Chapter 2 – An Introduction to Catastrophe Models and Insurance

Major Contributors:
Patricia Grossi
Howard Kunreuther
Don Windeler

This chapter provides an overview of the history of catastrophe models and their role in risk assessment and management of natural disasters. It examines the insurability of catastrophe risk and illustrates how the output from catastrophe models aids insurers in meeting their goals for risk management. Throughout the chapter, there is an emphasis on understanding catastrophe modeling for earthquake and hurricane hazards and how it is used to manage natural hazard risk. In the final section, a framework for integrating risk assessment with risk management via catastrophe modeling is presented.

2.1 History of Catastrophe Models

Catastrophe modeling is not rooted in one field or discipline. The science of assessing and managing catastrophe risk originates in the fields of property insurance and the science of natural hazards. Insurers may well argue that catastrophe modeling's history lies in the earliest days of property insurance coverage for fire and lightning. In the 1800's, residential insurers managed their risk by mapping the structures that they covered. Not having access to Geographic Information Systems (GIS) software, they used tacks on a wall-hung map to indicate their concentration of exposure. This crude technique served insurers well and limited their risk. Widespread usage of mapping ended in the 1960's when it became too cumbersome and time-consuming to execute (Kozlowski and Mathewson, 1995).

On the other hand, a seismologist or meteorologist may well argue that the origin of catastrophe modeling lies in the modern science of understanding the nature and impact of natural hazards. In particular, the common practice of measuring an earthquake's magnitude and a hurricane's intensity is one of the key ingredients in catastrophe modeling. A standard set of metrics for a given hazard must be established so that risks can be assessed and managed. This measurement began in the 1800's, when the first modern seismograph (measuring earthquake ground motion) was invented and modern versions of the anemometer (measuring wind speed) gained widespread usage.

In the first part of the twentieth century, scientific measures of natural hazards advanced rapidly. By the 1970's, studies theorizing on the source and frequency of events were published. Significant analyses include the U.S. Water Resources Council publication on flood hazard (USWRC, 1967), the Algermissen study on earthquake risk (Algermissen, 1969) and National Oceanic and Atmospheric Administration (NOAA) hurricane forecasts (Neumann, 1972). These developments led U.S. researchers to compile hazard and loss studies, estimating the impact of earthquakes, hurricanes, floods, and other natural disasters. Notable compilations include Brinkmann's summary of hurricane hazards in the United States (1975) and Steinbrugge's anthology of losses from earthquakes, volcanoes, and tsunamis (1982).

These two separate developments – mapping risk and measuring hazard – came together in a definitive way in the late 1980's and early 1990's, through catastrophe modeling as shown in Figure 2.1. Computer-based models for measuring catastrophe loss potential were developed by linking scientific studies of natural hazards' measures and historical occurrences with advances in information technology and geographic information systems (GIS). The models provided estimates of catastrophe losses by overlaying the properties at risk with the potential natural hazard(s) sources in the geographic area. With the ability to store and manage vast amounts of spatially referenced information, GIS became an ideal environment for conducting easier and more cost-effective hazard and loss studies.

Around the same time, several new modeling firms developed computer software for analyzing the implications of natural hazard risk. Three major firms emerged: AIR Worldwide was founded in 1987 in Boston; Risk Management Solutions (RMS) was formed in 1988 at Stanford University; and EQECAT began in San Francisco in 1994 as a subsidiary of EQE International. In 2001, EQE International became a part of ABS Consulting.

When introduced, the use of catastrophe models was not widespread. In 1989, two large-scale disasters occurred that instigated a flurry of activity in the advancement and use of these models. On September 21, 1989, Hurricane Hugo hit the coast of South Carolina, devastating the towns of Charleston and Myrtle Beach. Insured loss estimates totaled $4 billion before the storm moved through North Carolina the next day (Insurance Information Institute, 2000). Less than a month later, on October 17, 1989, the Loma Prieta Earthquake occurred at the southern end of the San Francisco peninsula. Property damage to the surrounding Bay Area was estimated at $6 billion (Stover and Coffman, 1993).

These two disasters sent a warning signal to the insurance industry. On the heels of these two events, Hurricane Andrew made landfall in Southern Florida in August of 1992. Within hours of landfall, AIR Worldwide issued a fax to its clients to the effect that losses, as estimated in real time by the AIR Worldwide hurricane model, might reach the astonishing amount of $13 billion. It was not until months later that the final tally, $15.5 billion, was issued by the Property Claim Services Office.

Nine insurers became insolvent as a result of their losses from Hurricane Andrew. Insurers and reinsurers realized that, in order to remain in business, they needed to estimate and manage their natural hazard risk more precisely. Many companies turned to the modelers of catastrophe risk for decision support. The modeling companies grew and catastrophe models increased in number, availability, and capability. By 2001, other organizations joined these front-runners in developing catastrophe models for assisting insurers and reinsurers in pricing their insurance policies and determining how much coverage to offer in hazard-prone areas of the country.

The series of natural disasters in 1989 and 1992 also sent a warning signal to the public sector of the United States. The government recognized the need for an accurate assessment of the impact of disasters for mitigation and emergency planning purposes. In 1992, the Federal Emergency Management Agency (FEMA) funded a study to assess the latest loss estimation methodologies for earthquakes. The agency issued a report in 1994 on the results of this study entitled: Assessment of the State of the Art Earthquake Loss Estimation Methodologies (FEMA 249, 1994).

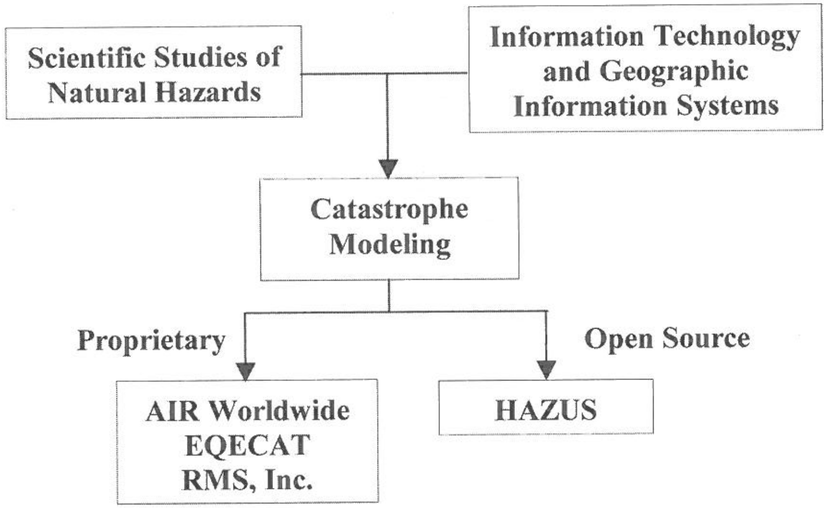

Figure 2.1. Development of catastrophe modeling.

This study convinced FEMA to fund the development of "Hazards U.S." (HAZUS), a catastrophe model in the public domain. HAZUS is labeled as an open source model in Figure 2.1. From the outset, one of FEMA's goals was to create a methodology that was the "standard national loss methodology for assessing losses from natural hazards" (FEMA, 2002). The first version of HAZUS was developed with a combination of public and private resources to estimate earthquake losses and was released in 1997 (NIBS, 1997). Updates to the HAZUS earthquake model have been in the form of data and software integration; methodologically, the software remains the same. In 2004, the latest HAZUS multi-hazard methodology, relabeled HAZUS-MH, integrates the earthquake module with two new modules for estimating potential losses from wind and flood (riverine and coastal) hazards.

