



# **A Bernoulli Formulation of the Land-Use Portfolio Model**

By Richard A. Champion

Report Series 2008–1310

**U.S. Department of the Interior**  
**U.S. Geological Survey**

**U.S. Department of the Interior**  
DIRK KEMPTHORNE, Secretary

**U.S. Geological Survey**  
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2008

For product and ordering information:  
World Wide Web: <http://www.usgs.gov/pubprod>  
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,  
its natural and living resources, natural-hazards, and the environment:  
World Wide Web: <http://www.usgs.gov>  
Telephone: 1-888-ASK-USGS

Suggested citation: Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted material contained within this report.

Champion, R.A., 2008, A Bernoulli formulation of the Land-use Portfolio Model: U.S. Geological Survey Open-File Report 2008-1310, 25 p. [<http://pubs.usgs.gov/of/2008/1310/>].

## Contents

<b>Introduction</b> .....	1
<b>Model Specification</b> .....	4
Definitions .....	4
Bernoulli Formulation.....	6
Axioms .....	7
LUPM Summary Statistics .....	10
Wealth and Losses .....	10
Return on Mitigation Investment.....	12
Simple Return .....	13
Probabilistic Return .....	13
<b>Discussion</b> .....	15
Parameterization and Data .....	15
Selection of Axioms.....	15
Rare Events.....	16
Uncertainty and Variation .....	16
<b>Scenario versus Risk Analysis</b> .....	19
<b>Conclusions</b> .....	20
<b>References</b> .....	21
<b>Appendix: Derivations</b> .....	23

## Figures

<b>Figure 1.</b> ArcGIS LUPM Extension.....	4
---	---

## Tables

<b>Table 1.</b> Natural-hazards mitigation-decision framework.....	1
<b>Table 2.</b> Numerical examples of loss and risk. ....	11
<b>Table 3.</b> A numerical experiment on measures of variation for the Land-Use Portfolio Model.....	18
<b>Table 4.</b> Scenario summaries for seismic risk for the San Francisco Bay Area.....	19
<b>Table 5.</b> Product of distinct Bernoulli random variables.....	23
<b>Table 6.</b> The square of a Bernoulli random variable.....	24

# A Bernoulli Formulation of the Land-Use Portfolio Model

By Richard A. Champion

## Introduction

Decision making for natural-hazards mitigation can be sketched as knowledge available in advance (*a priori*), knowledge available later (*a posteriori*), and how consequences of the mitigation decision might be viewed once future outcomes are known (table 1). Two outcomes—mitigating for a hazard event that will occur, and not mitigating for a hazard event that will not occur— can be considered narrowly correct. Two alternative outcomes—mitigating for a hazard event that will not occur, and not mitigating for a hazard event that will occur— can be considered narrowly incorrect. The dilemma facing the decision maker is that mitigation choices must be made before the event, and often must be made with imperfect statistical techniques and imperfect data.

