0364

UNCERTAINTY MODELING FOR DISASTER LOSS ESTIMATION

Felix S WONG[1], Hanyao CHEN[2] And Weimin DONG[3]

SUMMARY

This paper delineates the many kinds of uncertainties that must be properly accounted for in estimating losses caused by natural disasters, and the effects they have on risk management. Uncertainties associated with each step of the loss estimation chain, in the context of prevalent engineering models, are discussed. It is shown that the main sources of uncertainties such as occurrence, attenuation and vulnerability uncertainties may be classified into two general groups: aleatory and epistemic. The latter includes model and parameter uncertainties, as well as incompleteness of information such as site condition, structural details or policy terms. In particular, aleatory uncertainties lead to losses that are distributed, as compared with point estimates; distributed-ness affects loss allocations to the insured, insurer and reinsurer. The average annual loss and exceedance probabilities are also affected. Epistemic uncertainties widen the loss distribution further, but, most importantly, they cause losses at different locations (and losses of different portfolios) to be correlated. Such loss correlation has strong ramifications on the aggregate portfolio risk, and must be included in the modeling.

INTRODUCTION

As is well known, disaster loss estimation for the purpose of financial planning consists of the following parts:

- A source model, which predicts a series of events with various sizes, locations and frequencies.
- A site hazard model, which estimates the local severity of hazards including attenuation from the source.
- A vulnerability model, which estimates the asset loss given the local hazards.
- A financial model, which calculates the portion of the loss allocated to specific financial perspectives, e.g., the insured, insurer and reinsurer, based on the policy/portfolio structure and the uncertainty in the asset loss.

Uncertainties exist in each of the four modules in the assessment chain, and their cumulative effects are manifested in very important financial impacts. In particular,
For location loss, they cause a shift in the mean loss of specific financial perspectives.

- For policy/portfolio (multi-location) loss, they cause correlation between losses at different locations and, hence, a shift in the mean loss of specific financial perspectives.

They cause uncertainties in the average annual loss and exceedance probabilities, two key measures of risk in management decision.

The uncertainty models and their effects will be delineated in the following. To facilitate the discussion, we shall limit source uncertainties to uncertainties on the rate of occurrence. In other words, all events (their numbers and magnitudes) pertinent to the assets in question are assumed known; only how often they occur is uncertain, and uncertain occurrences lead to uncertainty in loss. We refer to these uncertainties as **OR uncertainties**. On the other hand, given an event, uncertainties in site hazards, vulnerability, etc., will also lead to uncertainty in loss.

[1] *Weidlinger Associates, Inc., 4410 El Camino Real, Suite 110, Los Altos, CA 94022-1049 USA. felix@ca.wai.com*
[2] *Risk Management Solutions, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025 USA hanc@riskinc.com*
[3] *Risk Management Solutions, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025 USA. weimind@riskinc.com*

We refer to these uncertainties collectively as **SV uncertainties**. Due to space limitation, the discussion favors not details, which are left to the references, but the impact these uncertainties have on loss estimation and portfolio risk quantification. An appreciation of the latter is especially important since it underscores why uncertainty modeling is essential in disaster loss estimation. Uncertainties are, quite simply, an indispensable part of any risk model.

It is also important to recall that uncertainties can be aleatory or epistemic in nature. **Aleatory uncertainty** is intrinsic, the irreducible uncertainty of the probabilistic phenomenon itself. **Epistemic uncertainty** is attributed to a lack of knowledge, and is reducible only at the expense of a very large sample size and accumulated knowledge. Epistemic uncertainty can be related to validity and accuracy of a model (hence, called **modeling uncertainty**), or the statistical basis for the parameters of the model (called **parameter uncertainty**). While both the aleatory and epistemic types contribute to the ultimate uncertainty in the loss, epistemic uncertainties are largely responsible for the correlation between losses at different locations or from different portfolios. Such correlation plays an essential role in financial risk management. Insurance companies that issue catastrophe policies such as for earthquakes and hurricanes are concerned with their probable maximum loss (PML). Their portfolios consist of many individual policies, and the maximum loss is an aggregation of the losses from the individual policies. Reinsurance companies that issue catastrophe treaties to primary insurers are also concerned with their portfolio risk. In this case, the risk is an aggregation of the risks of the primary insurers' portfolios.