2.2 Structure of Catastrophe Models

The four basic components of a catastrophe model are: hazard, inventory, vulnerability, and loss as depicted in Figure 2.2. First, the model characterizes the risk of natural hazard phenomena. For example, an earthquake hazard is characterized by its epicenter location and moment magnitude, along with other relevant parameters. A hurricane is characterized by its projected path and wind speed. The frequency of certain magnitudes or frequencies of events also describes the hazard in question.

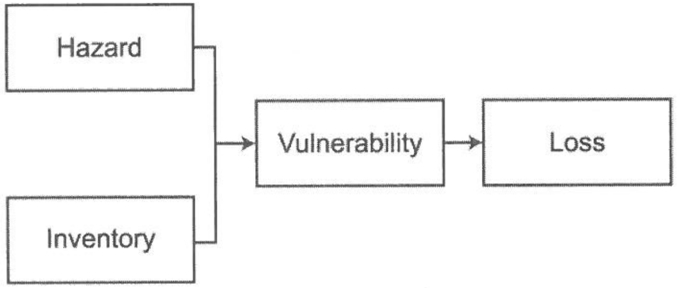

Figure 2.2. Structure of catastrophe models.

Next, the model characterizes the inventory or portfolio of properties at risk as accurately as possible. Arguably, the most important parameter used to characterize the inventory is the location of each property at risk. A process called geocoding is normally used to assign geographic coordinates such as latitude and longitude to a property based on its street address, ZIP code or another location descriptor. With a property's location in spatial terms, other factors that could aid in estimating the vulnerability of a property are added to its characterization. For a building, these parameters include such features as its construction type, the number of stories in the structure, and its age. If the property is insured, information on the nature of the policy, such as the deductible and coverage limit, is also recorded.

The hazard and inventory modules enable the calculation of the vulnerability or susceptibility to damage of the structures at risk. In essence, this step in the model quantifies the physical impact of the natural hazard phenomenon on the property at risk. How this vulnerability is quantified differs from model to model. For example, the HAZUS model classifies a structure as being in a Slight, Moderate, Extensive, or Complete damage state. Other models construct damage curves and relate structural damage to a severity parameter, such as peak gust wind speed or spectral acceleration. In all models, damage curves are constructed for the building, its contents and time element losses, such as business interruption loss or relocation expenses.

From this measure of vulnerability, the loss to the inventory is evaluated. In a catastrophe model, loss is characterized as direct or indirect in nature. Direct losses include the cost to repair and/or replace a structure. Indirect losses include business interruption impacts and relocation costs of residents forced to evacuate their homes. Proprietary models include the ability to analyze insurance policies, so that the loss can be properly allocated. More details on these elements of a catastrophe model are provided in Chapter 3.

2.3 Uses of a Catastrophe Model for Risk Management

A catastrophe model is employed to assess catastrophe risk and improve risk management decisions. But how is this accomplished? Briefly, the model output is quantified and presented in a way that is useful to the stakeholder. Once these metrics are in hand, alternate risk management strategies, such as mitigation, insurance, reinsurance and catastrophe bonds, can be assessed. Currently, insurers and reinsurers are the stakeholders with the most widespread interest and integrated use of catastrophe models. Reinsurance brokers in particular have enhanced the use of catastrophe models. It is fairly common for a broker to collect data for potential clients, run the models on that data, and provide the output to interested reinsurers.

The capital markets have also been eager users of this technology in order to more accurately price catastrophe bonds. In fact, their recent interest and involvement in natural hazards have been made possible by the quantification afforded by catastrophe modeling. Property owners are less likely to use catastrophe models themselves, but their decision processes are directly or indirectly influenced by the outcomes. At the governmental level, catastrophe modeling presents both a positive opportunity and a political dilemma for regulators and emergency management agencies.

As an example of a positive use of the models, consider the use of HAZUS to measure the impact of an earthquake. One model output option is to create a GIS map of the potential loss. Given the definition of the hazard, including the earthquake's epicenter location, and the concentration of the properties at risk, Figure 2.3 depicts a map of the displaced households for the Charleston, South Carolina region subject to an M 7.3 earthquake. The largest concentration of loss, measured by the number of individuals seeking shelter following the disaster, is near the scenario's epicenter. This map is potentially useful to emergency response and recovery officials responding to a disaster.

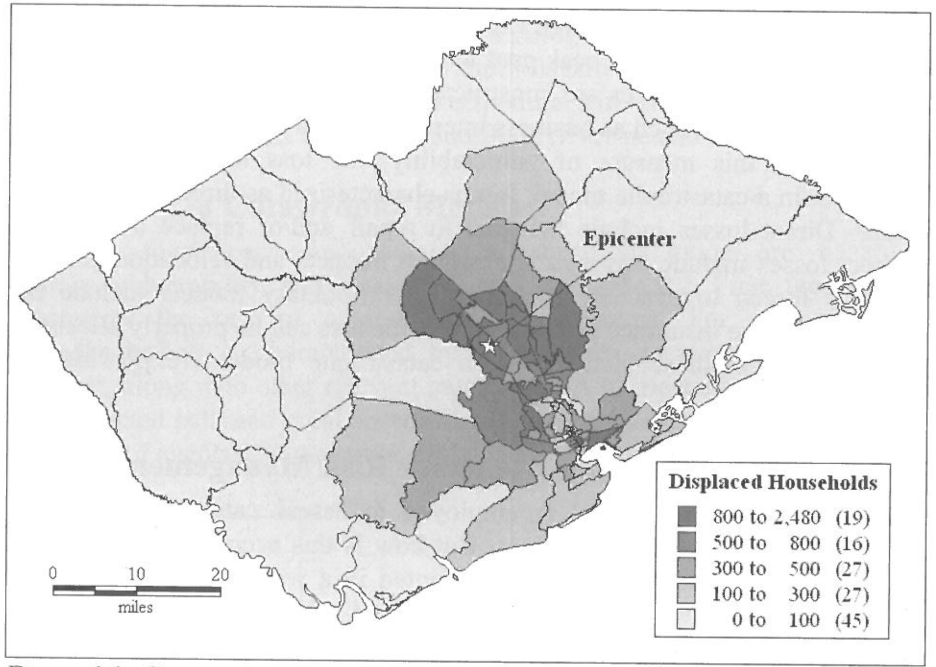

Figure 2.3. Catastrophe model output: Map of shelter requirements predicted by HAZUS for M 7.3 events in Charleston, South Carolina region.

Another output option is the exceedance probability (EP) curve. For a given portfolio of structures at risk, an EP curve is a graphical representation of the probability that a certain level of loss will be surpassed in a given time period. Special attention is given to the right-hand tail of this curve where the largest losses are situated. Figure 2.4 depicts an EP curve for an insurer with a portfolio of residential earthquake policies in Long Beach, California. In contrast to a GIS map of loss, which presents loss in a spatial manner, an exceedance probability curve portrays loss in a temporal manner.