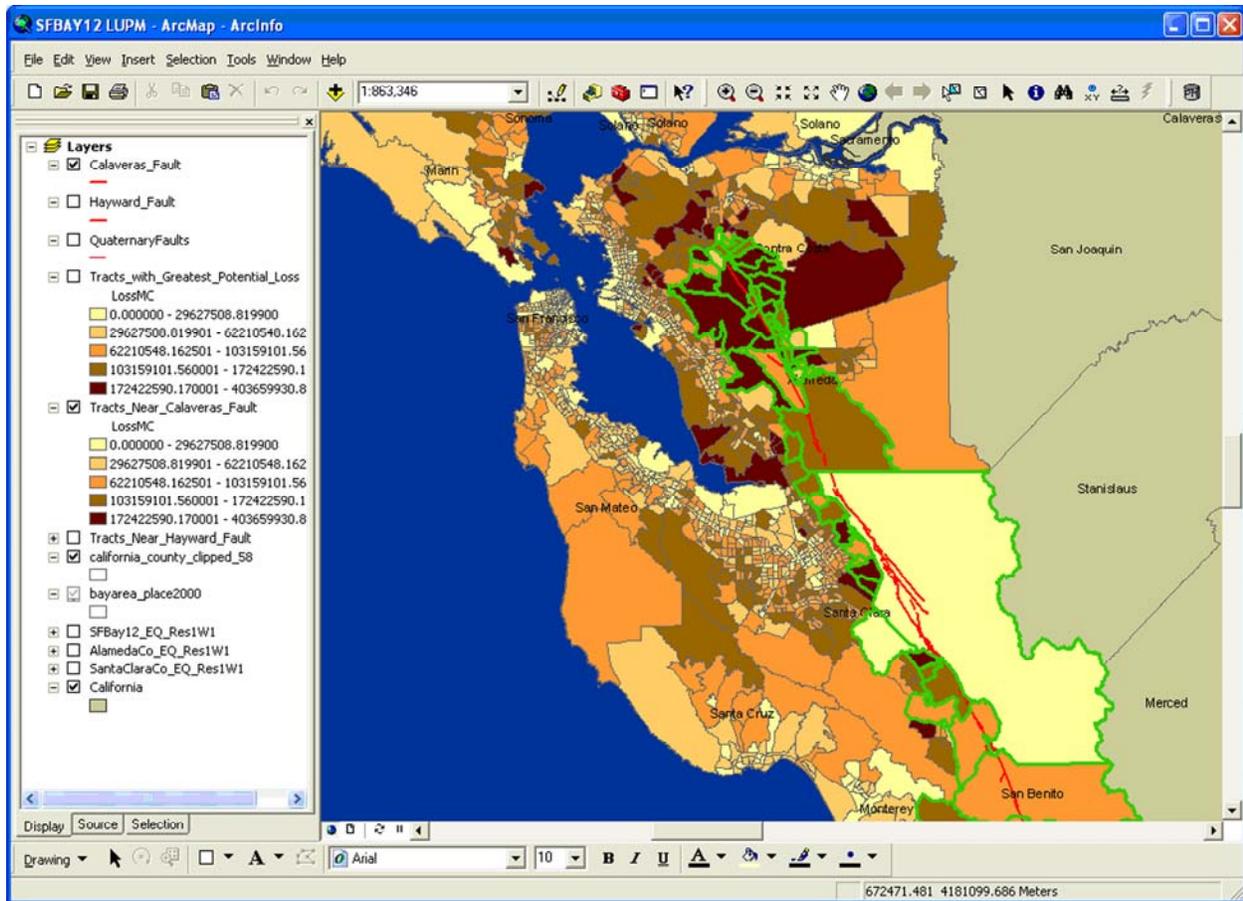
**Table 1.** Natural-hazards mitigation-decision framework.

<b>Planning choices</b>	<b>Natural world outcomes</b>	<b>Consequences</b>
Mitigate	Hazard event occurs	Money spent reduces losses
	Hazard event does not occur	Money spent but no benefit
Do not mitigate	Hazard event occurs	Losses from failure to mitigate
	Hazard event does not occur	No losses, but no money spent

The Land-use Portfolio Modeler (LUPM) is a computer analysis technique that associates numerical values with future natural-hazard outcomes (Bernknopf, and others, 2001; and Bernknopf, and others, 2006; Dinitz and others, 2003a). In financial investing, a portfolio is a combination of securities chosen to provide income at specified levels of risk and return, given a limited investment budget (Markowitz, 1959; and Intriligator, 1971). Analogously, a land-use or natural-hazards mitigation portfolio is a combination of mitigation options related to land use selected to reduce economic losses from natural-hazard vulnerabilities (Dinitz, and others, 2003b). To construct a financial portfolio, the investor identifies a risk preference—a willingness to accept potential losses in exchange for benefits—and then selects a combination of securities to optimize return given the risk preference. To construct a mitigation portfolio, the decision maker identifies a risk preference—such as a willingness to live in a region subject to earthquakes, fires or land slides in exchange for the views or the amenities of a desirable environment—and then applies the LUPM to weigh future outcomes given a limited mitigation budget. Both financial and land-use portfolios use estimates of risk and benefit described on average, with uncertainties of outcome given by statistical measures of variation. There are additional challenges for computing statistics for land-use portfolios that become apparent, for example, data quality. In land-use portfolios, the data may be incomplete and vague. Because natural-hazard events tend to be rare, the underlying statistical distributions tend to be highly skewed. Examples will illustrate how statistical measures derived from skewed statistical distributions may be misleading. Consequently, the statistical algorithms illustrated in this report should be considered as first approximations that will require further development.

