# NATURAL CATASTROPHE PROBABLE MAXIMUM LOSS

By Dr. Gordon Woo
Risk Management Solutions

ABSTRACT

The procedure for estimating Probable Maximum Loss (PML) for natural catastrophes has evolved over the past few decades from a rather simplistic deterministic basis to a more sophisticated methodology based on loss exceedance probability curves, generated using catastrophe modelling software. This development process is reviewed, with an emphasis on the earthquake peril, which, because of its widespread threat to critical industrial installations, has been at the forefront of most PML advances. The coherent risk definition of PML is advocated as an improvement over standard quantile methods, which can give rise to anomalous aggregation results failing to satisfy the fundamental axiom of subadditivity, and so discouraging the pooling of risks.

©2002, Risk Management Solutions, Inc. All Rights Reserved.
Originally Published in the *British Actuarial Journal*, Volume 8, Part V (2002)

## 1. THE DETERMINISTIC TRADITION

### 1.1 *Historical introduction*

The requirement to approximate an upper limit to the loss arising from a fire, explosion or other accidental peril, has always been a basic element of sound insurance practice. The insured will need to fix the ceiling on the cover required, and so will the insurer in deciding how much cover to grant or retain. Loss arises from the occurrence of a hazard event, and vulnerability to this hazard. Therefore any definition of maximum loss should reflect the dual hazard and vulnerability aspects of loss. According to the dictionary of insurance (Bennett, 1992): '*it should be the largest possible loss which it is estimated may occur, in regard to a particular risk, given the worst combination of circumstances'.*

The use of absolutes such as 'largest possible loss' and 'worst combination of circumstances' in the wording of this basic definition, reflects an understandable quest for clarity in the midst of ambiguity, and objectivity in the midst of subjectivity. However, such clarity and objectivity are often illusory. Risk is the possibility of loss occurring (Bennett, 1992). The notion that residual risk need not be quantified is reminiscent of the confident NASA assurance about the Challenger shuttle: 'Since the shuttle is a manned vehicle, the probability of mission success is necessarily very close to 1.0'. This is a bold statement, but prior to the 1986 disaster, no formal quantitative risk assessment had yet been carried out. Post-disaster risk analysis showed the shuttle failure probability to be greater than 1/100.

There is considerable uncertainty over the severity of loss that might arise under a broad variety of hazard event circumstances. But, instead of attempting to gauge the uncertainty, the determinist seeks a conservative loss bound; one which circumvents the complex issue of risk quantification. Scientific laws of Nature are at the core of the deterministic tradition; if a clear upper loss bound may be established from fundamental scientific principles, such as the physics of combustion in respect of the spread of fire, there should be no need to quantify risk.

Economy is one of the attractions of determinism. Diligence in assessing the largest possible loss, and identifying the worst combination of circumstances, may absolve a determinist from the numerical burden of having to model vulnerability and hazard in order to quantify risk. Given the computational labour involved in this modelling task, it is unsurprising that, in a pre-computer age, determinism proved such a compelling doctrine. Those unsure of the technical arguments against risk analysis would have been persuaded by the tedious hours of hand calculations. As with numerical weather forecasting, for as long as it depended upon the arithmetic stamina of clerical staff, numerical risk assessment was never viable. So it was that deterministic estimates of maximum insured loss prevailed more or less unchallenged into the last quarter of the 20th century. For an industry whose very business was risk, natural catastrophe insurance lacked for decades what might seem an essential tool: a workable quantitative risk index to measure catastrophic losses.

### 1.2. *Deterministic estimates of maximum loss*

Determinism was always the abiding philosophy of early practitioners of earthquake hazard. The professions directly concerned with this hazard were civil engineering and geology. Civil engineers traditionally followed deterministic construction practices and building code guidelines in the quest for absolute safety, and dealt with uncertainty not by quantifying it, but by incorporating explicit factors of safety in design procedures. Evaluating risk was not a traditional part of a civil engineer's

©2002, Risk Management Solutions, Inc. All Rights Reserved.

education. Neither were concepts of risk evaluation in the standard working vocabulary or training manuals of field geologists. Given a site location, an engineering geologist would follow a prescriptive procedure to arrive at a ground motion level for seismic design. Where there were practical difficulties in implementation, due to decisions being made in the presence of uncertainty, past case experience might be cited, as in the interpretation of legal statutes. The deterministic procedure consisted of the following steps:

- Identify from a regional fault map those faults posing the greatest threat
- From the length of each fault, estimate the maximum earthquake magnitude
- Estimate the severity of ground shaking at the site, arising from such a maximum earthquake rupture on each fault
- Rank the site-specific ground motion values for each of the faults considered
- Select the highest ranked ground motion value for seismic design purposes

One of the main reference regions for deterministic assessments of PML has been the state of California. The deterministic PML approach is implicit in the prescriptive language of the 1972 Alquist-Priolo Earthquake Fault Zoning Act, the purpose of which has been to prohibit the location of most structures for human occupancy across traces of active faults. The responsibility of the State Geologist to compile fault maps of California has relieved other geologists of this preliminary mapping task which is crucial for implementing a deterministic PML procedure. Utilizing the extensive work done for mapping fault zones in California, the California Department of Insurance (CDI, 1995-1996) define PML deterministically in several ways:

- The insured losses corresponding to an earthquake of magnitude up to 8.25 on the Richter scale (this is the great 1906 San Francisco earthquake scenario)
- The expected shake damage loss given the maximum size earthquake that is likely to occur in a fault zone

No account is taken by engineering geologists of the likelihood of occurrence of the event upon which the hazard assessment is based, except that a fault is only classified as active if there has been surface displacement within the last 11,000 years. (Faults in California typically display some manifestation of surface displacement if the earthquake rupture reaches the surface).

Within a deterministic approach, a very high loss potential might thus be assigned to a site, merely on the slim evidence of activity on a neighbouring fault, which may not have experienced a notable earthquake in many thousands of years. Conversely, a rather modest loss potential might be assigned to a site which is distant from any mapped active fault, but might be close to a fault as yet unmapped, or deemed, from available evidence, to be inactive. Allowing for the inevitable differences in geological opinion on the interpretation of partial evidence, instability and controversy have been hallmarks of many a deterministic seismic ground motion assessment.

©2002, Risk Management Solutions, Inc. All Rights Reserved.

## 2. INITIAL STEPS TOWARDS RISK QUANTIFICATION

### 2.1 *The rationale for a probabilistic approach*

As concern spread over deterministic assessments of seismic hazard, so rival risk-based methods began emerging in the late 1960's (Cornell, 1968). Probabilistic seismic hazard analyses began to be applied in the 1970's in studies for critical industrial installations such as nuclear facilities, petrochemical plants, liquefied natural gas plants, and dams. The rate of progress was slowed by the reluctance of regulators, mindful of the potential for legal challenge, to countenance major innovation such as the introduction of probabilistic methods. A notorious legal confrontation between deterministic and probabilistic assessments of seismic hazard took place over a nuclear test reactor at Vallecitos, California (Meehan, 1984).

Setting aside the deterministic tradition, which found favour with lawyers, the original rationale for the development of probabilistic risk assessment for such critical installations was stark. For an ordinary factory, the worst accident might destroy all the factory buildings and the workers within. But for a nuclear facility, the worst accident might transcend site boundaries, and lead to the deaths by irradiation of many people off site. Given that no nuclear facility, however well deterministically designed, can be entirely safe, the risk to the public has to be quantified. In acknowledgement of this necessity, the principles of probabilistic risk assessment were developed in the 1960's.

### 2.2 *Maximum Credible Earthquake (MCE) based on fault activity rate*

Earthquake risk studies for insurance applications began with important individual risks, such as industrial installations, exposed to seismic activity on nearby faults. Acknowledging that the neglect of fault activity rates was a serious weakness of the deterministic approach, seismic hazard consultants provided a risk-based definition for the expression 'Maximum Credible Earthquake', which had hitherto been used by geologists in a purely deterministic sense. The MCE for a fault was defined as the event with a 10% probability of exceedance in 50 years. Where earthquakes occur as a Poisson process on a fault, this is tantamount to a return period of about 475 years for the MCE, since: $1 - \exp(-50 / 474.6) = 0.1$.

This definition immediately resolved the deterministic conundrum of dealing with faults near an insured site which have not moved in many thousands of years. At the prescribed tolerance level, these can be ignored, and the attention focused on the more recently active faults. Once the MCE is specified for each threatening fault, the implications of this earthquake for a property portfolio remain to be assessed by structural engineers. The simplest option for these engineers is to estimate a conservative deterministic upper bound for the consequent portfolio loss, from a map of ground shaking generated by the MCE. Where such conservatism might be deemed excessive, an alternative is to estimate a high (e.g. 90%) confidence bound on the portfolio loss from the MCE. This refinement would require a statistical analysis of building vulnerability, in which likelihood weights are assigned to different states of earthquake damage from the MCE. If such 90% loss bounds were estimated for each individual fault, then the overall PML for a portfolio would be taken as the highest 90% confidence portfolio loss from any fault, where the defining earthquake on the fault is the MCE with a 475 year return period.

The use of the risk-based MCE in defining PML was certainly a notable improvement over the deterministic procedures which hitherto had held sway, despite their patent shortcomings. Fortunately, such an improvement came at a relatively

©2002, Risk Management Solutions, Inc. All Rights Reserved.

modest cost in terms of professional manpower and computing resources. In particular, no elaborate risk software or expensive hardware were required. As a consequence, this risk-based MCE method proved very popular with earthquake consultants and their industrial and insurance clients.

However, the MCE, being but one step towards a full risk methodology, was not without its own intrinsic problems. The use of 475 years as a reference return period for an earthquake on a fault is arbitrary; other long return periods might be chosen. Furthermore, regardless of return period, any PML definition which is based on fault-specific event recurrence neglects the diverse range of faults which collectively contribute to the earthquake risk to the portfolio. Not just one, but a sizeable number of faults may each generate 475 year return period earthquakes capable of causing major portfolio loss. A multiplicity of hazard sources should be taken into account.