An EP curve is particularly valuable for insurers and reinsurers to determine the size and distribution of their portfolios' potential losses. Based on the EP curve, they can determine the types and locations of buildings they would like to insure, what coverage to offer, and what price to charge. To keep the probability of insolvency at an acceptable level, insurers can also use an EP curve to determine what proportion of their risk needs to be transferred to either a reinsurer and/or the capital markets.

For example, suppose an insurer in Long Beach offers residential earthquake coverage and the insurer's exceedance probability curve for its portfolio is as depicted in Figure 2.4. Further suppose the insurer specifies $10 million as an acceptable level of loss at a 1% (1-in-100) probability of exceedance. Based on the graph, it can be seen that loss profile of the current portfolio would be unacceptable since the 1-in-100 loss for the portfolio is $15 million. The insurer would need to look for ways to reduce its portfolio, transfer $5 million of loss to a reinsurer, or purchase a catastrophe bond to cover it.

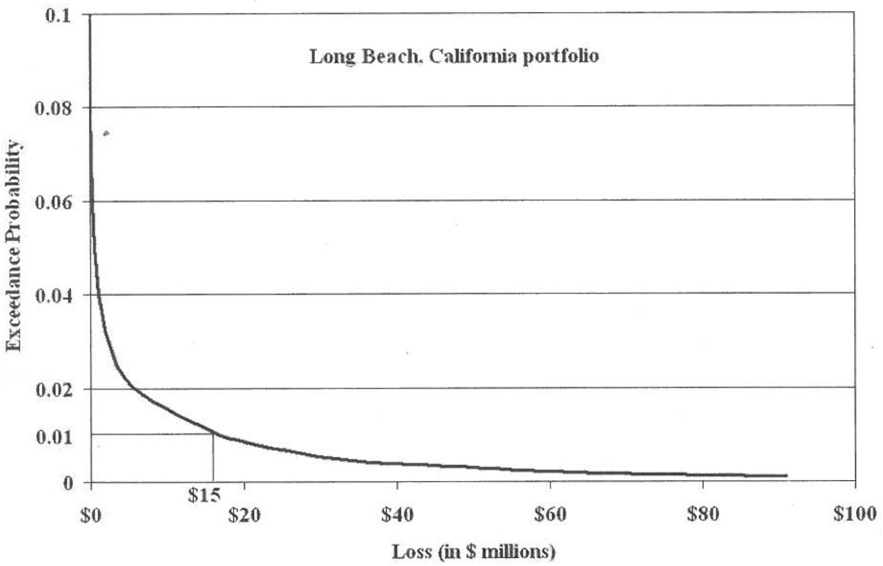

Figure 2.4. Catastrophe model output: Right-hand tail of exceedance probability curve predicted by EQECAT for all possible events.

2.4 Derivation and Use of an Exceedance Probability Curve

Given the importance of how insurers use catastrophe modeling and the EP curve to manage risk, it is essential to understand how the EP curve can be created from the loss output.

2.4.1 Generating an Exceedance Probability Curve

For the purposes of illustration, some simplifying assumptions are made to generate an EP curve. Suppose there is a set of natural disaster events, E_i, which could damage a portfolio of structures. Each event has an annual probability of occurrence, p_i, and an associated loss, L_i. The number of events per year is not limited to one; numerous events can occur in the given year. A list of 15 such events is listed in Table 2.1, ranked in descending order of the amount of loss. In order to keep the example simple and calculations straightforward, these events were chosen so the set is exhaustive (i.e., sum of the probabilities for all of the events equals one).

The events listed in Table 2.1 are assumed to be independent Bernoulli random variables, each with a probability mass function defined as:

$$P(E_i \text{ occurs}) = p_i$$

$$P(E_i \text{ does not occur}) = (1 - p_i)$$

If an event E_i does not occur, the loss is zero. The Expected Loss for a given event, E_i, in a given year, is simply:

$$E[L] = p_i L_i$$

The overall expected loss for the entire set of events, denoted as the average annual loss (AAL) in Table 2.1, is the sum of the expected losses of each of the individual events for a given year and is given by:

$$AAL = \sum_i p_i L_i$$

Assuming that during a given year, only one disaster occurs, the exceedance probability for a given level of loss, $EP(L_i)$, can be determined by calculating:

$$EP(L_i) = P(L > L_i) = 1 - P(L \le L_i)$$

$$EP(L_i) = 1 - \prod_{j=1}^{i} (1 - p_j)$$

The resulting exceedance probability is the annual probability that the loss exceeds a given value. As seen in the equation above, this translates into one minus the probability that all the other events below this value have not occurred. The exceedance probability curve for the events in Table 2.1 is shown in Figure 2.5. Sidebar 1 explains how the EP curve can be used to determine probable maximum loss (PML).

Table 2.1. Events, Losses, and Probabilities

Event (E_i)	Annual probability of occurrence (p_i)	Loss (L_i)	Exceedance probability [$EP(L_i)$]	$E[L] = (p_i * L_i)$
1	0.0020	$25,000,000	0.0020	$50,000
2	0.0050	15,000,000	0.0070	75,000
3	0.0100	10,000,000	0.0169	100,000
4	0.0200	5,000,000	0.0366	100,000
5	0.0300	3,000,000	0.0655	90,000
6	0.0400	2,000,000	0.1029	80,000
7	0.0500	1,000,000	0.1477	50,000
8	0.0500	800,000	0.1903	40,000
9	0.0500	700,000	0.2308	35,000
10	0.0700	500,000	0.2847	35,000
11	0.0900	500,000	0.3490	45,000
12	0.1000	300,000	0.4141	30,000
13	0.1000	200,000	0.4727	20,000
14	0.1000	100,000	0.5255	10,000
15	0.2830	0	0.6597	0
Average Annual Loss (AAL) =				$760,000

SIDEBAR 1: PML as a function of the EP Curve

The exceedance probability curve illustrated in Figure 2.5 enables an insurer to determine his PML or Probable Maximum Loss for a portfolio of structures in a given time period. The term PML is a subjective risk metric and is associated with a given probability of exceedance specified by the insurer. For example, suppose that an insurer specifies its acceptable risk level as the 0.4% probability of exceedance. The insurer can use the EP curve to determine how large a loss will occur at this probability level. Often, PML limits are framed in terms of a return period. The return period is simply the inverse of the annual probability of exceedance. In this example, a 1-in-250 year PML is the lower limit on the loss at a 0.4% probability of exceedance on the EP curve. From the inset of Figure 2.5, it can be seen that the PML is approximately $21 million.