This paper discusses how portfolio losses can be determined from their component (policy or cedant insurer) losses. For a portfolio with assets at many different locations, the mean of the aggregated loss is simply the sum of the mean location losses. However, the computation of the standard deviation of the aggregated loss is more complicated because losses at any two locations may be correlated. Such correlation is known to have significant impact on the distribution of the aggregated loss. Less well known but equally important is the fact that the allocation of losses to the reinsurer, such as under an excess-loss treaty, depends not only on the mean of the aggregated loss but even more so on the loss distribution. Hence, location loss correlation plays an important part in quantifying portfolio risk.

UNCERTAINTY MODELS

Or Uncertainties

State-of-the-art research indicates that the time-independent Poisson model is adequate for modeling the occurrence of earthquakes and hurricanes, although in parts of the country such as California where a larger database may support a time-dependent model (e.g., the USGS model for earthquakes [Cornell and Winterstein, 1986], [USGS, 1988], [USGS, 1990]). However, the average rate, λ, a parameter of the Poisson model, is subject to uncertainty due to (for earthquakes) relative short recording periods compared with the recurrence interval of the events of interest, and (for both earthquakes and hurricanes) due to limited knowledge. Many studies have been performed in which geological and geophysical data are used to reduce the uncertainty in the occurrence rates of earthquakes [Dong, et. al., 1987], [Wesnousky, 1986], [Mark, 1996], [Cramer et. al., 1996], [Dong, 1997]. For hurricanes, the corresponding physical quantities are not available, and occurrence data are used to estimate the statistical error in rate estimation [Gorden et. al., 1997].

Hence, physical quantities underlying the Poisson model contain uncertainty, leading to epistemic uncertainty in the estimate of rates in addition to the aleatory uncertainty inherent in the model. We assume that the occurrence rate, λ, follows the normal or lognormal distribution so that its uncertainty is defined by two parameters such as the mean m_λ and standard deviation σ_λ. m_λ and σ_λ have been related to uncertainties in creep and slip rate (e.g., see [USGS, 1988], [USGS, 1990], [USGS, 1995]).

Sv Uncertainties

Ground Motion Attenuation - Current methods for estimating ground motion belong to two groups: methods based on wave propagation and empirical methods, and the latter is more widely used. One popular empirical model is attributed to BJF [Boore, Joyner and Fumal, 1993]:

$$\log A = -0.95 + 0.23 M_0 - \log R_1 - 0.0027 R_1 + \varepsilon \qquad (1)$$

where A = peak ground acceleration in g, M_0 = the moment magnitude, $R_1 = \sqrt{d^2 + 8^2}$, d = closest distance to the surface projection of the fault rupture in km, and ε is a zero-mean normal random variable that represents the aleatory uncertainty. In particular, through comparison with available ground motion data from recent earthquakes [Joyner and Boore, 1988], ε is shown to have a standard deviation of 0.28 (i.e., $\sigma_{\log A} = 0.28$). Note that while the median values of peak ground acceleration of two neighboring locations are highly correlated through the function *logA* in Eq.1, their random variation portions denoted by ε in the equation are *independent*.

Because there is no consensus on which of the many attenuation relationships proposed is the best, epistemic ground-motion uncertainty can be modeled as different relationships between peak acceleration, earthquake magnitude and distance from fault (e.g., in a logic-tree approach such as that used in [Cramer et. al., 1996]). Hence, the resultant modeling uncertainty can be as significant as the respective aleatory uncertainties such as the ε for Joyner and Boore shown in Eq.1, and unlike the latter, is correlated for all locations. That is, if a model chosen gives “high” ground motions, then the ground motion for all locations based on that model will be “high”.