LUPM calculations have been implemented as experimental software applications in several configurations with a variety of interfaces. A geographic information systems (GIS) interface is used to show how the occurrence and intensity of natural hazards varies across the landscape.

Figure 1 shows the LUPM interface with experimental software running as an add-on for the ESRI ArcMap desktop application. In this example the software is being used to evaluate mitigation scenarios in the San Francisco Bay Area for earthquakes along the Hayward and Calaveras faults. The LUPM interface allows the user to calculate and rank mitigation options for a variety of natural hazards—fires, floods, earthquakes—as part of a single program run. Output is in the form of tables and charts that describe the scientific and mitigation assumptions; the geographic regions with associated economic data; statistics such as losses and effectiveness of mitigation actions; and measures of variation and uncertainty. Data, mitigation scenarios, and output are stored so that the user can rerun the program to evaluate the effectiveness of a variety of mitigation possibilities. The experimental software is described in an unpublished manuscript, “The Land-use Portfolio Model, Version 1.0” by Richard Taketa, P. Ng ([png@usgs.gov](mailto:png@usgs.gov)), and M. Hong. This report gives definitions of risk-analysis terminology and describes the equations used in the experimental software.



**Figure 1.** ArcGIS LUPM Extension.

## Model Specification

Model specification includes definitions of the economic and risk-analysis terms needed to rank mitigation options, the mathematical description of the economic and risk analysis terms as Bernoulli variables, modeling axioms, and summary statistics. Details of derivations are in the Appendix.

## Definitions

The LUPM associates specific meanings to the terms hazard, loss, risk, and damage. *Hazard* is the natural event of concern, such as an earthquake or flood. *Loss* is a quantitative

measure of the consequences of the hazard event, for example, the dollar value of structures damaged by an earthquake, or the number of fatalities in a flood. *Risk* is the expected value of the loss, a combination of the magnitude of the loss and the probability that the hazard will actually occur. For example, suppose that an earthquake causes a loss of \$100,000,000 if it occurs, but the probability of it occurring during any given year is only 1percent. Thus, the loss is \$100,000,000, and the risk per year is \$1,000,000. *Damage* measures the proportion of loss. If 10 percent of the value of a structure is lost during an earthquake, then the damage is 0.10. The LUPM quantifies the effectiveness of mitigation by using *damage reduction* and *avoided loss*. If 50 percent of the value of a structure would be lost without mitigation, but only 10 percent would be lost with mitigation, then the damage reduction is 40 percent. If the value of the structure is \$1,000,000, then the avoided loss is \$400,000. Comparing the avoided loss to the cost of mitigation gives a measure similar to the rate of return on a security, dollars saved per dollar invested.

The LUPM incorporates the geospatial variability of the severity and consequences of natural hazards. For the specified geographic area, each natural-hazard of concern has an associated *observation period* and a *probability of occurrence* that are used to quantify the time interval during which a hazard event may or may not occur and the probability that the natural hazard will occur during that interval. A *location* (or *site*) is the smallest unit to be mitigated, such as a single structure. A *community* (or *neighborhood*) is a grouping of sites, each with a specified probability of failure given a hazard event. Each location has associated parameters which describe monetary value; cost of mitigation of the site; and, optionally, the probability of loss and extent of damage given the hazard event. The consequences for the community are expressed as initial and end-of-period community wealth and community losses. The *initial community wealth* is the sum of the property values of all locations at the beginning of the observation period and the *end-of-period community wealth* is the sum of the property values of all locations at the end of the observation

period. The *community losses* are the sum of losses due to the hazard event that occurs during the observation period.