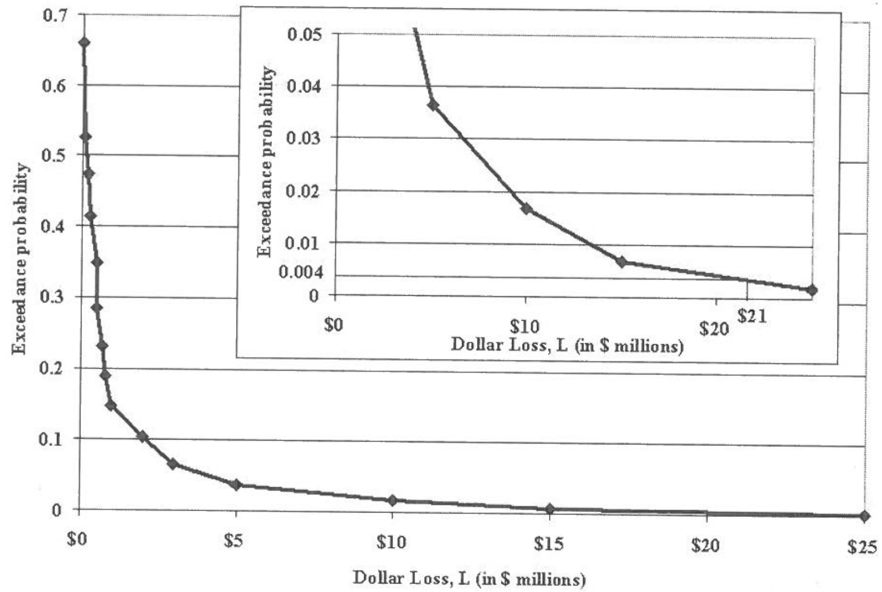

Figure 2.5. Exceedance probability curve

2.4.2 Stakeholders and the Exceedance Probability Curve

The exceedance probability curve can also be used to distribute the losses between stakeholders. Suppose there are three stakeholders who share the losses from a particular disaster. The owner retains the first part of the loss, a second party covers the middle portion and a third party covers the extreme portion. This scenario could represent a portfolio of homes with the homeowners having deductibles on their insurance policies such that they cover the first portion of the loss, an insurer covers the middle portion and a reinsurer handles the losses above a certain amount. Figure 2.6 shows a simple illustrative example. The potential loss for a portfolio with a total value of $100 million is split between three participants: P1, P2, and P3. The first $5 million of loss (L1) would be borne by P1 (homeowners), losses between $5M and $30M (L2) by P2 (insurer), and losses in excess of $30M (L3) by P3 (reinsurer). If the events facing the three parties were those given in Table 2.1, then the reinsurer would never experience any claim payments because the maximum loss would be $25 million.

Now suppose the three parties face the set of events in Table 2.1, but there is some uncertainty associated with the losses from each of the first 14 events (E_{15} has a loss of zero). In other words, the losses in Table 2.1 represent the mean estimates of loss; each event E_i has a distribution of loss associated with it. There is now a range of possible outcomes for each event, and some of these will penetrate the higher layer L3 (Figure 2.7). By combining the loss distributions for all the events, the probability of exceeding a specific loss level can be calculated. This then becomes the basis for developing EP curves for each of the parties with resources at risk.

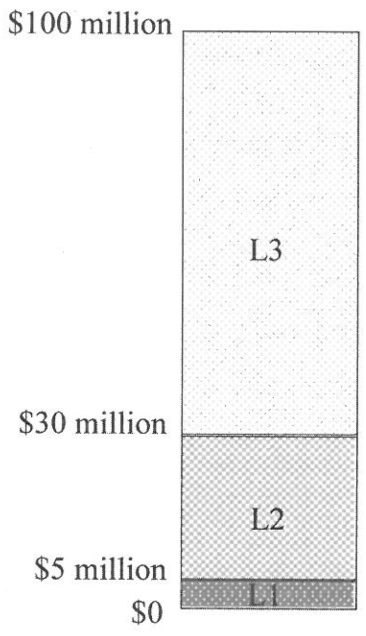

Figure 2.6. Layering for hypothetical portfolio, total value $100 million.

Figure 2.7 shows a set of loss-causing events with a high level of uncertainty in the loss distributions where the coefficient of variation (CV) on the event losses is 1.0.[1] By definition, the coefficient of variation is the ratio of the standard deviation to the mean. The effect of this high uncertainty is clearest on L3. If there were no variability in the losses, L3 would not be affected because no event is capable of reaching a $30 million loss, as previously stated. Based on the assumption (CV = 1.0), there is an annual probability of 0.28% that an event would cause some loss to L3.

This illustrative example shows how catastrophe modeling provides a means of both quantifying risks and allocating them among stakeholders. Using these metrics, it is possible to make rational, informed decisions on how to price risks and determine how much coverage is needed based on an

---

[1] Note that the assumption of a constant coefficient of variation for all events is not realistic and is used only for ease of illustration. The CV on the event loss generally decreases as the size of the loss increases; a portfolio CV of 1.0 for the most damaging event in this example is highly unlikely.

acceptable level of risk. However, there are uncertainties inherent in the catastrophe modeling process that can have a large impact on the distribution of risk among stakeholders. The quantification and disaggregation of uncertainty provides opportunities for stakeholders to reduce risk. As will be discussed in Part II, some of this uncertainty can be reduced by better data, but a significant component is an intrinsic part of the physical process.

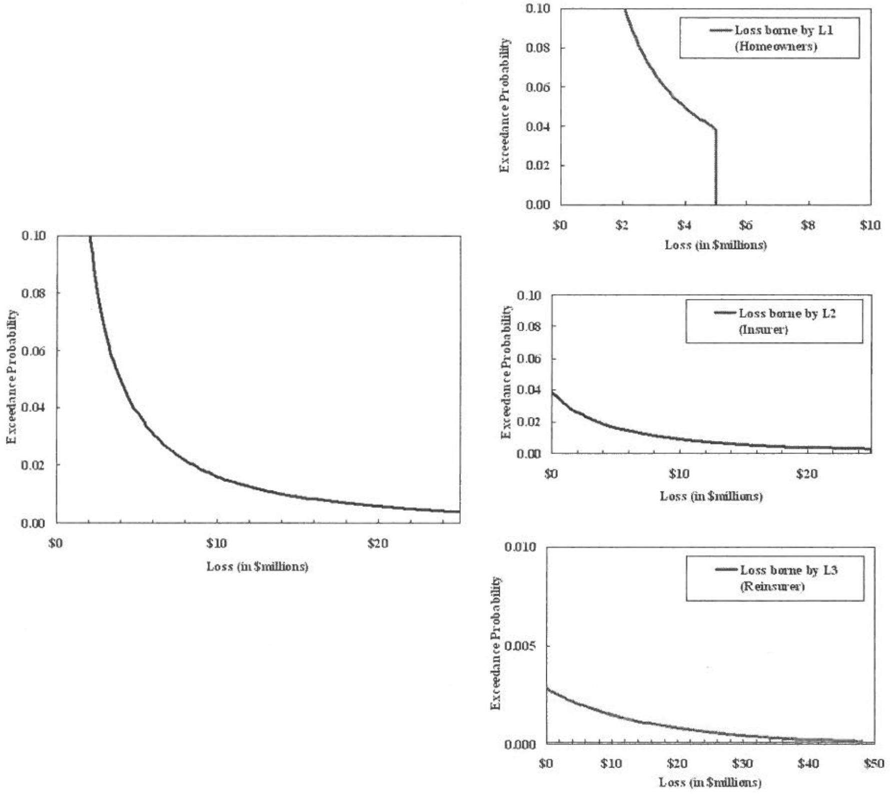

Figure 2.7. Exceedance probability curves for total portfolio and individual participants.

2.5 Insurability of Catastrophe Risks

In most developed countries, insurance is one of the principal mechanisms used by individuals and organizations to manage risk. Insurance allows the payment of a relatively small premium for protection against a potentially large loss in the future. In the United States, some property insurance coverage is required by law or by the lending institution. For example, homeowners normally have to purchase fire coverage as a condition for a mortgage. Automobile liability insurance is also required in most states as a condition for licensing a car. However, earthquake insurance is usually not required by lenders on single-family residences.