Vulnerability Uncertainties - While structures of the same type tend to perform similarly, they exhibit a deal of variability due to differences in details. Based on analysis of expert opinions, ATC-13 [ATC, 1985] recommends that the *mean* damage ratio be related to ground motion intensity and building class as:

$$m_R = g(MMI, class) \tag{2}$$

with the damage ratio *R* for a particular building within the class be given by:

$$f_R(r) = \frac{1}{B(\lambda,\nu)} \frac{r^{\lambda-1}(100-r)^{\nu-1}}{100^{\lambda+\nu-1}} \tag{3}$$

where *r* is the damage ratio in percent, and λ and ν are parameters of the Beta function *B* and can be related to the mean damage ratio given in Eq.2 and the coefficient of variation of the damage ratio. Hence, the model reflects that damage is random and, for the same intensity, independent of location. In particular, buildings of the same type at the same nominal location can have very different damages due to aleatory uncertainty.
Since the mean damage ratios solicited in the ATC-13 survey exhibit quite a wide range, and the mean damage affects λ and ν, parameters of the Beta function, variation in the mean damage ratio is parameter uncertainty – a variation that is totally correlated and constitutes a systematic error due to lack of knowledge of the real mean damage ratio. For example, if the real mean is 10%, and 8% is used in determining the Beta distribution, then all damage ratios randomly generated based on the Beta distribution with the lower mean will be underestimated.

Incomplete Information - It often happens in loss estimation that many pieces of information are not available. A common example is that only the aggregate exposure of *n* buildings at a particular zip code is known, but the individual buildings are not. One may have to assume that the buildings are uniformly distributed in all building classes, and the average damage ratio based on this assumption may be larger than the actual. The same situation exists in dealing with soil data with various degrees of accuracy/precision (regional, zip code, or census tract data).

Uncertainty associated with incomplete data is epistemic, and can only be reduced with more data collection effort. As things stand, various assumptions can be made regarding the distribution of such uncertainty and used in loss estimation, but the best approach is evaluation by experience and judgment on a case-by-case basis.

EFFECTS ON LOSS

Uncertainties in occurrence, attenuation, vulnerability, and incomplete information are combined to yield the uncertainty in building loss estimate, as depicted in Fig.1. To improve clarity in discussion, the effects due to OR uncertainties is further separated from that due to SV uncertainties. The latter is also called loss uncertainty given an event, which we shall address first.

