13th World Conference on Earthquake Engineering
Vancouver, B.C., Canada
August 1-6, 2004
Paper No. 1416

PORTFOLIO OPTIMIZATION AND VALUE OF INFORMATION ON CATASTROPHE INSURANCE

Carol HAYEK[1] and Roger GHANEM[2]

SUMMARY

The last decades have experienced large economical losses associated with natural disasters forcing insurance companies to better assess catastrophe losses and to work on their portfolio diversification. Given the lack of historical data required to perform statistical analysis on insurance claims, insurers are increasingly relying on the engineering community for catastrophe loss modeling. Consequently, loss estimation procedures are being amended to handle insurance issues.

Moreover, loss estimation procedures embody considerable amount of uncertainty at all levels, in the hazard quantification, inventory assessment and vulnerability analyses, and therefore a need of improved information representation and integration has emerged. A significant component in reducing this uncertainty can be associated with better information management tied to engineering knowledge. Efforts have been recently undertaken to acquire new information through the application of advanced technologies and damage sensing devices to the area of catastrophe loss estimation. In view of that, this work underscores the benefits of the use of such technologies and the information they carry in the field of catastrophe modeling by linking these new concepts to insurance decisions.

A sensitivity analysis is carried out pointing at the impact of improved data on the portfolio predictions and the risk behind making portfolio decisions with inaccurate data. The portfolio distribution is controlled by an optimization model for insurance profit maximization under inslovency and stability constraints. The insurance decision-making process has access to spatially and structurally resolved portfolio. Implementing the present research will help insurance companies decide on the level of information required, the means of data acquisition and translate these decisions into the portfolio selection criterion.

INTRODUCTION

[1] Graduate Student, Department of Civil Engineering, Johns Hopkins University, 3400 N. Charles street, Baltimore, MD 21218. E-mail: carolhayek@jhu.edu
[2] Professor, Department of Civil Engineering, Johns Hopkins University, 3400 N. Charles street, Baltimore, MD 21218. E-mail: ghanem@jhu.edu

The insurance industry is facing a difficult position with regard to portfolio optimization given the increasing trend in catastrophe losses. Furthermore, the predictions of future disasters inferred by catastrophe modelers suggest that a major event taking place in an urban area would lead to escalating losses. Unlike traditional insurance where frequent, independent, low consequence events and rich historical data is available, catastrophe insurers cope with low probability, high consequence events where historical data is neither common nor accurate. In light of that, the present work aims at bridging the gap between loss estimation procedures, the use of enhanced information through advanced technologies and engineering studies. The purpose behind utilizing remote sensing for inventory data is to bring new insight to existing loss estimation methodologies by means of data collection. The traditional way of collecting data is ground based, profile oriented, time consuming and expensive. Introducing advanced technologies changes the way data collection is perceived, from individual building characteristics assessment to large-scale structural and economical characteristics assessment [1;2].

To handle the flow of information, an information set is created composed of different data accuracy levels where each level contains some improved data acquired from either field survey, advanced technologies or engineering research. The improved data relates to a data enhancement of any kind consistent with the building type classification such as building specific inventory data or structural properties.
In short, this study aims at incorporating this information into loss estimation, and assessing its effect on insurance decision-making. The main objectives of the present paper are summarized as follows:

- Modeling the use and impact of information from field survey, advanced technologies and engineering knowledge on catastrophe insurance decision-making
- Modeling optimal coverage distribution per location as well as per building type
- Optimizing profit under insolvency and stability risks

The first part of the paper includes the development of the database needed from generation of earthquake catalog to economical losses. The next section presents the methodology for the portfolio optimization problem and the sensitivity analysis followed by a section illustrating the implementation. The last sections are intended for results and conclusion.

METHODOLOGY

Stochastic optimization models have been previously applied to the field of insurance decision-making [3-5]. In this study also, a chance constrained optimization is proposed to deal with insurance portfolio selection. The proposed optimization model enables insurers to have control over the spatial distribution as well as the buildings structural type distribution reflecting the buildings vulnerability to earthquakes. Furthermore, the methodology incorporates the value of better information through the use of advanced technologies, field survey and engineering research in loss estimation. Information is treated as a commodity that can be valuated to help insurers decide on the accuracy of information required to better assess their portfolio objectives. This section presents a summary of the proposed methodology focusing first on formulating an objective function and constraints for the optimization problem and then on the value of information.

Portfolio Optimization

In the catastrophe insurance business, corporations aim at maximizing their profit while controlling their risks [6;7]. Insurers worry primarily about insolvency and have to make sure that the probability of losses exceeding the insurance's capability is less than a certain threshold dictated by the management team. Another issue of concern is the stability risk which asserts that losses plus expenses stay within a range of the collected premiums, it basically ensures positive expected profits. This paragraph outlines the insurance profit and constraints required to define the optimization formulation.

Insurance Profit

The insurance profit consists of income minus losses and expenses, where income represents the collected premiums and expenses involve transactions cost, fees and overheads. Since such expenses can be reduced through good management and business operability, they are considered internal issues and therefore will not be treated here. They are accounted for in the premium loading factor.