Insurance pricing can be a signal of how risky certain activities are for a particular individual. To illustrate, consider automobile insurance. For cars that are the same price, younger, inexperienced drivers of sporty vehicles pay more in premiums than older drivers of more conservative cars. For life and health insurance, smokers pay more for coverage than nonsmokers. This allocation of risk seems appropriate since it is tied to the likelihood of outcomes resulting from the nature of an individual's lifestyle. If one individual is more susceptible to a specific risk, then the cost for coverage against a loss from that risk is greater. Of course, since insurance rates are subject to regulation, the price of the policy may not fully reflect the underlying risk.

The key challenge is how to allocate catastrophe risk among stakeholders in a manner similar to what is done for more frequent, non-extreme events. For automobile coverage, considerable historical data are available and utilized to estimate insurance premiums for individuals with different risk characteristics. The large number of data points and the absence of correlation between accidents allow the use of actuarial-based models to estimate risk (Panjer and Willmot, 1992). With respect to natural disasters, there are limited data available to determine the probabilities of events occurring and their likely outcomes. In the absence of past data, there is a need for insurers to model the risk. Catastrophe models serve this purpose by maximizing the use of available information on the risk (hazard and inventory) to estimate the potential losses from natural hazards.

2.5.1 Conditions for Insurability of a Risk

Consider a standard insurance policy whereby premiums are paid at the start of a given time period to cover losses during this interval. Two conditions must be met before insurance providers are willing to offer coverage against an uncertain event. The first condition is the ability to identify and quantify, or estimate at least partially, the chances of the event occurring and the extent of losses likely to be incurred. The second condition is the ability to set premiums for each potential customer or class of customers.

If both conditions are satisfied, a risk is considered to be insurable. But it still may not be profitable. In other words, it may be impossible to specify a rate for which there is sufficient demand and incoming revenue to cover the development, marketing, operating, and claims processing costs of the insurance and yield a net positive profit over a prespecified time horizon. In such cases, the insurer will opt not to offer coverage against this risk.

To satisfy the first condition, estimates must be made of the frequency of specific events and the likely extent of losses. Such estimates can be based on past data or catastrophe modeling, coupled with data on what experts know about a particular risk. The insurer can then construct an exceedance probability (EP) curve that depicts the probability that a certain level of loss will be exceeded on an annual basis.

With respect to the second condition, if there is considerable ambiguity or uncertainty associated with the risk, insurers may wish to charge a much higher premium than if they had more precise estimates of the risk (Kunreuther, Hogarth and Meszaros, 1995). Moreover, if the capacity of the insurance industry is reduced due to recent large losses, then premiums will rise due to a shortage in supply. The situation will be exacerbated if the recent losses trigger an increase in demand for coverage, as was the case after Hurricane Andrew in 1992 and the Northridge earthquake in 1994 (Kunreuther and Roth, Sr. 1998).

Once the risk is estimated, the insurer needs to determine a premium rate that yields a profit and avoids an unacceptable level of loss. There are a number of factors that influence an insurer's decision on what premium to set. State regulations often limit insurers in their rate-setting process, and competition can play a role in what may be charged in a given marketplace. Even in the absence of these influences, there are a number of issues that an insurer must consider in setting premiums: uncertainty of losses, highly correlated losses, adverse selection, and moral hazard. Neither adverse selection nor moral hazard appears to be a major problem with respect to natural hazard risks. Adverse selection occurs when the insurer cannot distinguish (or does not discriminate through price) between the expected losses for different categories of risk, while the insured, possessing information unknown to the insurer, selects a price/coverage option more favorable to the insured. Moral hazard refers to an increase in the expected loss caused by the behavior of the policyholder. One example of moral hazard is moving unwanted furniture into the basement so an impending flood can destroy it, but this behavior occurs very infrequently. Given the difficulty uncertainty of losses and highly correlated losses pose in setting premiums, they are discussed below.

2.5.2 Uncertainty of Losses

Natural disasters pose a set of challenging problems for insurers because they involve potentially high losses that are extremely uncertain. Figure 2.8 illustrates the total number of loss events from 1950 to 2000 in the United States for three prevalent hazards: earthquakes, floods, and hurricanes. Events were selected that had at least $1 billion of economic damage and/or over 50 deaths (American Re, 2002).

Looking across all the disasters of a particular type (earthquake, hurricane or flood), for this 50-year period, the median loss is low while the maximum loss is very high. Given this wide variation in loss distribution, it is not surprising that there is a need for catastrophe models to aid insurers and reinsurers in estimating the potential loss from events that have not yet occurred but are scientifically credible.

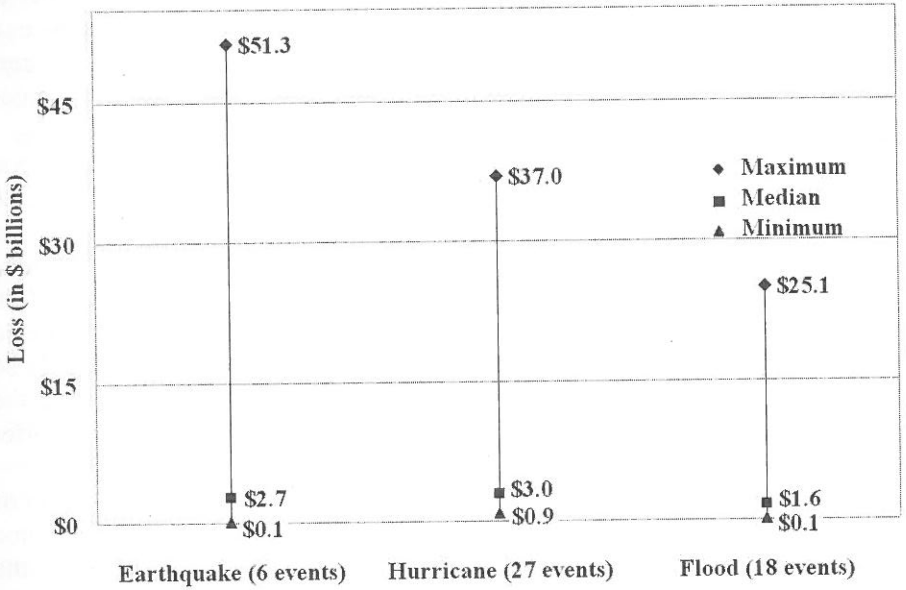

Figure 2.8. Historical economic losses in $ billions versus type of significant U.S. natural disaster. 1950-2000 (Source: American Re)

2.5.3 Highly Correlated Losses

Natural disasters involve spatially correlated losses or the simultaneous occurrence of many losses from a single event. If insurers sell a block of residential policies in a neighborhood, they could potentially experience a large (spatially correlated) loss should a disaster occur in the region. For example, due to their high concentration of homeowners' policies in the Miami/Dade County area of Florida, State Farm and Allstate Insurance paid $3.6 billion and $2.3 billion in claims respectively in the wake of Hurricane Andrew in 1992. Given this unexpectedly high loss, both companies began to reassess their strategies of providing coverage against wind damage in hurricane-prone areas (Lecomte and Gahagan, 1998).

