explicit incorporation of simulation models for catastrophic events opens up a way for analyzing the interplay between changes in frequencies, magnitudes, patterns of catastrophes and insurance strategies.

I took several different approaches to modeling catastrophic events. In one approach the insurance processes are simulated within the time interval [0,T]. In this case the terminal state of a company is associated with its bankruptcy ("stopping time"). In a second approach insurance processes are simulated until the first occurrence of the catastrophic event. In this case extreme events are associated with the worst case values of uncertain nonstochastic variables. It leads to the so-called stochastic maximin problems.

The adaptive Monte Carlo method is used for adjusting the feasible decision variables towards desirable outcomes. This method is based on stochastic optimization techniques.

The necessary proofs are only outlined since they are lengthy. For example, the convergence properties (5.3), (5.6) of the search procedures are equivalent to laws of large numbers for path-dependent nonstationary processes. Rigorous proofs of these assertions are beyond the scope of this paper. The analysis of generalized differentiability of the performance function and its generalized gradients with stopping times is rather lengthy. By using these gradients it is possible to formulate optimality conditions generalizing Borch's results.

Special attention has been given to the analysis of iterative importance sampling imbedded in stochastic optimization procedures as well as other specific variance reduction techniques. These ideas have been analyzed by using a number of special practical cases.

I have presented numerical experiments with fictitious data to illustrate the feasibility of the developed approaches. These experiments also demonstrate the capability of the stochastic quasi-gradient procedure for designing optimal insurance decisions in the presence of dependent catastrophic risks. The advantage of these methods stems from the lack of a tractable analytical structure of the sample performance function, which often excludes any alternative approach. The experiments show that the computer time required for the search of the optimal value of the performance function has the same order of magnitude as the time, required for estimating its value at a given initial point.

8. References

1. K.J. Arrow (1971). Insurance, Risk and Resource Allocation. Essays in the Theory of Risk-Bearing, New-York: American Elsevier.
2. R. Beard, T. Pentikainen, E. Pesonen (1984). Risk Theory. Printed in Great Britain at the University Printing House, Cambridge.
3. K. Borch (1962). Equilibrium in a reinsurance market. Econometrica, Vol.30, No.3.
4. H. Buhlmann (1970). Mathematical Methods in Risk Theory, Springer-Verlag, New York, Heidelberg Berlin.
5. P.J. Cook, D. A. Graham (1977). The demand for insurance and protection: the case of irreplaceable commodities. Published by John Wiley & Sons, Inc.
6. J.D. Cummins and N. Doherty (1996). Can insurers pay for the "Big One"? Measuring capacity of an insurance market to respond to catastrophic losses. Wharton Risk Management and Decision Processes Center, University of Pennsylvania, W.P.
7. C.D. Daykin, T. Pentikainen and M. Pesonen (1994). Practical Risk theory for Actuaries. Monographs on statistics and applied probability, Volume 53, Published by Chapman and Hall Ltd..
8. N. Doherty (1996). Insurance Markets and Climate change, University of Pennsylvania, W.P.
9. Y.Ermoliev and V. Norkin (1997). On nonsmooth and discontinuous problems of stochastic systems optimization, European Journal of Operation Research (forthcoming).

10. Y. Ermoliev, R. Wets (eds.) (1988). Numerical techniques of stochastic optimization. *Computational Mathematics*, Springer-Verlag.
11. Y. Ermoliev, V. Norkin (1997). Stochastic generalized gradient method with application to insurance risk management. IIASA Interim Report, IR-97-021.
12. T. Ermoliva, Y. Ermoliev, V. Norkin (1997). Spatial Stochastic Model for Optimization Capacity of Insurance Networks Under Dependent Catastrophic Risks: Numerical Experiments. IIASA Interim Report, IR-97- 028.
13. T. Ermolieva, Y. Ermoliev and V. Norkin (1997). Computational approaches to the design of insurance decisions (forthcoming).
14. Insurance Service Office (1994) "The Impact of Catastrophes on Property Insurance", New York.
15. J.M. Hammersley and D.C.Handscomb (1964). Monte Carlo Methods. John Wiley, New York.
16. P. Kall, A. Ruszczynski, and K. Frauendorfer (1988). Approximation Techniques in Stochastic Programming in Y.Ermoliev and R. Wets (Eds.), Numerical Techniques for tochastic Optimization, Springer-Verlag, Computational Mathematics, 1988.
17. H. Kunreuther (1996). Mitigation Losses and Providing Protection Against Catastrophic Risks. The Role of Insurance and Other Policy Instruments. University of Pennsylvania, W.P.
18. H. Kunreuther, N. Doherty, P. Kleindorfer (1994). The role of insurance and regulations in dealing with natural disasters. Proposal to the National Science Foundation, April 20, 1994.
19. G. MacDonald (1992). Persistance in Climate. JSR-91-340, The MITRE Cooperation, McLean, Verginia, 22102-3481.
20. D. Mayers and C. Smith (1983). The interdependence of Individual Portfolio Decisions and Demand for Insurance. Journal of Politiacal Economy, vol. 91, No.2.

21. T. Pentikainen, H. Bonsdorf, M. Pesonen, J. Rantala, M. Ruohonen (1989). Insurance solvency and financial strength. Finnish Insurance Training and Publishing Company Ltd., Helsinki.
22. G. Pflug (1996). Optimization of stochastic models. Kluwer Academic Publishers.
23. G. Pflug (1996). Risk-Reshaping Contracts and Stochastic Optimization. IIASA, WP-96-142.
24. E.L. Pugh (1966). A gradient technique of adaptive Monte Carlo. SIAM Review, vol. 8, No. 3.
25. Report of Swiss Reinsurance Company(1997). Natural catastrophes and major losses in 1996: high losses from man-made disasters, but no extremely costly losses from natural catastrophes. Sigma No.3.
26. Report of Swiss Reinsurance Company (1997). Alternative risk transfer via finite risk reinsurance: an effective contribution to the stability of the insurance industry. Sigma No.5.
27. Report of Swiss Reinsurance Company (1996). Insurance in Eastern Europe: a growth industry on the way towards market structures. Sigma No.8.
28. Snyder (1993).
29. M. Tucker (1997). Analysis: Climate change and the insurance industry: the cost of increased risk and the impetus for action. Ecological Economics, 22 (1997) 85-96, Natural hazards observer, volume 221, number 6, July 1997.