The general form of the insurance profit is expressed as follows, where m represents the number of locations and n represents the structural types included in the analysis,

$$\text{Profit} = \text{Premiums - Cost/Fees - Covered Losses} \tag{1}$$

$$\text{Premiums } \Pi = \sum_{j=1}^{m}\sum_{k=1}^{n}(1+\lambda_{kj})\mathrm{E}_{\Omega}\left[L_{kj}(w)d_{kj}\right] \tag{2}$$

$$\text{Covered Losses } L = \sum_{j=1}^{m}\sum_{k=1}^{n}L_{kj}(w)d_{kj} \tag{3}$$

$$d_{kj} = \frac{\text{Property value to be covered}}{\text{Total existing property value}} \quad \text{for all } k=1,\ldots,n \text{ and } j=1,\ldots,m \tag{4}$$

$$\text{Profit } P(\underline{d}) = \sum_{j=1}^{m}\sum_{k=1}^{n}\Big((1+\lambda_{kj})\mathrm{E}_{\Omega}\left[L_{kj}(w)\right]-L_{kj}(w)\Big)d_{kj} - C_{kj} \tag{5}$$

where:

w: earthquake event
d_{kj}: coverage ratio for each location j and structural type k
λ_{kj}: loading factor at each (k,j), reflecting transaction cost, profits, margins and so on
$L_{kj}(w)$\$: simulated economical losses at each (k,j) for every event w
C_{kj} information cost at each (k,j)
$\mathrm{E}_{\Omega}[L_{kj}(w)d_{kj}]$: expected value of covered losses for every (k,j) over all events w.

The premium pricing is based on the expected economical catastrophe losses for each pair (k,j) increased by a loading factor λ_{kj}. This factor covers transaction costs and fees and will be taken as a known constant for all locations and structures. Including λ_k as a decision variable in the optimization process requires an extended model that includes such factors as market competition and insured's willingness to pay for coverage. In order to emphasize the role of information, its cost C_{kj} will be treated separately. This cost denotes any charge incurred for information acquisition through satellite imaging, image processing, field survey or additional services requested by catastrophe modelers in exchange for enhanced database and vulnerability analysis. The assigned dollar value depends on the location, the structure and on the information resolution.
The optimization is carried out for a specific information level determined from the information analysis treated in the next section.

Insolvency Constraint

The maximization of the profit is subject to the insolvency risk representing the possibility of ruin of the insurance company. It is now appropriate to include the term insurance reserve R_0 which indicates a certain amount of money set aside by insurers for unexpected losses and payment of ongoing actual losses. Ruin occurs when the covered losses exceed reserve plus income reduced from all costs. The ruin constraint is expressed by,

$$\text{RC} \equiv \Pr[\text{ Reserve+Premiums-Costs/Fees-Covered Losses } \leq 0] \leq \alpha_1 \tag{6}$$

which means that for a particular information level the constraint is,

$$\Pr[R_0 + \sum_{j=1}^{m}\sum_{k=1}^{n}\left((1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)] - L_{kj}(w)\right) d_{kj} - C_{kj} \leq 0] \leq \alpha_1$$

$$\Pr[\sum_{j=1}^{m}\sum_{k=1}^{n} L_{kj}(w) d_{kj} \geq R_0 + \sum_{j=1}^{m}\sum_{k=1}^{n}(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)]\, d_{kj} - C_{kj}] \leq \alpha_1$$

This is interpreted as a chance constraint of the form Pr $[L \leq x] \leq \alpha$. To solve this problem, this probabilistic constraint is converted into its deterministic equivalents [8-10] by means of the cumulative distribution function F_L such that Pr $[L \leq x] = F_L(x)$ implying that,

$$\Pr[\sum_{j=1}^{m}\sum_{k=1}^{n} L_{kj}(w) d_{kj} \leq R_0 + \sum_{j=1}^{m}\sum_{k=1}^{n}(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)]\, d_{kj} - C_{kj}] \geq 1-\alpha_1$$

$$R_0 + \sum_{j=1}^{m}\sum_{k=1}^{n}[(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)]\, d_{kj} - C_{kj}] \geq F_L^{-1}(1-\alpha_1)$$

where α_1 is a deterministic target value set by the management team, and F_L is the cumulative distribution function of the total predicted insurance catastrophe losses.

Stability Risk

The stability constraint monitors the profitability of the insurance company by restricting the probability that losses and expenses exceed a certain proportion of the premiums,

$$\text{SC} \equiv \Pr\left[\frac{\text{Covered Losses + Costs}}{\text{Premiums}} \geq \beta\right] \leq \alpha_2 \tag{7}$$

It is expressed by,

$$\Pr\left[\frac{\sum_{j=1}^{m}\sum_{k=1}^{n} L_{kj}(w) d_{kj} + C_{kj}}{\sum_{j=1}^{m}\sum_{k=1}^{n}(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)] d_{kj}} \geq \beta\right] \leq \alpha_2$$

$$\Pr[\sum_{j=1}^{m}\sum_{k=1}^{n} L_{kj}(w) d_{kj} - C_{kj} \geq \sum_{j=1}^{m}\sum_{k=1}^{n}\beta(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)] d_{kj}] \leq \alpha_2$$

$$\Pr[\sum_{j=1}^{m}\sum_{k=1}^{n} L_{kj}(w) d_{kj} \leq \sum_{j=1}^{m}\sum_{k=1}^{n}\{\beta(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)] d_{kj} - C_{kj}\}] \geq 1-\alpha_2$$

$$\sum_{j=1}^{m}\sum_{k=1}^{n}\{\beta(1+\lambda_{kj})\mathrm{E}_\Omega[L_{kj}(w)] d_{kj} - C_{kj}\} \geq F_L^{-1}(1-\alpha_2)$$

Again α_2 and β are deterministic target values set by the management team, and F_L is the cumulative distribution function of the sum of the simulated economical losses multiplied by the relevant coverage ratio d_{kj} over all locations and buildings structural types.