## **Bernoulli Formulation**

The LUPM uses Bernoulli random variables (Feller, 1971) to express the chance outcomes of the natural-hazard event. Bernoulli random variables provide a mathematical way of expressing:

- The uncertainty that a natural-hazard event will occur during an observation period.
- The uncertainty that any location in the community will suffer damage if the natural-hazard event occurs during the observation period.
- The vulnerabilities, losses, and damages to locations in the community.
- The economic benefits of mitigation strategies at individual locations.
- The aggregate costs, benefits, and losses of the natural-hazard event to the community.

The Bernoulli random variable  $B_p = B(1, p)$  has two outcomes, 0 and 1:

$$\Pr(B = 0) = 1 - p, \quad \text{and}$$

$$\Pr(B = 1) = p .$$

The special case  $B_{false} = B(1, 0)$  indicates that the specified outcome occurs with probability 0.

This is convenient for describing conditions that cannot occur. For example, if there is no earthquake, then there is no earthquake damage. The special case  $B_{true} = B(1, 1)$  indicates that the specified outcome occurs with probability 1. This is convenient for describing conditions that are certain to occur.

Hazard computations can be made before the hazard event has occurred (*a priori*), or after the hazard event has occurred (*a posteriori*). The question in the *a priori* case is, “What are the costs and benefits of mitigation if we are not sure that the hazard event will actually happen?” The question in the *a posteriori* case is, “What are the losses and the benefits of mitigation now that the hazard event has occurred?” The formulas for the *a priori* and the *a posteriori* cases are related,

the major difference being that the *a priori* calculations contain an explicit probability that expresses the uncertainty as to whether or not the hazard event occurs during the observation period.

This formulation of the LUPM assumes that the outcome given the hazard event at one location does not affect the outcome at any other location. That is, the outcomes are assumed to be conditionally independent. For example, one house may suffer damage in an earthquake, while a nearby house may remain undamaged. For other natural-hazards, such as floods, proximity may imply similar outcomes. In this case the outcomes are said to be conditionally dependent. Conditional independence is a strong simplifying assumption that leads to manageable equations. Conditional dependence is a more general mathematical assumption that, in some cases, will give more realistic statistics of cost, benefit, and loss, however, conditional dependence may lead to a more complex formulation of the LUPM model.

## **Axioms**

If there is no hazard event during the observation period, then the failures or damages across the community will be zero. But if there is a hazard event, then any one site may or may not experience a failure independently of failure at any other site. These conditions are expressed as Bernoulli random variables by using (1) to (4). The *hazard event*  $\delta_H$  indicates the presence or absence of a hazard event during the time period  $T > 0$  with  $\delta_H = 1$  indicating that the event occurs, and  $\delta_H = 0$  indicating that the event does not occur. The probabilities of the presence or absence of a hazardous event are:

$$\Pr(\delta_H = 1) = P_H \quad \Pr(\delta_H = 0) = 1 - P_H . \quad (1)$$

The collection of outcomes across the individual sites is described as an array or list of Bernoulli variables. Each element in the array (2) describes the outcome at the corresponding site given the chance occurrence of the hazard event and the chance occurrence of failure:

$$F = (F_1, F_2, F_3, \dots, F_n). \quad (2)$$

Equations (1) and (2) describe a probability branch. One side of the branch is the absence of the hazard event during the observation period. The other side is the presence of the hazard event during the observation period. If there is no hazard event, then there can be no damage at any site, so that:

$$(F_i | \delta_H = 0) = B_{false} \quad (3)$$

and thus 
$$F(\delta_H = 0) = (B_{false}, B_{false}, B_{false}, \dots, B_{false}).$$