### 2.3 *Maximum Credible Event based on local hazard return period*

Consider a particular urban location where there is a concentration of insured exposure, under threat from a range of scenarios of a regional natural peril. Instead of defining a PML in terms of just one specified source event, (e.g. a single fault rupture or single storm track), it can encompass all sources by being defined in terms of the local hazard exceedance curve, which charts the annual probability that a given hazard level is locally exceeded. All regional hazard sources contribute to this curve, and a PML can be defined in terms of the curve at some specific return period.

A hurricane illustration of this PML approach is provided by a study undertaken of critical infrastructure in parts of the Caribbean, including the port of Kingston, Jamaica (Stubbs, 1999). This study was an important component of the Caribbean Disaster Mitigation Project supported by the Organization of American States and the US Agency for International Development. The objective was to outline options for strengthening the regional insurance industry, and increasing risk retention through reducing risk exposure and vulnerability. An essential step towards this objective was the estimation of PML.

Statistical analysis of 112 years of meteorological data from the port of Kingston yielded local windspeed values for 50 and 100 year return periods. From fitting a two-parameter Weibull distribution using Maximum Likelihood methods, median (i.e. 50%) and 90% confidence windspeeds were output for both return periods. Four extreme windspeed values were thus obtained:

- 50% confidence 50 year return period = 91 knots
- 90% confidence 50 year return period = 105 knots
- 50% confidence 100 year return period = 103 knots
- 90% confidence 100 year return period = 124 knots

For each of these four extreme windspeed values, the probability of failure of the infrastructure elements was assessed, and the expected loss calculated as the product of the failure probability and the replacement cost. By aggregating the loss over all infrastructure elements, and assuming no significant spatial variation in the extreme windspeed, a simple approximate PML estimate was obtained without recourse to a more sophisticated hurricane catastrophe model. The degree of flexibility in the choice of windspeed return period and in the confidence level was used to output four alternative PML values, which were helpful in formulating insurance strategy.

©2002, Risk Management Solutions, Inc. All Rights Reserved.

2.4 *Seismic hazard exceedance probability curves*

Unlike hurricanes, there is no annual season for earthquakes, and for earthquake shaking at a site, the long time intervals between major event occurrences render impractical, if not dubious, attempts to estimate the seismic hazard by fitting statistical distributions to local measurements of ground motion. However, a mathematical procedure for quantifying hazard, by considering all earthquake sources that might affect an insured site, was outlined by Cornell as far back as 1968. Being comparatively data intensive, it took several decades for this procedure to be justified for insurance applications.

The basis of this procedure was just the total probability theorem. Let $f(M, \mathbf{x})$ be the frequency of earthquakes of Magnitude M, epicentred at **x,** and let $P(Z \geq z \mid M, \mathbf{x})$ be the conditional probability that the site ground motion $Z$ (e.g. peak acceleration) is greater or equal to $z$, given the occurrence of this event. Then the expected annual number of exceedances of this ground motion level at the site is given by the expression:

$$\nu(z) = \int dM\, d\boldsymbol{x}\; f(M, \boldsymbol{x})\, P(Z \geq z \mid M, \boldsymbol{x})$$

For a Poisson process, the annual probability of exceedance of ground motion $z$ is just $1 - \exp(-\nu(z))$.

On the resulting hazard curve, a specific point of earthquake engineering interest is the exceedance probability of 1/475, which correponds to a 475 year return period for the site. This return period is quite often used as an international building code criterion for a seismic design basis. The primary focus of building codes is human safety. The Uniform Building Code (ICBO, 1994) includes the 475 year return period requirement for life-safety design for a building or structure. If the construction at a given site has been designed for ground shaking with a 475 year return period, then the loss resulting from this level of seismic action should not be so calamitous as to be life-threatening.

Irrespective of whether such a design basis has been used, a site PML can be defined in terms of the ground motion with this site return period. As with the fault-based MCE, a conservative estimate of maximum site loss could be made deterministically, from knowledge of the construction on the site. Less conservatively, a 90% confidence loss estimate may be assigned by reference to vulnerability functions for structures of the type found on the site. The 90% figure reflects the wording of definitions, such as developed by the California Department of Insurance, that PML should be the expected shake damage loss experienced by 9 out of 10 buildings.

In this PML procedure, hazard and damage are treated here in distinct ways, and not integrated within an overall systematic risk-based approach. On the one hand, seismic hazard is probabilistically quantified, with the hazard contribution of all earthquake event scenarios weighted according to their likelihood of occurrence. On the other hand, damage is estimated either deterministically or statistically at a specified confidence level. Accordingly, this is a quasi-probabilistic hybrid approach, which has the merit of encompassing all sources of seismic threat in a probabilistic manner, but it still falls short of representing loss in a unified and consistent way.