Optimization Formulation

The objective is then to maximize the expected insurance profit under calculated and defined risks. Given the profit linearity in covered losses, the expected profit is proportional to the expected losses. The optimization problem for a given information resolution I can then be stated as,

$$\max_{\underline{d}} \; \mathrm{E}_{\Omega}\left[P(\underline{d})\right] \;\equiv\; \max_{\underline{d}} \; \sum_{j=1}^{m}\sum_{k=1}^{n}\Big(\lambda_{kj}\mathrm{E}_{\Omega}[L_{kj}(w)]\, d_{kj} - C_{kj}\Big) \tag{8}$$

subject to the following constraints,

$$1.\;\; F_L^{-1}(1-\alpha_1) - R_0 - \sum_{j=1}^{m}\sum_{k=1}^{n}\Big((1+\lambda_{kj})\mathrm{E}_{\Omega}[L_{kj}(w)]\, d_{kj} - C_{kj}\Big) \le 0 \tag{9}$$

$$2.\;\; F_L^{-1}(1-\alpha_2) - \sum_{j=1}^{m}\sum_{k=1}^{n}\Big(\beta(1+\lambda_{kj})\mathrm{E}_{\Omega}[L_{kj}(w)]\, d_{kj} - C_{kj}\Big) \le 0 \tag{10}$$

$$3.\;\; 0 \le d_{kj} \le 0 \quad \text{for all } \; k, j \tag{11}$$

Value of Improved Information

Modeling catastrophe losses has been restrained by the considerable uncertainty included in the models. An analysis done on the quantification of uncertainty in disaster related losses shows the sensitivity of these losses with respect to information [11;12]. A sensitivity analysis is proposed here to deal with the impact of better information. The analysis focuses on the changes in the portfolio parameters with respect to changes in the information accuracy. The larger this sensitivity, the more valuable is the additional information.

Let I be an indexing set for all possible levels of information accuracy then I_i , $i \in \mathrm{I}$ defines the information associated with level i. This data is in a form commensurate with the damage analysis tool (e.g. Hazus). Once the optimal portfolio is attained using the optimization model for a certain level of information, the insurance model is further altered to cater for the impact of true or improved information. The new data produces changes only in covered losses given that the portfolio distribution has been chosen earlier. As an example, if one takes a structure along with a rough estimate of its structural properties, a corresponding optimal coverage ratio can be computed. Suppose now that the actual building state varies from the one adopted in the decision-making process, then the assumed optimal coverage reaches erroneous and misleading outcomes. In that case, the premiums and costs are fulfilled via the data I_1 at hand, whereas the losses are controlled by the actual building data I_a. This divergence translates into the profit as follows,

$$\text{Actual Profit} \;\; = \text{Premiums}\left|I_1\right. - \text{Cost/Fees}\left|I_1\right. \; - \; \text{Covered Losses}\left|I_a\right.$$

$$\text{Premiums } \Pi\left|I_1\right. = \sum_{j=1}^{m}\sum_{k=1}^{n}(1+\lambda_{kj})\mathrm{E}_{\Omega}[L_{kj}(w) d_{kj}]\left|I_1\right. \tag{12}$$

$$\text{Actual Losses } AL\left|I_{1,a}\right. = \sum_{j=1}^{m}\sum_{k=1}^{n} L_{kj}(w)\left|I_a\right. d_{kj}\left|I_1\right. \tag{13}$$

$$\Rightarrow \text{Actual Profit } \; AP\left|I_{1,a}\right. = \sum_{j=1}^{m}\sum_{k=1}^{n}\Big((1+\lambda_{kj})\mathrm{E}_{\Omega}[L_{kj}(w)]\left|I_1\right. - L_{kj}(w)\left|I_a\right.\Big) d_{kj}\left|I_1\right. - C_{kj}\left|I_1\right. \tag{14}$$

In the above equations, the notation $\Pi|I_i$ refers to the value of Π computed using the level of information I_i and the notation $AP|I_{l,a}$ corresponds to the actual profit obtained by taking the portfolio distribution consistent with I_l used in the decision-making process and the losses according to I_a.

The high level detail building data I_a, is in most cases very costly to achieve, and therefore insurers have to settle for less precise data depending on the available resources. This justifies the quest for an optimal data accuracy level given the cost of that information, the associated sensitivity in the predicted losses and its effect on the insurance coverage. Let the optimal portfolio distribution be determined by the information associated with level I_1, then the corresponding profit $P|I_1$ as defined in equation (5) is equal to,

$$P\big|I_1 = \text{Premiums}\big|I_1 - \text{Cost/Fees}\big|I_1 - \text{Covered Losses}\big|I_1$$

$$P\big|I_1 = \sum_{j=1}^{m}\sum_{k=1}^{n}\Big(\Big((1+\lambda_{kj})\,\mathrm{E}_\Omega[L_{kj}(w)] - L_{kj}(w)\Big)d_{kj} - C_{kj}(w)\Big)\Big|I_1 \tag{15}$$

Now, let the level of accuracy under scrutiny be I_i which is considered more refined than I_1. In that case, the profit AP in agreement with I_i is more precise than the one related to I_1 and follows from equation (14),

$$AP\big|I_{1,i} = \text{Premiums}\big|I_1 - \text{Cost/Fees}\big|I_1 - \text{Actual Losses}\big|I_{1,i}$$

$$AP\big|I_{1,i} = \Pi\big|I_1 - C\big|I_1 - AL\big|I_{1,i} \tag{16}$$

$$AP\big|I_{1,i} = \sum_{j=1}^{m}\sum_{k=1}^{n}\Big((1+\lambda_{kj})\,\mathrm{E}_\Omega[L_{kj}(w)\big|I_1\, d_{kj}\big|I_1] - L_{kj}(w)\big|I_i\, d_{kj}\big|I_1\Big) - C_{kj}\big|I_1 \tag{17}$$

$$AP\big|I_{1,i} = \sum_{j=1}^{m}\sum_{k=1}^{n}\Big((1+\lambda_{kj})\,\mathrm{E}_\Omega[L_{kj}(w)]\big|I_1 - L_{kj}(w)\big|I_i\Big)d_{kj}\big|I_1 - C_{kj}\big|I_1 \tag{18}$$