If there is a hazard event, then damage may occur at some sites with outcomes described by Bernoulli variables:

$$(F_i | (\delta_H = 1)) = B(1, p_i) \quad (4)$$

so that

$$F(\delta_H = 1) = (B_1, B_2, B_3, \dots, B_n).$$

The corresponding conditional probabilities of failure along the two probability branches are:

$$\begin{aligned} \text{Probs}(\text{Failure} | \text{No Hazard Event}) &= (0, 0, 0, \dots, 0) \\ \text{and} & \\ \text{Probs}(\text{Failure} | \text{Hazard Event}) &= (p_1, p_2, p_3, \dots, p_n) \end{aligned} \quad (5)$$

Equation (3) indicates that, in the absence of a hazard event, failure at a site is a Bernoulli variable of constant-value zero (or false). Equation (4) indicates that, in the presence of an event, the failure at a site is a Bernoulli variable with probability  $p_i$ . The probabilities in (5) are the result of these probability branches.

Assessing economic loss requires associating an property value to each site:

$$\text{Property Values} = (V_1, V_2, V_3, \dots, V_n) . \quad (6)$$

The simplest possible model assumes that, if there is failure at a site, then the economic loss  $\lambda_i$  is either zero, or one (complete). Losses across the community are then expressed as:

$$\Lambda = (\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n) \quad (7)$$

with  $\lambda_i = V_i * F_i$  .

A more realistic expression allows partial damage. Let the *damage vector* be given by

$$\Phi = (\phi_1, \phi_2, \phi_3, \dots, \phi_n) \quad (8)$$

with  $0 \leq \phi_i \leq 1$  .

The definition on the right in (7) can then be modified so that:

$$\lambda_i = \phi_i * V_i * F_i . \quad (9)$$

The economic value of mitigation is a reduction in partial losses. This is expressed as a *partial damage vector given mitigation*:

$$\bar{\Phi} = (\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3, \dots, \bar{\phi}_n) \quad (10)$$

with  $\bar{\phi}_i \leq \phi_i$  . The reduced damage measures effectiveness of mitigation:

$$\Delta\Phi = (\Delta\phi_1, \Delta\phi_2, \Delta\phi_3, \dots, \Delta\phi_n) \quad (11)$$

with  $\Delta\phi_i = \phi_i - \bar{\phi}_i$ .

## LUPM Summary Statistics

### Wealth and Losses

The *community wealth* is the sum of the property values across all locations of the community. The community wealth at the beginning (*initial community wealth*) and at the end of the observation period (*final community wealth*) are:

$$W_{Init} = \sum_{i=1}^n V_i \quad \text{and} \quad W_{Final} = \sum_{i=1}^n (V_i - \Delta V_i) . \quad (12)$$

The change in community wealth for the observation period is the sum of the losses across the individual locations:

$$\begin{aligned} \Delta W &= W_{Init} - W_{Final} \\ &= \sum_{i=1}^n \Delta V_i \end{aligned} \quad (13)$$

If there is no hazard event, then the loss at each location is zero. If there is a hazard event and mitigation, the maximum loss and the benefit of mitigation (the *avoided loss*) are:

$$\phi_i * V_i \quad \text{and} \quad \Delta\phi_i * V_i \quad (14)$$

Losses occur, given the hazard event, with probabilities as specified in (14). The expected loss and the expected benefit of mitigation for a single location given the occurrence of a hazard event are:

$$p_i * \phi_i * V_i \quad \text{and} \quad p_i * \Delta\phi_i * V_i . \quad (15)$$

However, it is not known, in advance, whether the hazard event will or will not occur. Including this uncertainty in the expected loss and the expected benefit of mitigation gives:

$$P_H * p_i * \phi_i * V_i \quad \text{and} \quad P_H * p_i * \Delta\phi_i * V_i . \quad (16)$$

Perspectives for decision making are suggested by equations (14) to (16), and table 1. Because *risk* is defined as expected loss, the left side of equation (14) shows loss, while the left sides of equations in (15) and (16) show risks. The numerical examples (table 2) show that the risks are less than the maximum losses, and that the unconditional risk is less than the conditional risk. The unconditional risk is less than the conditional risk because the unconditional risk incorporates the uncertainty as to whether or not the natural-hazard event occurs.