Furthermore, where a portfolio is sufficiently spatially dispersed for there to be significant variations in 475 year return period ground motion, aggregation of PML has the inconsistency that the ground motion levels at the constituent sites may be

©2002, Risk Management Solutions, Inc. All Rights Reserved.

mutually incompatible. The 475 year return period shaking at different portfolio locations might never be physically realizable in any actual event. Thus the PML, conservative though it might be, would only represent some notional loss.

## 3. CALCULATION OF LOSS PROBABILITY

### 3.1 *Portfolio loss exceedance probability curves*

Earthquake insurers have always ideally needed to construct an exceedance probability curve for loss to a portfolio covering multiple sites distributed over a wide geographical region. No specific fault would be singled out; this exceedance probability curve would implicitly encompass all conceivable hazard event scenarios, designated by Magnitude M and epicentre **x**, weighted according to their frequency of occurrence (e.g. Woo, 1999).

The appropriate procedure to achieve this aspiration is essentially that of the seismic hazard analysis described above for ground motion exceedance at a single site. In order to extend this analysis to cover loss at a portfolio of sites, the same Cornell approach can be followed, except that instead of the exceedance parameter being ground motion at one site, it is the portfolio loss aggregated over all sites.

Portfolio loss exceedance curves began to be developed in the late 1980's, at a time when such analysis became feasible on a desktop computer; this was neither practically convenient nor economically viable on a mainframe computer. However, due to restrictions of available PC power, the early earthquake models were rather skeletal in seismic source representation. As computer performance has escalated in the years since, so also has the complexity and spatial resolution of the earthquake models. Commercial impetus to improve insurance loss modelling was given by two destructive earthquakes in California: Loma Prieta (1989) and Northridge (1994).

A natural risk index for PML purposes is a high quantile of the annual exceedance probability (EP) curve. A loss with an annual exceedance probability of P has a return period of 1/P years. The establishment of insurance limits by reference to a loss return period has the great virtue of simplicity, and the concept of a return period is readily understandable across tiers of insurance management. The success of the quantile index for PML purposes has been so resounding that, for those perils and territories for which catastrophe modelling software exists, this quantile index has largely superseded its less elaborate and less sophisticated predecessors.

Official endorsement has encouraged its usage still further: the Natural Disaster Coalition (NDC, 1995) has defined PML in this probabilistic manner, setting a low annual exceedance probability of 0.002, equivalent to a return period of 500 years. The computation of a loss exceedance probability curve is far more intensive in model development and hardware resources than any analysis previously considered for PML estimation. But this is not an excessive effort: PML estimation is of course only one application of the EP curve, which is a foundation for many diverse actuarial calculations.

### 3.2 *PML anomalies in portfolio aggregation*

In the earlier sections 1 and 2, methods were described for estimating PML for a single site or localized urban concentration of exposure, or for a regional portfolio of sites. In a manner which is economical with both resources and rigour, aggregation using these methods has often been achieved by simple addition of PML's. This gross

©2002, Risk Management Solutions, Inc. All Rights Reserved.

simplification deftly side-steps the problem of calculating the geographical correlation of loss between different hazard events.

Thus if the PML for a Southern California portfolio was based on a Los Angeles earthquake scenario, and the PML for a Northern California earthquake was based on a San Francisco Bay earthquake scenario, the overall California PML might be taken to be the sum, even though the event scenarios were quite distinct. The LA scenario might generate some loss in Northern California, but it would not cause the PML loss.

Similarly, the San Francisco Bay earthquake scenario might cause some loss, but not the PML loss, in Southern California.

This coarse aggregation procedure would typically be adopted for any of the traditional methods reviewed in sections 1 and 2. PML's would be added regardless of whether each individual scenario was selected deterministically; or as a Maximum Credible Earthquake associated with a 475 year fault activity rate; or as a Maximum Credible Earthquake having a given hazard exceedance probability in a metropolis. The simple addition of PML's may be excused as being convenient, as well as commensurate with the simplistic character of these methods.

By contrast, the loss exceedance probability approach is more sophisticated than its deterministic or quasi-probabilistic counterparts, and can be used by a conscientious analyst to aggregate PML's more methodically. Thus if PML's for Northern and Southern California were obtained from separate regional EP curves, then a PML for the whole state might be obtained from a state-wide EP curve. Through making the effort to pool the regional portfolios, some diversification benefit would be expected in the form of a joint PML smaller than the sum of the two regional PML's. This benefit is realized with California earthquake hazard, but is this always the case?

The exceedance probability curve quantile has been widely adopted for PML purposes over the past decade. Although this index has been well tried and tested over the years, its usage has been the source of confusion in the interpretation of some results, which has led on occasion to the false suspicion of numerical error or computer programming bug. The problem is that apparent inconsistencies can arise which seem to defy common sense, and thus are very hard for insurance analysts to understand for themselves or explain convincingly to their managers. An illustration of inconsistency is given below. This is a very elementary version of the California example previously cited.

Consider two factories A and B which are sufficiently distant not to be affected by the same earthquake. Given that no earthquake could affect both factories, the earthquake PML for (A+B) should be less than the sum of the PML's for A and B.