In general, insurance markets flourish when companies can issue a large number of policies whose losses are spatially and otherwise independent. The portfolio follows the law of large numbers, and is thus predictable. This law states that for a series of independent and identically distributed random variables, the variance around the mean of the random variables decreases as the number of variables increases. Losses from natural hazards do not follow the law of large numbers, as they are not independent.

2.5.4 Determining Whether to Provide Coverage

In his study, James Stone (1973) sheds light on insurers' decision rules as to when they would market coverage for a specific risk. Stone indicates that firms are interested in maximizing expected profits subject to satisfying a constraint related to the survival of the firm. He also introduces a constraint regarding the stability of the insurer's operation. However, insurers have traditionally not focused on this constraint in dealing with catastrophic risks.

Following the disasters of 1989, insurers focused on the survival constraint in determining the amount of catastrophe coverage they wanted to provide. Moreover, insurers were caught off guard with respect to the magnitude of the losses from Hurricane Andrew in 1992 and the Northridge earthquake in 1994. In conjunction with the insolvencies that resulted from these disasters, the demand for coverage increased. Insurers only marketed coverage against wind damage in Florida because they were required to do so and state insurance pools were formed to limit their risk. Similarly, the California Earthquake Authority enabled the market to continue to offer earthquake coverage in California.

An insurer satisfies the survival constraint by choosing a portfolio of risks with an overall expected probability of insolvency less than some threshold, p_1. A simple example illustrates how an insurer would utilize the survival constraint to determine whether the earthquake risk is insurable. Assume that all homes in an earthquake-prone area are equally resistant to damage such that the insurance premium, z, is the same for each structure. Further assume that an insurer has \$A dollars in current surplus and wants to determine the number of policies it can write and still satisfy its survival constraint. Then, the maximum number of policies, n, satisfying the survival constraint is:

$$\text{Probability [Total Loss} > (n \cdot z + A)] < p_1$$

Whether the company will view the earthquake risk as insurable depends on whether the fixed cost of marketing and issuing policies is sufficiently low to make a positive expected profit. This, in turn, depends on how large the value of n is for any given premium, z. Note that the company also has some freedom to change its premium. A larger z will increase the values of n but will lower the demand for coverage. The insurer will decide not to offer earthquake coverage if it believes it cannot attract enough demand at any premium structure to make a positive expected profit. The company will use the survival constraint to determine the maximum number of policies it is willing to offer.

The EP curve is a useful tool for insurers to utilize in order to examine the conditions for meeting their survival constraint. Suppose that an insurer wants to determine whether its current portfolio of properties in Long Beach is meeting the survival constraint for the earthquake hazard. Based on its current surplus and total earthquake premiums, the insurer is declared insolvent if it suffers a loss greater than \$15 million. The insurer can construct an EP curve such as Figure 2.4 and examine the probability that losses exceed certain amounts. From this figure, the probability of insolvency is 1.0%. If the acceptable risk level, $p_1 < 1.0\%$, then the insurer can either decrease the amount of coverage, raise the premium and/or transfer some of the risk to others.

2.6 Framework to Integrate Risk Assessment with Risk Management

Figure 2.9 depicts a framework for integrating risk assessment with risk management and serves as a guide to the concepts and analyses presented in this book. The risk is first assessed through catastrophe modeling. Catastrophe modeling combines the four components (hazard, inventory, vulnerability, and loss) to aid insurers in making their decisions on what type of protection they can offer against a particular risk.

The key link between assessing risk via catastrophe models and implementing risk management strategies is the stakeholders' decision processes. The types of information stakeholders collect and the nature of their decision processes are essential in developing risk management strategies. With respect to insurers, catastrophe models are the primary sources of information on the risk. Their decision rule for developing risk management strategies is to maximize expected profits subject to meeting the survival constraint. Property owners in hazard prone areas utilize simplified decision rules in determining whether or not to adopt mitigation measures to reduce future losses to their property and/or to purchase insurance.

For purposes of this book, risk management strategies are broadly classified as either risk reduction measures, such as mitigation, or risk transfer measures, such as reinsurance. For example, strategies for residential property owners often involve a combination of measures, including mitigation, insurance, well-enforced building codes, and land-use regulations. In California and Florida, all these initiatives exist in some form. Strategies for insurers could involve charging higher rates to reflect the uncertainty of the risk, changing their portfolio so they can spread the risk across many areas, or reassigning the risk using risk transfer instruments such as reinsurance and/or catastrophe bonds.

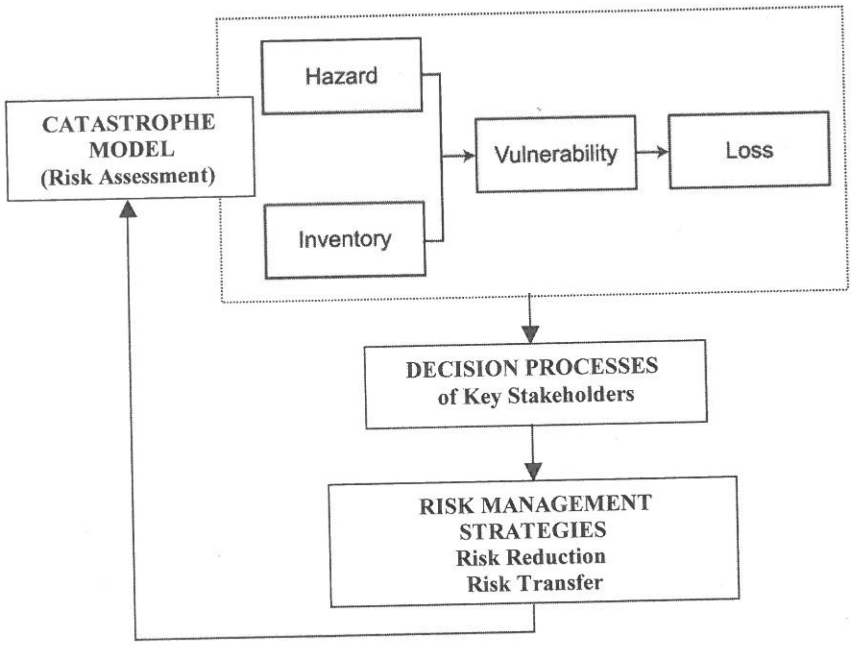

Figure 2.9. Framework for linking risk assessment with risk management.

2.7 Summary and Relationship to Parts II-IV

This chapter examined the history of catastrophe modeling and the role catastrophe models play in making a risk insurable. Part II provides a more detailed discussion of catastrophe modeling for earthquakes and hurricanes. The output from catastrophe models provides important information for insurers to manage their risk. By modeling the risk, insurers can more accurately estimate the premiums to charge for insurance coverage from natural disasters. In addition, insurers and reinsurers are able to tailor their coverage to reduce the chances of insolvency. They can develop new strategies for managing their portfolios so as to avoid losses that might otherwise cause an unacceptable reduction in surplus. These strategies are discussed in Part III of the book.

The impact of insurers' risk management strategies on profitability and probability of insolvency are explored further in Part IV of the book. Exceedance probability curves are constructed using real market data for insurers in Oakland, California, Long Beach, California and Miami/Dade County, Florida and alternative strategies are examined, including requiring mitigation to homes in these disaster-prone areas and using risk transfer instruments to satisfy an insurer's survival constraint. The book concludes with a chapter on the future role of catastrophe models in dealing with the risks associated with terrorism as an extreme event.