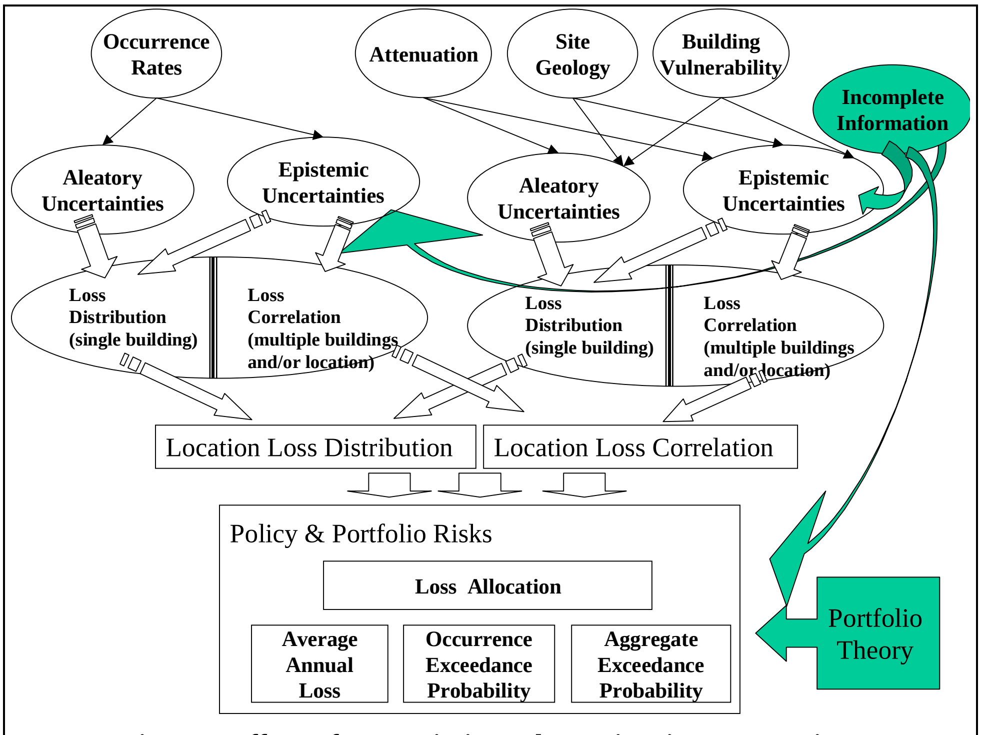

Figure 1. Effects of uncertainties on loss estimation, an overview.

Given an event, uncertainties in attenuation and vulnerability as well as incomplete information can be channeled into the two groups marked "aleatory" and "epistemic". Aleatory uncertainties contribute to the probabilistic variation in building loss at a location; as a result, the building loss is now distributed, as compared with deterministic. We refer to this effect as **distributed loss** for emphasis. The effects of epistemic uncertainties are more complex and can best be delineated with reference to a single building (location), or multiple buildings (locations). With respect to the loss of a single building, epistemic uncertainties contribute further to the distributed-ness in the loss (e.g., incomplete information on the soil type admits all possible soil types, and, hence, begets a wider range of hazards). They widen the loss distribution. With respect to buildings at multiple locations, epistemic uncertainties contribute to **correlation** of the losses of the buildings due to an event (i.e., when the loss at one location is higher than the average, the loss at a different but correlated location is also high). Major factors contributing to loss correlation are: Geographic (buildings in the same soil pocket), vulnerability (buildings of the same class), geologic (buildings on the same soil) and attenuation (buildings at same distance to fault). The epistemic effect relationships are singled out in Fig.1 for emphasis.

To appreciate the effects due to OR uncertainties, it is expedient to assume for the moment that SV uncertainties are zero. Building loss given that a specific event has occurred is then deterministic. Call it L_i where i denotes the event. Building loss corresponding to an event whose occurrence is governed by the Poisson process is probabilistic, and the average annual loss, denoted by AAL is $\lambda_i L_i$ where λ_i is the rate of occurrence. This accounts for the aleatory uncertainty in occurrence. When the rates themselves are uncertain, i.e., epistemic uncertainty in the model parameter, the probabilities of one event, two events, …., occurring in a year will vary, and, hence, the resultant loss distribution is affected. In particular, the AAL is random with mean and variance given by:

$$\overline{AAL} = \bar{\lambda}_i L_i \quad \text{and} \quad Var(AAL) = \sigma_{\lambda_i}^2 L_i^2 \tag{4}$$

where $\bar{\lambda}_i$ and σ_{λ_i} are the mean and standard deviation of the rate of occurrence, respectively. If we now incorporate the SV uncertainties into the process, Eq.4 becomes:

$$\overline{AAL} = \bar{\lambda}_i \bar{L}_i \quad \text{and} \quad Var(AAL) = \bar{\lambda}_i^2 \sigma_{L_i}^2 + \bar{L}_i^2 \sigma_{\lambda_i}^2 + \sigma_{\lambda_i}^2 \sigma_{L_i}^2 \tag{4a}$$

where $\bar{L}_i$ and σ_{L_i} are the mean and standard deviation of building loss due to SV uncertainties.