There should be some diversification benefit in pooling the two factories together into a joint portfolio. However, the quantile PML does not automatically satisfy this requirement, which is fundamental enough to be deemed to be axiomatic. If earthquake loss distributions were Gaussian, then the quantile PML would comply with this axiom; but of course earthquake loss distributions are much heavier-tailed.

To highlight this apparent paradox, consider two widely separated independent portfolios A and B. Each regional portfolio is exposed to a tectonic environment where moderate magnitude earthquakes are quite frequent, but not particularly damaging; but where large destructive events occur with a longer return period of, say, 200 years. Then the 100 year PML for both portfolio A and B would be small. However, the 100 year PML for the joint portfolio is large, because this takes into account the 200 year return period destructive events striking either portfolio A or B.

Further examples of anomalies in the use of quantiles for insurance risk

©2002, Risk Management Solutions, Inc. All Rights Reserved.

management have been given by Evans (2001). These examples of the discouragement of risk sharing between independent risks are all the more disquieting for being fairly easy to contrive. If the Pareto distribution is used as a model for a loss distribution, then examples are readily obtainable by assuming a reduction of the shape parameter for a joint portfolio, which corresponds to a thickening of the tail.

## 4. COHERENT RISK MEASURES

### 4.1 *Four risk axioms*

The history of the development of risk indices has been an evolutionary process, with less well adapted methods falling into disuse, and the advent of new technology catalyzing the introduction of more elaborate computer-intensive risk definitions.

From time to time, this history has been punctuated by episodes of accelerated progress. One such current episode is linked with addressing directly the following question: *what fundamental characteristics should a measure of financial risk have?*

This is the kind of challenging question that one might imagine would have at most an academic answer, given that the question must have been pondered many decades ago. But whatever little motivation there may have been to address this question in the past, the development of derivatives trading has turned this ostensibly philosophical question into one of acute practical interest in the banking world. It is from this commercial source that the initiative came to establish an answer, which introduced the essential concept of a coherent risk measure. Perhaps surprisingly, this is an answer of real and immediate practical application. Artzner, Delbaen, Eber and Heath (1999) presented four axioms which collectively define a coherent risk measure. These axioms cover any financial position, but here, for illustrative purposes, these four axioms are restated in terms of PML for one natural peril, which might be earthquake, windstorm, flood, etc..

From the viewpoint of risk to a property portfolio, the set of states of nature of insurance concern is the set of possible regional peril event sequences during a year.

For each such sequence, a random variable A is defined as the highest loss a particular portfolio sustains from an event happening during the year. The loss to the portfolio is the net loss, accounting for deductibles, coverage limits, reinsurance etc.. (For ease of notation, the portfolio is also labelled as A). One historical realization of this random variable might then be {0,0,0,10,0,0,5}, if the only net losses were in the 4th and 7th years, and the highest net losses in these years were 10 and 5 (in a standard monetary unit).

©2002, Risk Management Solutions, Inc. All Rights Reserved.

The question is how to define a meaningful measure of maximum annual event net loss in terms of such a random variable. There are many possible choices. But the alternatives are greatly reduced in number if some algebraic constraints are imposed. Artzner et al., define a measure as coherent if it satisfies four basic criteria, which are listed below.

(a) Translation Invariance:

$$PML( A + c ) = PML( A ) + c$$

If a constant loss amount c (e.g. a fixed cost or expense for positive c; or a fixed subsidy or reimbursement for negative c) is applied to the (net) portfolio loss for the worst annual event, then the PML is changed by the same amount. This holds even where, for example, PML(A) is 100%. In this case, a universal loss surcharge corresponds implicitly to an effective increase of c in the limit.

(b) Positive Homogeneity:

$$PML( k \,.\, A) = k \,.\, PML(A) \quad \text{(where } k > 0)$$

A change k in the monetary scale of the portfolio exposure (e.g. from a currency conversion) changes the PML by the same factor. (For large values of k, the effects of demand surge inflation may introduce some nonlinearity, but this is explicitly taken into account in modelling the cost of damage).

(c) Monotonicity:

$$\text{If } A \geq B \qquad \text{then } PML( A ) \geq PML( B )$$

If, whatever peril events occur during the year, the largest event loss to portfolio A is greater or equal to the largest event loss to portfolio B, then the PML for portfolio A must be greater or equal to that of B.

(d) Subadditivity:

$$PML( A+B ) \leq PML( A ) + PML( B )$$

There should be some diversification benefit in pooling risks. The PML for the joint A and B portfolios should not exceed the sum of the PML's for A and B separately.

The first two (a) and (b) rule out complicated nonlinear algebraic functions of event loss such as: $PML = f(Loss^{\gamma}) \quad (\gamma \neq 1)$. A root-mean-square formula such as $PML = \sqrt{E(Loss^2)}$ would satisfy the scaling criterion (b), but fail the translation criterion (a). The third criterion rules out standard deviation terms such as appear in some pricing formulae: $PML = E(Loss) + c \,.\, \sigma(Loss)$. With such expressions, small but volatile losses may give rise to a lower PML than sizeable but stable losses. The fourth criterion rules out quantile-based PML's.