2.8 References

American Re (2002). *Topics: Annual Review of North American Natural Catastrophes 2001.*

Algermissen, S.T. (1969). *Seismic risk studies in the United States,* 4th World Conference on Earthquake Engineering Proceedings, Chilean Association for Seismology and Earthquake Engineering, Santiago, Chile.

Brinkmann, W. (1975). *Hurricane Hazard in the United States: A Research Assessment.* Monograph #NSF-RA-E-75-007, Program on Technology, Environment and Man, Institute of Behavioral Sciences, University of Colorado, Boulder, Colorado.

FEMA 249 (1994). *Assessment of the State of the Art Earthquake Loss Estimation Methodologies,* June.

FEMA (2002). *HAZUS99 SR2 User's Manual,* Federal Emergency Management Agency.

Insurance Information Institute (2000). Catastrophes [online]. The Insurance Information Institute 10 June 2000 <http://www.iii.org/media/issues/catastrophes>.

Kozlowski, R. T. and Mathewson, S. B. (1995). "Measuring and Managing Catastrophe Risk," *1995 Discussion Papers on Dynamic Financial Analysis,* Casualty Actuarial Society, Arlington, Virginia.

Kunreuther, H., R. Hogarth, J. Meszaros and M. Spranca (1995). "Ambiguity and underwriter decision processes," *Journal of Economic Behavior and Organization,* 26: 337-352.

Kunreuther, H. and Roth, R. (1998). *Paying the Price: The Status and Role of Insurance Against Natural Disasters in the United States.* Washington, D.C: Joseph Henry Press.

Neumann, C.J. (1972). *An alternate to the HURRAN tropical cyclone forecast system.* NOAA Tech. Memo. NWS SR-62, 24 pp.

NIBS (1997). *HAZUS: Hazards U.S.: Earthquake Loss Estimation Methodology.* NIBS Document Number 5200, National Institute of Building Sciences, Washington, D.C.

Panjer, H. H. and Willmot, G.E. (1992). *Insurance Risk Models.* Illinois: Society of Actuaries.

Steinbrugge, K. V. (1982). *Earthquakes, Volcanoes, and Tsunamis: An Anatomy of Hazards.* Skandia America Group: New York, New York.

Kennedy, R.P. and Ravindra, M.K. (1984). *Seismic Fragilities for Nuclear Power Plant Risk Studies*, Nuclear Engineering and Design, 79: 47-68.

Kurihara, Y., Tuleya, R.E., Bender, M.A., and Ross R (1992). *Advanced Modeling of Tropical Cyclones*, Proceedings of the ICSU/WMO International Symposium on Tropical Cyclone Disasters, pp.190-201.

Mendez, A. and Anderson, J. G. (1991). *The temporal and spatial evolution of the 19 September 1985 Michoacan earthquake as inferred from near-source ground motion records*, Bulletin of the Seismological Society of America 81: 844-861.

Reiter, L. (1990). *Earthquake Hazard Analysis*, New York, Columbia University Press, New York.

SCEC (Southern California Earthquake Council) (1995). Working Group on California Earthquake Probabilities, *Seismic hazards in southern California: probable earthquakes, 1994-2024*, Bull. Seis. Soc. Am., 85: 379-439.

Schwerdt, R.W., Ho, F.P., Watkins, R.R. (1979). *Meteorological Criteria for Standard Projects Hurricane and Probable Maximum Hurricane Windfields, Gulf and East Coast of the united States*, United States Department of Commerce, National Oceanic and Atmospheric Administration, NOAA Technical report NWS23, September.

Simiu, E. and Scanlan, R.H. (1996). *Wind Effects on Structures- Fundamentals and Applications to Design*, Wiley Interscience.

USGS (1996). *National Seismic Hazard Maps Documentation: USGS Open-File Report 96-532.*

USGS (1999). *Earthquake Probabilities in the San Francisco Bay Region: 2000 to 2030 - A Summary of Findings, USGS Open-File Report 99-51.*

Wells, D.L., and Coppersmith, K.J. (1994). *New Empirical Relationships Among Magnitude, Rupture Length, Rupture Width, Rupture Area and Surface Displacement*, Bulletin of the Seismological Society of America, 84: 974-1002.

Chapter 4 – Sources, Nature, and Impact of Uncertainties on Catastrophe Modeling

Major Contributors:
Patricia Grossi
Don Windeler

4.1 Introduction

Catastrophe modeling is a complex tool used to assess the risk from natural hazards. The four components of hazard, inventory, vulnerability, and loss depicted in Figure 3.1 and discussed in detail in Chapter 3 require information from a range of sources and the expertise of an array of professionals. Natural hazard, engineering and economic data are the foundation of catastrophe models. Limitations in data and assumptions about the model's parameters, in the hazard, inventory, and vulnerability modules, affect a catastrophe model's loss estimates and the uncertainty associated with these estimates.

This chapter explores the sources, nature, and impact of uncertainties in a catastrophe model. Prevalent methods to represent and quantify uncertainty through the components of the catastrophe model are discussed. Finally, the impact of uncertainty on exceedance probability (EP) curves used by risk managers to quantify their catastrophe risk potential is illustrated by examining potential losses to residential property from hurricanes in Florida and earthquakes in Charleston, South Carolina. Quantification and classification of uncertainty provides opportunities to reduce risk. With accurate measures of uncertainty, stakeholders can potentially lower the cost of dealing with catastrophe risk. Furthermore, since the risk affects stakeholders in dissimilar ways, the robustness of a risk management strategy can be made clear to each stakeholder if uncertainty is delineated.

4.2 Classifications of Uncertainty

As indicated in Chapter 3, there is a great deal of information needed to develop the hazard, inventory, vulnerability, and loss components of a catastrophe model. Therefore, all stakeholders in the management of risk value new information regarding these modules. For example, an insurer values additional information on the likelihood of disasters and potential damage to properties in its portfolio in order to more accurately manage the risk. Local government officials value a thorough understanding of hazards in their regions in order to plan for emergency response and recovery efforts following a disaster. Model developers value any additional information to validate and calibrate their catastrophe models.

Since catastrophe modeling is a fairly new field of application, there are no historical classifications of catastrophe modeling uncertainty, per se. However, building on the concepts from probabilistic hazard analyses, uncertainty can be characterized as either aleatory or epistemic in nature (Budnitz et al., 1997). Aleatory uncertainty is the inherent randomness associated with natural hazard events, such as earthquakes, hurricanes, and floods. It cannot be reduced by the collection of additional data. In contrast, epistemic uncertainty is the uncertainty due to lack of information or knowledge of the hazard. Unlike aleatory uncertainty, epistemic uncertainty can be reduced by the collection of additional data.

While the advantage of differentiating between aleatory and epistemic uncertainty in an analysis is clear (only epistemic uncertainty can be reduced), the necessity of distinguishing between aleatory and epistemic uncertainty is not. “Epistemic and aleatory uncertainties are fixed neither in space…nor in time. What is aleatory uncertainty in one model can be epistemic uncertainty in another model, at least in part. And what appears to be aleatory uncertainty at the present time may be cast, at least in part, into epistemic uncertainty at a later date” (Hanks and Cornell, 1994). Therefore, developers of catastrophe models do not necessarily distinguish between these two types of uncertainty; instead, model developers concentrate on not ignoring or double counting uncertainties and clearly documenting the process in which they represent and quantify uncertainties.