Allocation Method	Client Loss	Insurer Loss
Point-Estimate	7%	0%
Distributed	5.2%	1.8%

In a similar but more complex fashion, the occurrence exceedance probability (OEP) and the aggregate exceedance probability (AEP) for a threshold value can be shown to evolve from "deterministic" to probabilistic due to OR uncertainties. Details are too cumbersome to include herein, but

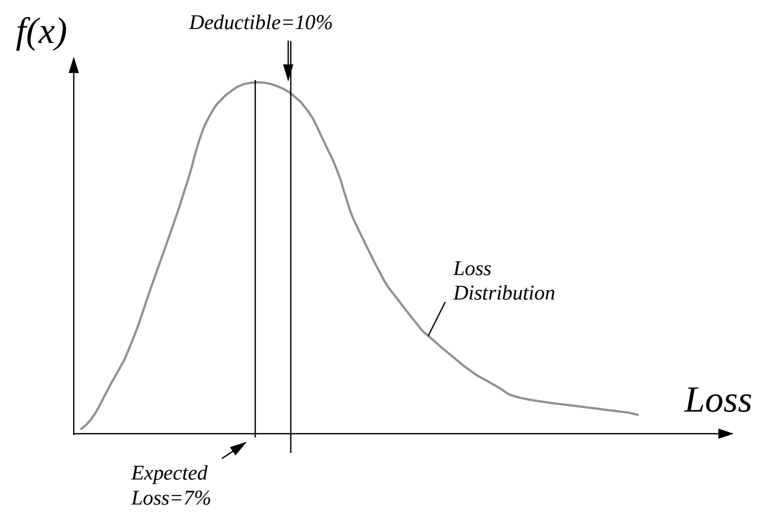

Figure 2. Simple example on the difference between expected and distributed loss allocations.

suffice it to say that higher layers tend to have larger rate variability; large losses are caused by severe events which by nature are infrequent, and there is scarce data for those events.

Hence, in summary, aleatory and epistemic components of uncertainty contribute to the distributed-ness in the loss (pertaining to a single building) and epistemic components contribute to loss correlation (pertaining to multiple buildings). Loss correlation increases the spread in the aggregate loss for multiple buildings or portfolios. These effects, in turn, influence decision parameters such as the average annual loss, exceedance probabilities and allocated loss. We shall illustrate these main points in the following.

IMPORTANCE TO INSURANCE RISK MANAGEMENT

Loss Allocation To The Insured And Insurer (Effect Of Distributed Loss)

Suppose the estimated damage of a building has a probability density distribution as depicted in Fig.2, with an expected (mean) damage of 7%. Suppose the coverage deductible is 10%. We compare the loss allocations made based on a point-estimate of damage, say, the expected damage, and based on the distributed damage in the table below.
In the former case, because the point-estimate damage (at 7%) is less than the deductible, the client incurs all of the loss. In the latter, the damage is sampled from the distribution curve, and the loss allocated base on the sample damage and the deductible. Hence, for sample damages that are below the deductible, 100% of the loss goes to the client. For sample damages greater than the deductible, the client accepts the deductible, and the excess goes to the insurer. The sample allocations are then weighted by the appropriate probability density. For this example, the net mean loss allocations are 5.2% for client and 1.8% for insurer.

Hence, mainly because of the presence of deductibles, the loss allocation can be warped if the point-estimate damage only is used. Deductibles are said to influence the allocation at the lower end, as they are designed to do. Similarly, the presence of limits affects the allocation at the high end, and other policy structures and portfolio applications have interactions with the distributed loss. While the example shows the mean values of the allocated losses are affected, the standard deviations of the allocated losses are also affected.