©2002, Risk Management Solutions, Inc. All Rights Reserved.

If it were difficult to define and evaluate a PML satisfying all these four criteria, then insurers might have to settle for partial fulfillment of these criteria. However, it turns out that there is a PML definition which both satisfies all these criteria, and is computationally relatively easy to evaluate. This definition makes more extensive use of the tail of the loss distribution, as explained next.

4.2 *Expected Shortfall*

An obvious characteristic of the simple quantile definition of PML is that it ignores the tail of the loss distribution beyond the quantile point. In Fig.1, several possible alternative tails beyond the risk tolerance level $\alpha$ are extrapolated.

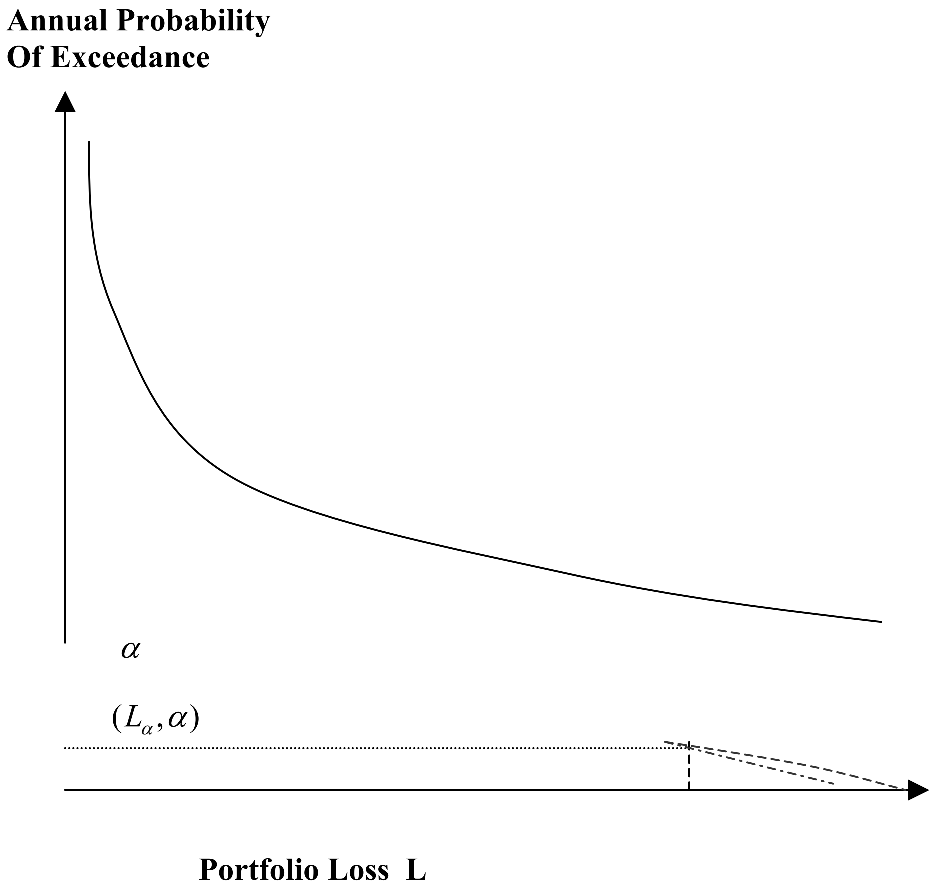


Fig.1: Schematic Exceedance Probability plot, showing alternative tail behaviour.

This deficiency can be rectified through introducing a coherent definition of PML.Of the possible candidates, the most natural and simplest definition in the present context is the Expected Shortfall (Acerbi and Tashe, 2002). This is most conveniently explained and motivated using the framework of a Monte Carlo simulation.

To start, set the risk tolerance level to be $\alpha$. Suppose that N years of regional activity of a natural peril are simulated. These many random simulations are assumed to be equally probable. For each year, the worst (net) portfolio peril loss is identified.

Rank these N losses in descending order. Let M be $N\alpha$, rounded up to the nearest integer. Then the Expected Shortfall $ES^{(\alpha)}$ is estimated simply as the average of the top M losses. Note that, even if the M'th loss is shared by further years, it is only the top M losses which are averaged. It is easy to see that $ES^{(\alpha)}$ is subadditive:

©2002, Risk Management Solutions, Inc. All Rights Reserved.

the set of top M losses to a joint portfolio is, at worst, the same as the set of top M losses to the individual portfolios.

As an elementary example, suppose that 1000 years of regional peril activity are simulated, and that the risk tolerance level is 1/200. Then M is 5. Suppose the top five losses (in a standard monetary unit) are: 5, 6, 8, 9, 12. Then the Expected Shortfall = 40/5 = 8. By comparison, a quantile-based estimate $L_\alpha$ would be the fifth largest loss, which is 5. Suppose that the top sixth loss happens also to be 5. This makes no difference to the Expected Shortfall. However, it would dilute a risk index based on the average over all losses of size $L_\alpha$ or above.