An error measure taking into account the effect of this improved information on the profit is defined,

$$\mathrm{E}_\Omega[AP\big|I_{1,i} - P\,|\,I_1] = \mathrm{E}_\Omega[\sum_{j=1}^{m}\sum_{k=1}^{n}\Big(L_{kj}(w)\big|I_1 - L_{kj}(w)\big|I_i\Big)d_{kj}\big|I_1] \tag{19}$$

By computing this error for the entire information set, the largest deviation underlines the key data that one should seek and thus decide on the information level to be used in the final insurance decision-making process. A negative deviation implies overestimating profit as losses are higher than predicted. On the contrary, a positive deviation results in a higher profit that can be interpreted as overcharging premiums or in a lost opportunity as the coverage could have been larger.

Furthermore, the improved information will reshape the ruin and stability constraints as well. The predicted constraints $RC|I_1$ and $SC|I_1$ follow respectively from equations (9) and (10) for a given level of information I_1 while the more precise constraints AR, AS linked to I_i comply with the following,

$$AR\big|I_{1,i} = AF_L^{-1}(1-\alpha_1)\big|I_{1,i} - R_0 - \sum_{j=1}^{m}\sum_{k=1}^{n}\Big((1+\lambda_{kj})\,\mathrm{E}_\Omega[L_{kj}(w)]d_{kj} - C_{kj}\Big)\Big|I_1 \tag{20}$$

$$AS\big|I_{1,i} = AF_L^{-1}(1-\alpha_2)\big|I_{1,i} - \sum_{j=1}^{m}\sum_{k=1}^{n}\Big(\beta(1+\lambda_{kj})\,\mathrm{E}_\Omega[L_{kj}(w)]d_{kj} - C_{kj}\Big)\Big|I_1 \tag{21}$$

The errors showing the deviation from the anticipated constraints, where a positive variation suggests less safe constraints and possible violation are expressed by,

$$\mathrm{E}_\Omega[AR\big|I_{1,i} - RC\big|I_1] = AF_L^{-1}(1-\alpha_1)\big|I_{1,i} - F_L^{-1}(1-\alpha_1)\big|I_1 \tag{22}$$

$$\mathrm{E}_\Omega[AS\big|I_{1,i} - SC\big|I_1] = AF_L^{-1}(1-\alpha_2)\big|I_{1,i} - F_L^{-1}(1-\alpha_2)\big|I_1 \tag{23}$$

where the distribution function $AF_L|I_{1,i}$ represents the distribution of $AL|I_{1,i}$.

To sum up, the optimization model is carried out according to a specific information accuracy level in order to assess a preliminary optimal distribution. Then the information sensitivity is performed to highlight the deviation from the objectives in case of untruthful inventory and to set the information level that should be considered to determine the final overall optimal portfolio distribution.

IMPLEMENTATION

This section includes the database generation, the information sensitivity analysis and the optimization scheme. In this paper, the model will be implemented for the San Francisco Bay Region. The database of the decision-making model entails first the creation of an earthquake catalog and an exposure inventory to simulate the seismic risk of the region. Then, the information classification necessary for the sensitivity analysis is delineated followed by the loss estimation procedure. After that, the information sensitivity and the insurance optimization model are presented.

Database Generation

This task focuses in turn on the generation of earthquake events, the inventory list, the information set and the economical losses evaluation. It creates a database that shall be used as input for the insurance decision-making process.

1. Earthquake Events Catalog:
   The earthquake catalog is a crucial part in loss estimation, it is responsible for hazard characterization and probabilistic assessments. A report issued by ``U.S. Geological Survey (USGS)'' [13] depicts all possible threatening earthquakes of magnitude 6.7 and larger affecting the San Francisco Bay Area. The analysis is based on the rates of seismic moment accumulation and release across the region. The rupture sources taken into account are San Andreas, Hayward/Rogers Creek, Calaveras, Concord/Green Valley, San Gregorio, Greenville faults and Mount Diablo thrust.
   For the present research, the earthquake catalog is formed on the basis of the USGS report and includes event's rupture location, magnitude and long term rate of occurrence needed to evaluate the hazard parameters across the region.

2. Exposure List:
   The inventory exposure is consistent with Hazus building classification. Hazus is a public domain software prepared by the ``Federal Emergency Management Agency (FEMA)'' [14] for catastrophe loss estimation. A basic exposure database based on the 1990 United States Census is supplied with the software. Also, a recent feature that has been added to the software enables the user to forecast economical losses for a specific set of buildings. The corresponding user-input data consists of:
   - Detailed inventory data for building occupancy type, structural type, location, size and replacement value
   - Performance data for structure's response to earthquake (capacity curves, damage response)
   - Fragility parameters necessary for constructing fragility curves

   The resulting economical losses depend on the information accuracy of the user-specified data and on the hazard parameters.

1. Information Classification:
   The information set I is composed of different information accuracy levels. The accuracy relates to data refinement related to structure inventory or structural properties. Elements of this set are crafted in a way to be compatible with Hazus user-input building data and according to the structures classification. The default data level I_1 is taken from the data supplied with Hazus software. The next

level is designed to match the level of accuracy reached via remote sensing or naturally through traditional field survey mainly for assessing building height and footprint. Many ongoing studies related to remote sensing are focusing on the dissemination of structure type and usage. These properties can be obtained through a combination of remote sensing and statistical analysis such as correlation between the building height and footprint [15].