Loss Allocation To The Insured, Insurer And Reinsurer (Effect Of Loss Correlation)

A policy/portfolio with assets at many different locations will suffer loss when any of the locations suffer a loss. It is well known that the mean loss for a portfolio is the sum of the mean location losses. But the standard deviation of the portfolio loss depends not only on the standard deviation of the location losses, but the correlation that may exist between losses at pairs of locations. For example, consider a portfolio with 100 locations and the policy coverage is $100,000 per location, for a total coverage of $10 million (100 × $100,000). Suppose the policy has a 5% deductible so that the amount of deductible is 5% of $10 million or $0.5 million. Suppose further that the mean damage to all locations is the same, at 10% damage. Hence, the expected loss to the portfolio is $1 million (10% of $10 million is $1 million). We shall compare the portfolio losses for two cases: when the location losses are totally independent, and when they are totally correlated.

Loss results for the two cases are generated by Monte Carlo simulations with a sample size of 10,000 and a coefficient of variation of 1.6 for location loss. They are presented in Fig.3 and compared with results for the no-uncertainty case. With reference to the figure, we see that when location losses are totally independent, the aggregate loss distribution is concentrated (i.e., a coefficient of variation of approximately 0.16). The distribution is centered around the deterministic value of $1 million. When the location losses are totally correlated, the aggregate distribution is spread out over a wide range; the coefficient of variation is about 10 times the independent case.

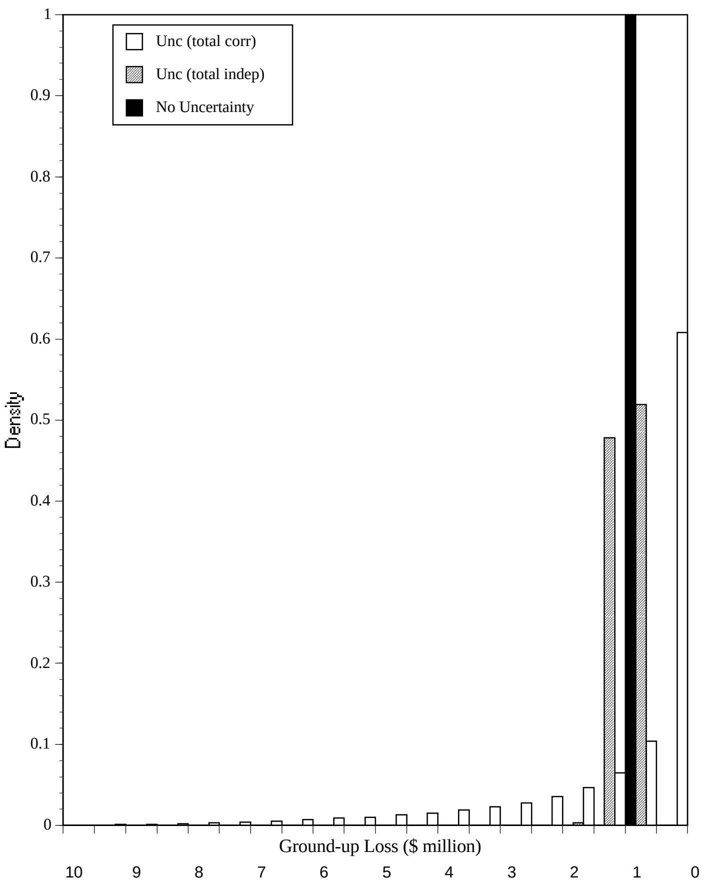

Figure 3. Density distribution of total asset loss with and without consideration of uncertainties.

These differences in distribution have significant impact on the allocation of insurer/reinsurer losses. We use the treaty structure shown in Fig.4b to illustrate this important point. Applying this structure to the loss results when there are no uncertainties (Fig.4a), the loss of $1 million is above the deductible of $0.5 million, and, hence, affects FAC1 but not FAC2. The allocation to FAC1 is (1-0.5)*40%=0.2. When the effects of uncertainties are included but their effects are independent (Fig.4c), the aggregate loss is distributed but the distribution is narrow banded. For this example, the distribution does not spread beyond $2 millions so that FAC1 is still the only treaty affected, as in the no-uncertainty case (although the losses allocated will be different, as is obvious from the figure). When uncertainties are included and their effects are totally correlated (Fig.4d), the aggregate loss is widely spread and both treaties are activated; all have allocated losses and the numerics is straightforward.