This observation leads on to a somewhat less conservative risk measure than Expected Shortfall: the Tail Conditional Expectation (TCE). This is defined as the conditional expectation of the losses greater or equal to $L_\alpha$. The inclusion of a potentially large probability weight for a loss equal to $L_\alpha$, such as is quite common with excess of loss contracts, may reduce the conditional expectation below that of Expected Shortfall. This dilution of loss has the effect that TCE is not always subadditive. But, as with Expected Shortfall, it is obvious that TCE satisfies the other three criteria for coherence.

For a continuous loss distribution, the conditional expected loss over the tail may be expressed as the tail integral:

$$(1/\alpha) \int_{L_\alpha} dL\, L\, [\, dP(Loss \le L)/dL]$$

Acerbi and Tasche (2002) have shown that, through introduction of a generalized inverse function of the cumulative loss distribution, this tail integral defines the Expected Shortfall in situations where the loss distribution is discrete.

This tail integral is comparatively simple to evaluate for a number of parametric probability distributions. A particularly interesting example, because of its close relation to extreme value distributions (Embrechts et al., 1997), is the Generalized Pareto Distribution.

$$P(Loss \le L) = \quad 1 - (1 + \xi L / \beta)^{-1/\xi} \qquad (\xi < 1)$$

In this case, the risk tolerance level $\alpha = (1 + \xi L_\alpha / \beta)^{-1/\xi}$, which defines the quantile loss value $L_\alpha$. Integration by parts yields the simple expression: $(L_\alpha + \beta) / (1 - \xi)$. When the scaling index $\xi$ is close to zero, the fall-off with L is rapid, and the expression may not be much greater than the quantile loss $L_\alpha$. However, when the index $\xi$ is close to unity, the fall-off with L is slow, and the expression may be a sizeable multiple of $L_\alpha$.

### 4.3 *Practical usage of Coherent Risk*

A contributory factor in the non subadditivity of the quantile PML measure $L_\alpha$ is the lack of sensitivity to the length and shape of the tail beyond this point. Ideally, the PML should be increased if the tail were extended; and the PML should be reduced if the tail were short. Since $ES^{(\alpha)}$ is explicitly tail-dependent, this desirable characteristic of a PML can be attained by using this measure instead.

Extended tails are often associated with intraplate areas of moderate seismicity. Montreal, in eastern Canada, is a classic example of a city several hundred kilometres

©2002, Risk Management Solutions, Inc. All Rights Reserved.

distant from the most likely source of regional earthquakes. Thus a 250-year quantile PML would be small. However, there is small probability of a highly damaging magnitude 6 earthquake occurring very close to Montreal. An analysis of an actual Montreal insurance portfolio using the RMS Canadian earthquake model yields an Expected Shortfall which is three times the quantile value.

Clearly, for a fixed $\alpha$, $ES^{(\alpha)}$ is more conservative than $L_{\alpha}$, since it is an average of losses at least as great as this. However, there is no reason why the tolerance level for the Expected Shortfall should equate with that for the quantile PML measure. Thus, in practice, $ES^{(\beta)}$ might be defined with $\beta$ chosen so that $ES^{(\beta)} \approx L_{\alpha}$, for a typical form of tail distribution. Thus if a traditional tolerance level of 1/250 or 1/200 is used for the quantile PML, then a higher value of about 1/100 might be appropriate for $\beta$.

Exceedance probability curves are most typically generated for aggregated portfolios comprising a collection of properties. Refinement of these curves is restricted practically by the poor quality of information often provided to reinsurers on property location. As a supplement, exposure databases synthesized by risk modellers from economic information provide surrogate data on property location. The acquisition of more detailed site-specific information, and increasing portfolio spatial resolution, will enable exceedance probability curves to be generated for important individual properties such as industrial installations, factories, and commercial building complexes. If an energy insurer covered petrochemical plants located around the world, and generated loss exceedance probability curves for each, it might estimate PML's using quantiles of the individual curves. However, the loss exceedance curve for a well designed, constructed, and maintained plant might have a long tail, so the Expected Shortfall might be a preferable PML measure.

Urgent developments in the theory and practice of probabilistic fire modelling, hastened by the World Trade Center disaster, may allow fire loss exceedance curves to be produced for individual centres of high exposure concentration, such as landmark towers. The issue of PML definition and aggregation will become ever more acute as more and more property-specific long-tailed exceedance probability curves are generated.

### 4.4 *Implementation in a catastrophe risk model*

Since the late 1980's, commercial catastrophe models have been developed to assist insurers with managing their long-tail property risks. Inter alia, such models output loss exceedance probability curves, the quantiles of which may be used in applications such as reinsurance pricing, apart from setting PML's. Some catastrophe models provide users with scenario loss tables, with columns for event frequency and associated portfolio loss. Simple spreadsheet manipulation would then allow a user to calculate Expected Shortfall values. In due course, this spreadsheet labour will become redundant when Expected Shortfall is evaluated and encoded within the model software itself.