**Table 2.** Numerical examples of loss and risk.

Parameters		Unmitigated	Mitigated
Value	\$1,000,000		
Damage	0.30	Maximum loss	\$300,000
Mitigation	0.25	Avoided loss	\$250,000
Probability damage	0.05	Conditional risk	\$15,000
Probability event	0.02	Unconditional risk	\$300
			\$2,500
			\$50

The losses and risks for the community are obtained by summing across the locations. The maximum losses for the unmitigated and the mitigated cases are:

$$\sum_{i=1}^n \phi_i * V_i \quad \text{and} \quad \sum_{i=1}^n \bar{\phi}_i * V_i . \quad (17)$$

The maximum avoided losses are:

$$\sum_{i=1}^n \Delta\phi_i * V_i . \quad (18)$$

The risks for the community are:

$$\text{Conditional community risk} = \sum_{i=1}^n p_i * \phi_i * V_i . \quad (19)$$

$$\text{Unconditional community risk} = P_H * \sum_{i=1}^n p_i * \phi_i * V_i , \quad (20)$$

$$\text{Mitigated conditional community risk} = \sum_{i=1}^n p_i * \bar{\phi}_i * V_i , \text{ and} \quad (21)$$

$$\text{Mitigated unconditional community risk} = P_H \sum_{i=1}^n p_i * \bar{\phi}_i * V_i . \quad (22)$$

The avoided losses for the community are:

$$\text{Conditional community avoided losses} = \sum_{i=1}^n p_i * \Delta\phi_i * V_i \text{ and} \quad (23)$$

$$\text{Unconditional community avoided losses} = P_H \sum_{i=1}^n p_i * \Delta\phi_i * V_i . \quad (24)$$

Letting  $q_i = P_H * p_i$  the unconditional variance of the losses  $\Lambda$  is:

$$\begin{aligned} \text{Var}(\Lambda) = & \left[ \sum_{i=1}^n q_i * (1 - q_i) * V_i^2 * \phi_i^2 \right] \\ & + P_H * (1 - P_H) * \left[ \sum_{\substack{i,j \\ i \neq j}}^n p_i * p_j * \phi_i * \phi_j * V_i * V_j \right] \end{aligned} \quad (25)$$

From (13) the initial and final community wealth are related through the losses (17) and the avoided losses (18), so that summary statistics analogous to (19) through (24) are straightforward.

## Return on Mitigation Investment

The return on mitigation investment is a comparison of the cost of mitigation to the avoided losses. One measure of the return is *simple*, and another is *probabilistic*. Simple rate of return is a ratio of the total cost of mitigation to the total avoided losses. Probabilistic rate of return computes

expected values and variances of avoided losses. Computations are done from the perspective of the community, rather than from the perspective of the individual, which expresses community rather than private benefit. The investment in mitigation is given by:

$$CstTot \equiv Total\ Cost\ of\ Mitigation.$$

### Simple Return

The maximum reduction in losses that the community can achieve through mitigation is given by (18). Comparing this directly to the cost of mitigation gives the simple rate of return (SR):

$$\begin{aligned} SR &= \frac{Sum\ Avoided\ Losses}{Sum\ Mitigation\ Costs} \\ &= \frac{\sum_{i=1}^n \Delta\phi_i * V_i}{CstTot} \end{aligned} \quad (26)$$

### Probabilistic Return

Even with mitigation, there still may be partial losses due to a hazard event. These are described as an array of random variables:

$$\begin{aligned} RCoef &\equiv \frac{\Delta\Phi}{CstTot} \\ &\equiv \frac{(\Delta\phi_1, \Delta\phi_2, \Delta\phi_3, \dots, \Delta\phi_n)}{CstTot} \end{aligned}$$