Average Annual Loss And Exceedance Probabilities (Effect Of Epistemic Uncertainties)

OR uncertainties, even when acting alone, affect the confidence in the estimate of the loss exceeding probabilities and the average annual loss, two measures that are of paramount importance to catastrophe risk managers. For example, the 95-percentile value of the exceeding probability for a particular loss threshold can be 1.15-1.75 times the best estimate, depending on the regional diversification and threshold position. The higher the threshold and the more concentrated the portfolio area, the wider will be the 95-percentile bounds. The effect of occurrence uncertainty on the average annual loss is to enlarge its confidence interval. The coefficient of variation with rate uncertainty is about twice that without. Depending on the geographical diversification, the 90-percent confidence interval can be 100± 26% of the mean for large areas, and 100±45% of the mean for concentrated areas.

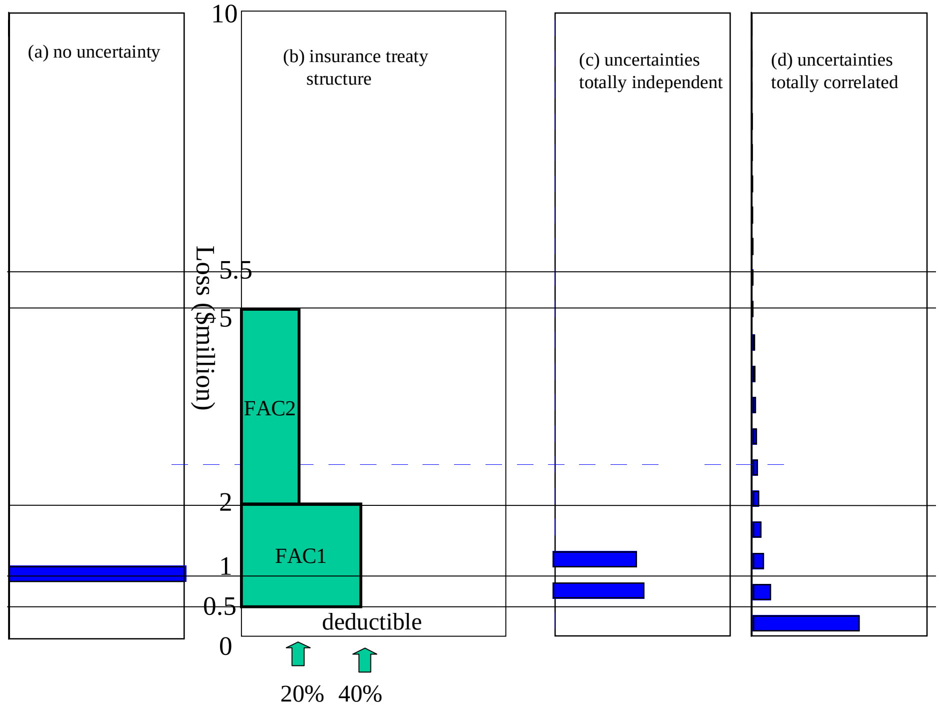

Figure 4. Insurance structure used in example.

SUMMARY

Inherent in the difficulty of managing catastrophe risks is the presence of uncertainties. There are uncertainties associated with the occurrence of an event, and, for an event, the loss induced is uncertain due to uncertainties in the hazards and vulnerability. The phenomenology models themselves, even with the state-of-the-art engineering sciences, are inherent uncertain due to scarce database and limited knowledge. These uncertainties, some epistemic and some aleatory, were discussed in the paper to emphasize that they have important bearing on key financial decision parameters. Hence, they must be properly quantified and their effects included in loss estimation, difficult though the task may be.