The levels after that involve field inspection for refinement of construction and design quality. Construction quality is rated as inferior, moderate, superior and design level is rated as low, moderate and high according to the building code requirements. Then, the resolution is again increased to enhance performance data, followed by improved fragility data. Performance data consists of buildings specific capacity curves while fragility data refers to specific fragility curves, both of which are obtained through elaborate engineering studies. The latter will be dealt with in a future study.

Table 1 displays a schematic information catalog, showing the default and improved accuracy at each level. It is important to note that this list can be reworked to suit any required resolution refinement.

Table 1. Information Set

Inf. Level	Inventory	Construction	Design	
I_1	Default	Default	Default	
I_2	Updated	Default	Default	
I_3	Updated	Updated	Default	
I_4	Updated	Updated	Updated	

The means of acquiring the enhanced data as defined for each level will fix the information cost C_{kj} necessary to calculate the insurance profit and constraints. The additional charge for improved inventory information either emanates from the cost of satellite imaging and image processing or from the cost of field surveying. The charge for improved performance and fragility data emanates from the cost of engineering simulations and research.

In this approach, the premiums must have latitude to cater for the cost of information. If the premiums are too low or the information cost is too high relative to the improvement in the losses accuracy, then an additional charge required for the information acquisition might not be justifiable.

4. Loss Estimation:

   The economical losses are estimated using Hazus software. The software estimates the economical losses for a given hazard scenario. It provides damage estimates through built-in or user-supplied capacity curves and fragility curves. A deterministic economical loss is given for each building category.

   As a result, each earthquake event is assigned predicted economical losses from Hazus simulation and a rate of occurrence consistent with USGS report. The loss estimation procedure is considered as a black box that assigns for every specific event and specific building information a deterministic value for the corresponding economical losses.

Information Sensitivity

The sensitivity analysis is implemented for each building type or class independently. First, the economical losses over all events are calculated and the related optimal coverage ratio d is computed in accordance with the optimization formulation in equations (8)-(11) with k,j=1. After that, the deviations from the predicted profit and constraints are evaluated for the different information levels. The error magnitude will identify the main information needed to reduce the risk of erroneous predictions.

It is important to indicate that in the case of small portfolios, the model depicted above can be directly applied. On a regional scale however, the practicality of obtaining information regarding every property and feeding it into the optimization model might not be workable. In this case, the entries (k,j) will no longer represent individual buildings but a collection of properties for each structural type k and location j. The losses linked to each building class portray the total losses of all the properties pertaining to that particular class. Moreover, one might consider to optimize the spatial distribution only and drop the structurally resolved distribution by taking k=1. In this manner, the losses associated with each region represent the total losses of all the buildings in this region.

Optimization Scheme

The optimization is solved using Dakota software prepared by the ``Sandia National Laboratories" [16]. DAKOTA stands for Design Analysis Kit for Optimization and Terascale Applications, it is a toolkit that deals with linear and nonlinear constrained problems. Dakota contains gradient and non-gradient optimization algorithms and several strategies such as multilevel hybrid optimization. Most importantly however, this powerful toolkit provides a flexible interface between users simulation codes and iterative methods. The users can develop their own codes to evaluate objective functions and constraints at each iteration.

Here, a Fortran code is built to generate the cumulative distribution function of the total covered losses $L=\Sigma_j\Sigma_k\, L_{kj}\, d_{kj}$ as well as the necessary statistics. It also evaluates the portfolio's objective function and constraints that feed into Dakota's optimization algorithms. In brief, the approach is,

1. Set initial $\underline{d}$
2. Compute cumulative distribution function and statistics for the covered losses L
3. Calculate objectives and constraints
4. Update $\underline{d}$ using DAKOTA's optimization algorithms

The multilevel hybrid strategy is exploited which allows the use of different algorithms for detecting global optima and local optima. A low fidelity model is adopted for the global search versus a refined model for the local search. The best solution from the global search is used as the starting point for the local one.

In this study, the method of feasible directions for nonlinear programming is used for the local search. The method involves the finding of a feasible descent direction and then finding the related step length along this direction. The process is repeated until termination criterion is satisfied.

As for the global search, to comply with Dakota's algorithms, the penalty function is used and the Fortran code is modified to transform the nonlinear constrained problem into the minimization of $\Psi(\underline{d})$ formed by the initial objective function $P(\underline{d})$ plus a penalty function $f(\underline{d})$. The penalty function ensures the feasibility of the solution by penalizing the objective function when the constraints are violated. When the solution is feasible, the penalty function is ignored. The new problem is formulated as follows,

$$f(\underline{d}) = \sum_i \max(0,\ g_i(\underline{d}))$$

$$\Psi(\underline{d}) = -P(\underline{d}) + c\sum_i \max(0,\ g_i(\underline{d})) \quad \text{with} \quad c > \sum_i |\mu_i|$$

where each g_i represents an inequality constraint and μ_i is the Lagrange multiplier associated with each constraint. The insurance portfolio optimization becomes,

$$\min_{\underline{d}} \quad -\mathrm{E}_\Omega\left[P(\underline{d})\right] + c\left\{\begin{array}{l}\max\left(0, F_L^{-1}(1-\alpha_1) - R_0 - \sum_{j=1}^{m}\sum_{k=1}^{n}(1+\lambda_{kj})E_\Omega[L_{kj}(w)]\, d_{kj} - C_{kj}\right) \\ + \max\left(0, F_L^{-1}(1-\alpha_2) - \sum_{j=1}^{m}\sum_{k=1}^{n}\beta(1+\lambda_{kj})E_\Omega[L_{kj}(w)]\, d_{kj} - C_{kj}\right)\end{array}\right\}$$

$$0 \le d_{kj} \le 1 \qquad \text{for all} \qquad k, j \tag{24}$$

The global search is carried out using genetic algorithms. For a given population size, the algorithm randomly selects solutions and evaluate the corresponding objective function. When the evaluation is complete, the best solutions are retained and used to determine the next set of solutions. This iterative analysis is pursued until the convergence limits are satisfied.