Following the pattern in (3) and (4) gives the *location-specific rate of return* random variable  $RR_i$  on community investment as:

$$(RR_i | \delta_H = 0) = B_{false} \quad \text{and} \quad (RR_i | (\delta_H = 1)) = \frac{\Delta\phi_i}{CstTot} * V_i * B(1, p_i).$$

These definitions say that the rate of return at a location is 0 if there is no hazard event during the observation period, otherwise, it is given by the ratio of a random variable for the avoided loss at

that location to the total (community wide) cost of mitigation. The community rate of return RR is then:

$$(RR|\delta_H = 0) = B_{false} ,$$

and

$$(RR|\delta_H = 1) = \frac{1}{CstTot} * \sum_{i=1}^n \Delta\phi_i * V_i * B(1, p_i) .$$

These equations affirm that the community rate of return is 0 if there is no hazard event during the observation period. Otherwise, it is the sum of random variables for the location-specific rates of return; or equivalently, it is the sum of the random variables for location-specific avoided losses divided by the total cost of mitigation. The expected value and variance of the community-wide rate of return follow from (32) and (36) in the Appendix:

$$E(RR) = \frac{p_H * \sum_{i=1}^n \Delta\phi_i * V_i * p_i}{CstTot}$$

and

$$Var(RR) = \frac{1}{CstTot^2} \left\{ \left[ \sum_{i=1}^n q_i * (1 - q_i) * (\Delta\phi_i * V_i)^2 \right] + \left[ p_H * (1 - p_H) * \sum_{\substack{i=1 \\ i \neq j}}^n p_i * p_j * \Delta\phi_i * \Delta\phi_j * V_i * V_j \right] \right\}$$

## **Discussion**

The LUPM provides a systematic method to assess quantitatively natural-hazard mitigation choices. The GIS interface provides a method for managing geographic data relevant to the evaluation of mitigation options, for specifying mitigation scenarios, and for managing statistical output.

## **Parameterization and Data**

The LUPM is parameterized by using data describing the probability of natural-hazard occurrence, the engineering response of structures, and economic costs. Preliminary experimentation has required the accommodation to data available rather than data specifically tailored to the requirements of the LUPM. Important avenues for future research will be the acquisition of data better tailored to the axioms specified for the LUPM. For a case study of LUPM parameterization see “Applying HAZUS-MH and the land-use portfolio model to estimate natural-hazard loss and risk: A hypothetical demonstration for Ventura County, California” (Laura B. Dinitz, [ldinitz@usgs.gov](mailto:ldinitz@usgs.gov)).

## **Selection of Axioms**

The LUPM axioms are based on Bernoulli random variables because these lead to statistics that are algebraically tractable and computationally convenient; however the current axioms also are extreme simplifications. For example, the assumption of conditional independence may be too strong given that the occurrence of a hazard event at one location may imply a higher probability of a hazard consequence at a nearby location. The probabilities of damage and of the occurrence of the hazard event currently are assumed to be constant over time. The current formulation of the model ignores the discount rate, that is, the time value of money. Simulation (Ross, 2002) may be

an approach to capturing more realistically probabilistic characteristics of natural-hazards without resorting to complex algebraic reformulation of the LUPM axioms.

### **Rare Events**

Hazard events can be rare, but of high consequence. Although losses from an event may be large, the risk—loss on average—can be small. For sufficiently rare events, losses on average may also be small relative to mitigation cost, which suggests that factors in addition to losses on average ought to be considered when making mitigation decisions. For small probabilities  $p_i$  the Bernoulli random variables in (4) converge to Poisson random variables (Simons and Johnson, 1971; Freund and Walpole, 1987). For these applications another possible improvement might be to formulate the LUPM axioms in terms of Poisson rather than Bernoulli random variables.