### 4.5 *Subadditivity as a criterion for choice of PML*

Although the concept of risk coherence is one which only recently has been formalized, the idea of using the conditional expectation of tail losses as a PML index is far from novel. From time to time, commercial studies for insurers have been undertaken in which this index has been computed. Furthermore, in other areas of insurance interest, such as capital allocation and dynamic financial analysis, the significance of expected deficit as a risk index has been long recognized, and it has

©2002, Risk Management Solutions, Inc. All Rights Reserved.

also been suggested as a parameter for understanding the pricing variability of catastrophe bonds. The specific use of a coherent measure of risk for risk-based capital allocation is being intensively researched (e.g. Denault, 2001, Singh, 2002).

What makes a difference is the existence of a sound mathematical reason for preferring this PML index to one based on a loss quantile. Until the work of Artzner et al., there was no imperative rationale for switching PML indices. Now there is such a rationale: to uphold the principle of subadditivity of risks. Acerbi and Tasche (2002) view this as such a key characteristic of a risk measure, that the adjective coherent may eventually become redundant.

The breakdown of subadditivity is not a mere mathematical nicety; it may lead to misconceived schemes for risk management. Suppose that an insurer accepts the premise that the PML for joint portfolios (A+B) exceeds the sum of the separate PML's for portfolios A and B. Then by ceding a proportion $\lambda$ of business, he would imagine he might reduce the PML by an even larger proportion. In reality, the PML might actually be reduced by a smaller proportion than $\lambda$. It is clear that this kind of misunderstanding could result in risk mismanagement, of a kind that eventually may attract the attention of financial regulators.

It is universally appreciated (e.g. Zadeh, 2000) that the proliferation of PML definitions is confusing, and can lead to problems for insurance decision-makers.

Different PML values are liable to be obtained, even with the same underlying models and data. The dispersion of PML values may have masked some of the problems arising from the breakdown of subadditivity: even if the joint PML exceeded the sum of the separate PML's for one model, it might be less than the sum for another. Insistence on coherence would help to resolve the PML quandary.

©2002, Risk Management Solutions, Inc. All Rights Reserved.

## 5. REFERENCES

ACERBI, C. & TASCHE, D. (2002) Expected shortfall: a natural coherent alternative to value at risk, *Economic Notes by Banca Monte dei Paschi di Siena SpA*, **31**, 379-388.

ARTZNER, P., DELBAEN, F., EBER, J-M & HEATH, D. (1999) Coherent measures of risk. *Mathematical Finance*, **9**, 203-228.

ARTZNER, P. (2000) Application of coherent risk measures to capital requirements in insurance. *North American Actuarial Journal*, **2**, 11-25.

BENNETT, C. (1992) *Dictionary of insurance*. Pitman, London.

CALIFORNIA DEPARTMENT OF INSURANCE (1995-1996) *California earthquake zoning and probable maximum loss evaluation program,* Los Angeles.

CORNELL, C.A. (1968) Engineering seismic risk analysis. *Bull. Seism. Soc. Amer.*, **58**, 1583-1606.

DENAULT, M. (2001) Coherent allocation of risk capital. *Journal of Risk*, **4**, 1-34.

EMBRECHTS, P., KLUPPELBERG, C. & MIKOSCH, T. (1997) *Modelling extremal events*. Springer, Berlin.

EVANS, J. (2001) Pitfalls in the probability of ruin type risk management. *Casualty Actuarial Society Forum*, 501-510.

INTERNATIONAL CONFERENCE OF BUILDING OFFICIALS (1994). *Uniform Building Code.* Whittier, California.

MEEHAN, R.L. (1984) *The Atom and the fault*. M.I.T. Press.

NATURAL DISASTER COALITION (1995) *Catastrophe Risk: A National Analysis of Earthquake, Fire Following Earthquake, and Hurricane Losses to the Insurance Industry.* Washington D.C.

SINGH, M.K. (2002) Risk-based capital allocation using a coherent measure of risk. *Journal of Risk Finance*, **3**, 34-45.

STRUCTURAL ENGINEERS ASSOCIATION OF CALIFORNIA (1995) *Probable maximum loss, a rational approach.* SEASC Convention, Indian Wells, California.

STUBBS, N. (1999) *A probable maximum loss study of critical infrastructure in three Caribbean island states.* Report for USAID Office of Foreign Disaster Assistance.

TASCHE, D. (2002) Expected shortfall and beyond. *Journal of Banking and Finance,* **26**, 1523-1537.

WOO, G. (1999) *The mathematics of natural catastrophes*. Imperial College Press, London.

ZADEH, M.M. (2000) Understanding risk management. *In Financial Management of Earthquake Risk.* EERI, Oakland.

©2002, Risk Management Solutions, Inc. All Rights Reserved.
Originally Published in the *British Actuarial Journal*, Volume 8, Part V (2002)