RESULTS

This section depicts the sensitivity analysis results and the optimal portfolio distribution for a hypothetical building exposure in two locations of San Francisco county. The first location corresponds to the center of San Francisco and is positioned at the latitude and longitude coordinates (37.75415; -122.44415) and the second location is closer to the bay and positioned at (37.7700; -122.4300). The exposure list comprises commercial low/mid/high rise and concrete/steel structures and complies with Hazus building data classification.

Sensitivity Analysis Results

The sensitivity analysis is investigated for the different information levels. First, the impact of accurate building footprint and height is examined. In Hazus, losses are proportional to the square footage data and therefore they vary linearly with any additional information on this matter. The building height however, is used to classify the building into low-rise, mid-rise, high-rise depending on the number of stories where an average story height is considered for each building type. For illustration purposes, the high-rise class is made of buildings of eight stories and higher and is represented by an average structure of twelve stories. This implies that all the buildings belonging to that class have the same structural behavior as the average structure. Also, any discrepancy in height measurement might swing the building from a high-rise to a mid-rise. As a result, the footprint and height information are essential pieces of information in conducting accurate loss evaluation.

Second, the construction quality and design level are incorporated into the model. Table 2 describes the related information classification along with the improved data.

Table 2. Information Levels Description

Information Level	Const. Data Accuracy	Construction Status	Design Data Accuracy	Design Status	Expected Losses
I_11	Default	Moderate	Default	Moderate	3503
I_21	Inspected	Superior	Default	Moderate	1649
I_22	Inspected	Inferior	Default	Moderate	5060
I_31	Default	Moderate	Inspected	High	2471
I_32	Default	Moderate	Inspected	Low	5399
I_41	Inspected	Superior	Inspected	High	1395
I_42	Inspected	Superior	Inspected	Low	2978
I_43	Inspected	Inferior	Inspected	High	3568
I_44	Inspected	Inferior	Inspected	Low	7120

The sensitivity analysis is carried out for a concrete mid-rise commercial type building. The parameters used in the calculations are R_0 =10000, λ = 2.5, α_1 = 1%, α_2 = 4% and β = 95%. The information cost on the other hand is considered as linear with respect to information where an arbitrary unit cost is equal to 200 and any single data improvement is assigned one unit. The optimal coverage distribution using the different information levels in the decision-making phase is displayed in figure 1. In the graph, Lc_1 and Lc_2 refer to the location designation.

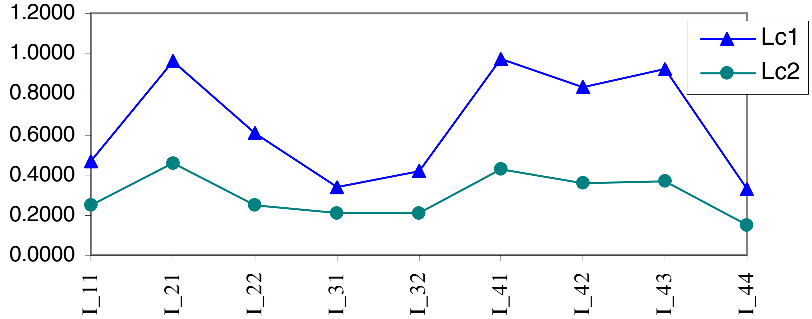

Figure 1. Coverage Ratios vs. Information Levels

Figures 2-4 describe the impact of enhanced data on the predicted profit and constraints. The improved and erroneous legends identify respectively the objective and constraints computed using equations (8-11) for each investigated data level I_i and the ones computed with erroneous data in accordance with equations (16-21) for $AP|I_{1,i}$, $AR|I_{1,i}$ and $AS|I_{1,i}$.