4.3 Sources of Uncertainty

Limited scientific knowledge, coupled with a lack of historical data, leave open several possible and competing explanations for the parameters, data, and mathematical models underlying each of the components in a catastrophe model. Simply put, the science and impact of natural hazards are not completely understood; in addition, the cross-disciplinary nature of a catastrophe model leads to complexity. Experts in seismology or meteorology who model the hazard must interact with structural engineers

who model the vulnerability; similarly structural engineers who model the vulnerability must interact with actuaries who model the loss. Basically, as each discipline’s modeling assumptions are added to the process, more uncertainty is added to the estimates.

In catastrophe modeling, both epistemic and aleatory uncertainties are reflected in the four basic components of a model. Aleatory uncertainty is reflected via probability distributions. The frequency of a hazard occurrence and the fragility of a building, as discussed in Chapter 3, are examples of aleatory uncertainty. Since the exact time of occurrence and the precise level of structural damage cannot be known in advance of a hazard event, the recurrence rate and the vulnerability of the inventory exposed to the natural hazard are characterized using probability distributions. Similarly the capacity of individual structural elements of a building during a severe event, and the resulting cost of repair cannot be determined beforehand. Probability distributions are also used to characterize these parameters in a catastrophe model.

A larger issue in quantifying uncertainty is the lack of data for characterizing the four components in a catastrophe model. For example, as discussed in Chapter 3, the recurrence of earthquake events on fault sources can be modeled using a magnitude-frequency model (Richter, 1958), a characteristic earthquake model (Youngs and Coppersmith, 1985), or a combination of both models. In California, estimates of ground shaking probabilities on certain fault segments are established by combining the two recurrence models for earthquake magnitude-frequency distributions (Peterson et al. 1996). Historical earthquake records are used to establish a recurrence curve, or the Gutenberg-Richter relationship, for the smaller magnitude events, while geologic data (most importantly, a fault’s slip rate) is used to estimate the recurrence of the larger, characteristic events.

The availability of seismological data describing earthquake occurrence in California for only a few hundred years makes the updating of the recurrence distributions problematic. When more data become available, in the form of fault slip rates or seismograph recordings, these relationships could potentially be improved. Similar issues arise in modeling the recurrence of hurricane events. Past data describing the location and occurrence of hurricanes on the eastern seaboard of the United States are also limited to a few hundred years (Powell and Aberson, 2001).

The deficiency of information regarding repair costs and business interruption costs affect the accuracy of the loss component of a catastrophe model. For example, the increased cost to repair or rebuild after an event is often taken into account using a demand surge adjustment. This is simply the percentage increase in costs due to the limited supply of construction material and labor immediately following a disaster. Further, due to the growing understanding of indirect losses, estimates of business interruption costs to commercial property owners are continually validated and calibrated with the latest loss information.

Another source of epistemic uncertainty in a catastrophe model is the lack of available data to create the Geographic Information Systems (GIS) databases within the modeling software. For any model, recognizing the importance of input data is essential. The “garbage in, garbage out” principle holds irrespective of how advanced or state-of-the-art a model may be. GIS maps of hazard sources, geologic features and topographic landscape characterize hazards. GIS maps of the locations of structures characterize inventory.

An incomplete description of a hazard source, the geology or the topography can cause erroneous results. For example, in earthquake modeling, having accurate information on the underlying soil in a region is very important. A structure built on rock-like material is likely to sustain much lower losses compared to a structure built on soft clay-like material. Inaccurate information on soil conditions can lead to large errors in estimation of loss due to an earthquake.

In fact, past observations from earthquakes confirm that soil condition plays a very important role in building performance. As expected, buildings on soft ground or steep slopes usually suffer more significant damage in comparison to those on firm and flat ground. Since soil condition may vary dramatically within a small area, such as the Marina District in San Francisco (where soil conditions vary from bay mud to rock site), using ZIP code to identify a location may not be sufficiently accurate. At a particular location, high-resolution geocoding should be used as it can more accurately pin down the soil condition.

Partial information on a structure’s characteristics can also result in an inaccurate estimate of future damage. For example, most structural engineers would agree that the construction type, age, height, occupancy, assessed value, and the location of a structure are needed – at a minimum – for the inventory component of a catastrophe model. If more specific information regarding the structure such as its location relative to other structures and previous damage to the structure were available, a more accurate estimate of damage or vulnerability would result.

Lack of accurate data on true market values of the properties under consideration is an additional source of epistemic uncertainty in the modeling process. For determining the appropriate coverage limit, many residential policies use property tax assessment data, which are generally outdated and under-valued. Under-valued exposures will result in under-estimating potential loss. For example, suppose a home’s property value is assessed at $600,000 when its true worth is $1 million. Furthermore, suppose it is insured with a 15% deductible and full coverage based on the lower assessed value. If an earthquake occurs and causes major damage and the cost to repair the structure is 35% of the true value of the home, the resulting monetary loss is $350,000. A $600,000 insurance policy with a 15% deductible translates to the homeowner being responsible for $90,000, with the insurer covering the remaining $260,000 of the loss. If the insurance coverage had been based on the home’s true worth of $1 million, the homeowner would have to cover the first $150,000 of the loss and the insurer would only have claim payments of $200,000.

Incomplete or inaccurate information on an inventory's description is a concern not only to insurers but also to all risk management stakeholders. To improve on the amount of such information available, an effort to document the types of housing structures worldwide was initiated in 2000 to assess the vulnerability of the world’s population to earthquake hazard. Under the guidance of the Earthquake Engineering Research Institute (EERI) and the International Association of Earthquake Engineering (IAEE), the World Housing Encyclopedia has a web-based listing of housing construction types from earthquake-prone countries around the world (EERI, 2003). In addition, the Institute for Business and Home Safety (IBHS) relies on INCAST, a data inventory tool used in conjunction with the HAZUS catastrophe model, to store inventory information on the homes that are a part of their "Fortified...for safer living" program. These homes are reinforced to withstand many natural hazards, including high winds, wildfire, flood, hail, and earthquake.

Epistemic uncertainty is also found in the use of laboratory testing (shake table tests for earthquake hazard or wind-tunnel tests for hurricane hazard) and expert opinion to develop the vulnerability component of a catastrophe model. For a portfolio risk assessment, damage functions such as the one illustrated in Figure 3.7 in Chapter 3, have traditionally been constructed using these sources along with damage surveys of actual structures. Given that laboratory testing has been restricted to certain types of structural materials, there is a limited understanding of how other materials withstand lateral loading.

In the earliest versions of catastrophe models, damage ratios were estimated using the Applied Technology Council report of Earthquake Damage Evaluation Data for California (ATC-13, 1985). This report was generated using the Delphi method of collecting information from a group of experts (Dalkey, 1969). In this method, a series of questionnaires interspersed with controlled opinion feedback resulted in a group judgment. In the ATC-13 study, 71 earthquake engineering experts were asked to indicate their low, best, and high estimates of damage ratios for 78 types of structures subject to earthquakes with Modified Mercalli Intensity (MMI) levels of VI through XII. Catastrophe model developers used these estimates in their earliest versions of their earthquake loss software, skewing estimates of damage due to the use of the Delphi Method and limiting the interpretation of damage due