While existence of these uncertainties is commonly acknowledged, how they interact with the “bottom line” does not appear to have been well publicized. Part of the reason is that the interaction is complex and difficult to summarize without doing injustice. Generalization is even more dangerous. Nevertheless, it may be said that the presence of uncertainties has two major consequences in loss estimation. First, they contribute to uncertainties, i.e., probabilistic variation in the loss estimate – which we emphasized by calling the result distributed loss. Lesser known is the fact that distributed loss interacts with the structures of a policy/portfolio, and the allocation of the mean losses to the insured, insurer and reinsurer can be greatly affected. Simply put, treating the loss estimate as deterministic, i.e., ignoring uncertainties, will have grave ramifications.

The second major consequence of uncertainties is that the distributed losses at different locations of interest (such as those covered by a policy/portfolio) may also be correlated – which we called loss correlation. Loss correlation is attributed to epistemic uncertainties, and its essential role in portfolio risk management, though well-acknowledged in general concept, is seldom delineated in detail because of the complexities involved. This paper indicated how it evolved from the basic (epistemic) uncertainties in the loss estimation chain, and how it could be quantified. More on its impact on financial decisions is given in a companion paper in the same proceedings [Dong and Wong, 2000].

REFERENCES

1. Applied Technology Council, ATC-13, *Earthquake Damage Evaluation for California*, 1985.
2. Cramer, C. H., Petersen, M. D. and Reichle, M. S., “A Monte Carlo Approach in Estimating Uncertainty for a Seismic Hazard Assessment of Los Angeles, Ventura, and Orange Counties, California”, *Bulletin of Seismological Society of America*, Vol. 86, Number 6, December 1996, p.1681,

3. Cornell, C. A. and Winterstein, S. R., “Applicability of the Poisson Earthquake Occurrence Model”, Technical Report NP-4770, Electric Power Research Institute, Palo Alto, CA, 1986.
4. Dong, W., et al., “Utilization of Geophysical Information in Bayesian Seismic Hazard Model”, *Journal of Soil Dynamics and Earthquake Engineering*, Vol. 3, No. 2, pp. 103-11, 1987.

5. Dong, W. M., “Quantification of Parameter Uncertainty and Its Impact in USA Earthquake Model”, *RMS White Paper*, 1997.
6. Dong, W., and Wong, F., “Portfolio Theory for Earthquake Insurance Risk Assessment”, Proceedings of the 12WCEE, New Zealand, January-February 2000.
7. Gorden, R. et. al., “Hurricane Primary Parameter Uncertainty Estimation”, RMS White Paper, 1997.
8. Joyner, W. B., and Boore, D. M., “Measurement, Characterization and Prediction of Strong Ground Motion: Proceedings of Earthquake Engineering and Soil Dynamics, Park City, June 27-30, 1988”, American Society of Civil Engineers, 1988, No. 43, pg. 102.
9. Mark, D. P., et al., “Probabilistic Seismic Hazard Assessment for the State of California”, Department of Conservation, Division of Mines and Geology DMG Open-File Report 96-08, USGS Open-File Report 96-706, 1996.

10. U.S.G.S., Working Group on California Earthquake Probabilities, “Probabilities of Large Earthquakes Occurring in California on the San Andreas Fault”, USGS Open File Report 88-398, 1988.
11. U.S.G.S., Working Group on California Earthquake Probabilities, “Probabilities of Large Earthquakes in the San Francisco Bay Region, California”, U.S. Geological Survey, Circular 1053, 1990.
12. U.S.G.S., “Seismic Sources and Recurrence Rates As Adopted by USGS Staff for the Production of the 1982 and 1990 Probabilistic Ground Motion Maps for Alaska and the Conterminous United States”, Open-File Report 95-257, 1995.

Wesnousky, S. G., “Earthquakes, Quaternary Faults, and Seismic Hazard in California”, *Journal of Geophysical Res.*, 91(B12):12587-12631, 1986