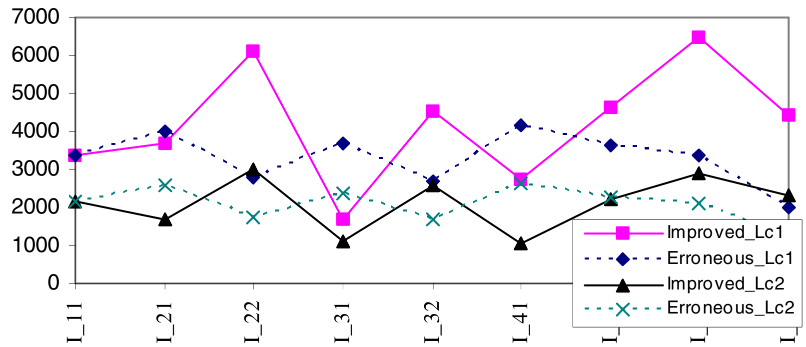

Figure 2. Objective Function vs. Information Levels

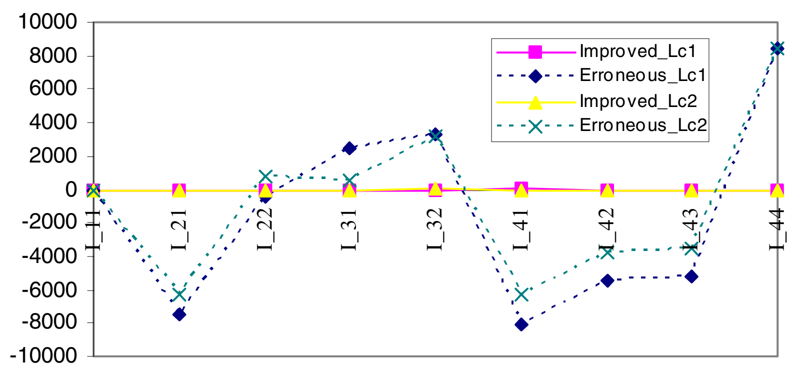

Figure 3. Ruin Constraint vs. Information Levels

Figure 5 shows the sensitivity of the objective function according to equation (19). Here, in the one-dimensional case, the stability constraint is always satisfied and the insolvency constraint is the critical restraint. The insolvency constraint is the one that determines the coverage ratio implying that it is equal to zero over all information levels in the improved data case and that the graph for the sensitivity is the same as figure 3.

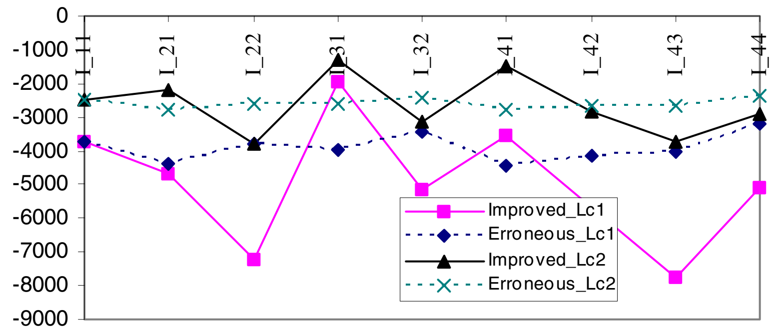

Figure 4. Stability Constraint vs. Information Levels

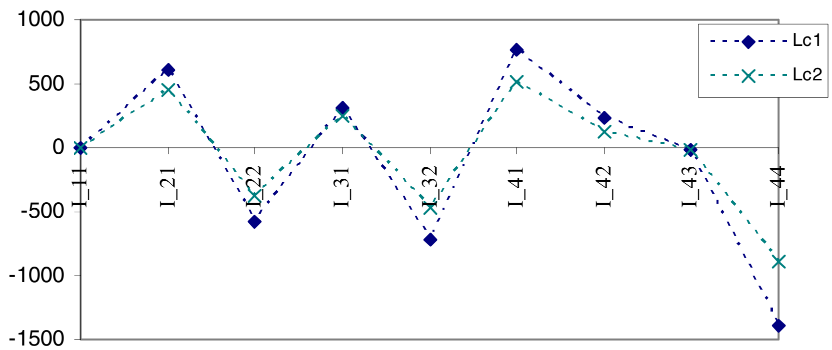

Figure 5. Sensitivity in Objective Function

The graphs make obvious that the objectives vary greatly with improved data. For example, if after inspection the construction quality turns out bad, the results in agreement with I_{22} point at an almost -500 or 20% loss in profit and constraints violation due to unanticipated large losses in comparison to the forecasted objectives based on the default data. On the opposite, if the quality status is good I_{21}, the analysis leads to almost 20% increase in profit and safer constraints. This positive outcome however is misleading, as in this case the premiums are overpriced and the insurance reserve is much higher than the one needed to avoid insolvency. As an alternative, the coverage ratio should have been higher. Lastly, if the investigated construction quality is moderate, consistent with its default data, then the simulated losses are unchanged but the insurer would still have to bear the cost of inspecting the building. This analysis clearly exhibits the untrue objectives obtained with low level detail data and the significance of improved information.

Optimal Portfolio Distribution

In this paragraph, the spatially and structurally resolved portfolio distribution is denoted. The model is executed on the hypothetical exposure using the default data and improved construction data. Other information levels can also be treated in a similar manner and analyzed following the same decision-making approach. Table 3 shows the portfolio exposure and distribution considering the default information level where the buildings construction quality status is taken as moderate. The same parameter values used for the sensitivity analysis are considered except for R_0 taken equal to 1.5E5.

Table 3. Portfolio Exposure and Distribution – Default Level

Building #	Building Classification (k,j)	Building Type	Building Specification	Expected Losses	Coverage Ratio
1	1,1	Mid-rise	Steel Moment Frame	3688	1.000
2	2,1	Mid-rise	Concrete Moment Frame	3221	1.000
3	3,1	Low-rise	Concrete Shear Walls	4404	0.000
4	4,1	Mid-rise	Concrete Shear Walls	2935	1.000
5	5,1	High-rise	Concrete Shear Walls	2549	1.000
6	1,2	Mid-rise	Steel Moment Frame	4444	1.000
7	2,2	Mid-rise	Concrete Moment Frame	3876	1.000
8	3,2	Low-rise	Concrete Shear Walls	4466	0.000
9	4,2	Mid-rise	Concrete Shear Walls	3503	0.000
10	5,2	High-rise	Concrete Shear Walls	3127	0.485
				Profit	***55,576***

This table defines the buildings coverage ratios for the two locations and over all structural types revealing that some structures such as mid-rise steel moment frames should be fully insured whereas others should be totally abandoned.

Table 4 lays the distribution of the same portfolio after investigating the construction quality of the buildings and updating the related loss data. The chosen construction quality for each structure is arbitrary and is only used to highlight the impact of improved information accuracy.

Table 4. Portfolio Exposure and Distribution – Updated Construction Quality

Building #	Building Classif. (k,j)	Building Type	Building Specification	Updated Const. Quality	Expected Losses	Coverage Ratio
1	1,1	Mid-rise	Steel Moment Frame	Low	5650	1.000
2	2,1	Mid-rise	Concrete Moment Frame	Mod.	3221	1.000
3	3,1	Low-rise	Concrete Shear Walls	Mod.	4404	0.006
4	4,1	Mid-rise	Concrete Shear Walls	High	1602	1.000
5	5,1	High-rise	Concrete Shear Walls	Mod.	2549	1.000
6	1,2	Mid-rise	Steel Moment Frame	Low	6999	0.213
7	2,2	Mid-rise	Concrete Moment Frame	Low	6262	1.000
8	3,2	Low-rise	Concrete Shear Walls	Mod.	4466	0.009
9	4,2	Mid-rise	Concrete Shear Walls	High	1649	0.446
10	5,2	High-rise	Concrete Shear Walls	High	1945	0.992
					Profit	***56,761***

In the case of good construction quality, the simulated losses are less than the ones attained for the moderate case, and hence these structures are most likely subject to an increase in the insurance coverage. For instance, at location 2 the mid-rise concrete shear wall structure has experienced a coverage increase from 48.5% to 99.2% due to good construction, while the mid-rise steel frame coverage changed from full coverage to 21% because of low construction quality. This change in coverage depends partly on the expected losses but mainly on the total loss distribution as imposed by the constraints. As for the profit, the result from the updated data after accounting for all information cost is almost the same as the one obtained from the default data. Consequently, the information acquisition is justified due to the fact that the inflicted cost is recovered and that the end result yields more accuracy in predictions.

CONCLUSION

Models for insurance decision-making are proposed involving the value of enhanced information and portfolio optimization. Results from the sensitivity analysis demonstrate that the use of improved information acquired from advanced technologies, engineering studies and field survey affect greatly the insurance catastrophe portfolio predictions. It also points to the fact that an optimal portfolio distribution based on inaccurate data might lead to erroneous and misleading outcomes. Furthermore, the chance constrained optimization establishes the importance of the portfolio distribution among different structure types and locations in order to avoid insolvency and control business stability. The model highlights the role of the structural vulnerability of the buildings and their particular ability to withstand earthquakes.

Moreover, given the fact that high level detail data is most of the time very costly to achieve, insurers have to search for an optimal accuracy level to use in their decision-making process depending on the cost of that information, and its impact on the end results. In light of that, incorporating enhanced data into the insurance portfolio selection criterion would minimize the uncertainty in the predicted profit and constraints and ensure that the anticipated objectives are close to the factual ones.

REFERENCES

1. Bell E.J., Westerlund F.V. “Inventorying Building and Critical Structures Using Statistical Sampling and Remote Sensing” University of Washington Department of Urban Design and Planning
2. Tralli David M. “Assessment Of Advanced Technologies for Loss Estimation” Multidisciplinary Center for Earthquake Engineering Research Report MCEER-00-SP02, 2002
3. Amendola A., Ermoliev Y., Ermolieva T.Y., Gitis V., Koff G., and Linnerooth-Bayer J. “A Systems Approach to Modeling Catastrophic Risk and Insurability” Natural Hazards 21:381-393,2000. 2000 Kluwer Academic Publishers
4. Ermoliev Y.M., Ermolieva T.Y. and Macdonald G.J. “Stochastic Optimization of Insurance Portfolios for Managing Exposure to Catastrophic Risks” Annals of Operations Research 99, 207-225, 2000. 2001 Kluwer Academic Publishers
5. Kleindorfer Paul R., Klein Robert W. “Regulation and Markets for Catastrophe Insurance” Murat R. Sertel and Semih Koray, Advances in Economic Design, Springer Verlag, Heidelberg (2003)
6. Stone J.M. “A Theory of Capacity and The Insurance of Catastrophe Risks, Parts 1 and 2” The Journal of Risk and Insurance 40, 231-244 and 339-355 (1973)
7. Kenneth Froot “The Financing of Catastrophe Risk” University of Chicago Press (1999)
8. Birge John R., Louveaux Francois “Introduction to Stochastic Programming” Springer Series in Operations Research (1997)
9. Cooper Leon and Steinberg David “Introduction to Methods of Optimization” W.B. Saunders Company (1970)

10. Rustagi Jagdish S. “Optimization Techniques in Statistics” Academic Press (1994)
11. Hayek Carol, Ghanem Roger “Impact of Uncertainty in Catastrophe Losses on Insurance Derivatives” Proceeding of 15th ASCE Engineering Mechanics Conference, Columbia University New York, NY
12. Grossi Patricia “Quantifying the Uncertainty in Seismic Risk & Loss Estimation” Presented at IIASA Euro-Conference (2000)
13. “Earthquake Probabilities in the San Francisco Bay Region: 2002-2031” U.S. Geological Survey WG02 Open-File Report 03-214 (2003)
14. “Earthquake Loss Estimation Methodology - HAZUS 99 Technical Manual” Federal Emergency Management Agency and National Institute of Building Sciences
15. Eguchi R.T., Huyck C.K., Houshmand B., Tralli D.M., Shinozuka M. “A New Application for Remotely Sensed Data: Construction of Buildings Inventories using Synthetic Aperture Radar Technology” MCEER Publication presented at the 2nd Multi-Lateral Workshop on Development of Earthquake and Tsunami Disasters Mitigation Technologies and Their Integration for the Asia-Pacific Region, Japan 2000
16. “DAKOTA: A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis” Sandia National Laboratories (2003)