### **Uncertainty and Variation**

Mitigation decisions may need to be made given uncertainty (Bernknopf and others, 2007). Statistical measures of variation, such as summary statistics (variation or standard deviation) and probability distributions are ways of quantifying uncertainty. Choices among mitigation options are clearer when uncertainty is smaller rather than larger. For example, a small range of variation in potential losses and a smaller range of variation in mitigation costs present the decision maker with easier choices than do large uncertainties. One approach to assessing variation is to compare the risk (20) to the standard deviation as derived from the variance (25). This compares losses on average to a range of variation around the average.

The variance (25) contains four independent variables. These are the property values, the partial damage, and two hazard probabilities (the probability of the event occurrence and the probability of damage given the event). All the independent variables, except the probability of the event occurrence, can take on site-specific values. The following two examples suggest possibilities

for statistical output from the LUPM. The first example (table 3) fixes all the variables in (25) except the probability of the event occurrence. The annualized probability of the hazard event  $P_H$  is calculated from the average waiting time for the hazard event to occur. If the hazard event occurs, on average, once every 5 years, the annualized probability is 0.20. The remaining model parameters are set at  $p_i = .01$ ,  $V_i = 500,000$ ,  $\phi_i = 1$ , and  $n = 10,000$ . Table 3 shows that the risk drops steadily with time as is to be expected from (20). The steady decline in the standard deviation is suggested by the fact that the coefficient of the second term in (25) declines to zero as the average waiting time to the hazard event increases. The last column of table 3 shows the ratio of the risk to the standard deviation. The closer this ratio is to zero, the narrower the variation in potential losses will appear to the decision maker.

The second example is based on LUPM mitigation outcomes for a 7.0 earthquake near the Calaveras and Hayward faults in the San Francisco and Monterey Bay Areas (Peter Ng, oral communication, Aug. 2008, US Geological Survey, [png@usgs.gov](mailto:png@usgs.gov)). There are ten scenarios chosen to vary distance from the fault, vulnerability, extent of mitigation, and mitigation cost (table 4). In this case there are variations among the risks and standard deviations across the ten scenarios that reflect the geospatial variation of property values and vulnerabilities. The constant ratios (153 percent) suggest that the uncertainties around the mean are large and, on a proportional basis, identical. This suggests statistical similarities among the ten mitigation options.

**Table 3.** A numerical experiment on measures of variation for the Land-Use Portfolio Model.

<b>Average waiting time in years</b>	<b>Annualized probability of hazard event</b>	<b>Risk in U.S. dollars</b>	<b>Standard deviation, in U.S. dollars</b>	<b>Ratio, percent</b>
2	0.5000	25,000,000	252,463	1.0
3	0.3333	16,666,667	237,446	1.4
4	0.2500	12,500,000	217,931	1.7
5	0.2000	10,000,000	201,234	2.0
7	0.1429	7,142,857	175,971	2.5
10	0.1000	5,000,000	150,823	3.0
20	0.0500	2,500,000	109,539	4.4
25	0.0400	2,000,000	98,484	4.9
50	0.0200	1,000,000	70,353	7.0
75	0.0133	666,667	57,636	8.6
100	0.0100	500,000	49,997	10.0
200	0.0050	250,000	35,442	14.2
300	0.0033	166,667	28,962	17.4
400	0.0025	125,000	25,092	20.1
500	0.0020	100,000	22,449	22.4

**Table 4.** Scenario summaries for seismic risk for the San Francisco Bay Area.

<b>Scenario</b>	<b>Selection</b>	<b>Mitigation cost, in percent</b>	<b>Loss</b>	<b>Risk (expected loss) in U.S. dollars</b>	<b>Standard deviation, in U.S. dollars</b>
[GPL, Greatest potential loss; Near, within 300 meters. Mitigation cost is a fraction of structure value]					
1	GPL	5	Total	29,891,764,865	45,835,021,380
2	GPL	5	Partial	5,628,137,250	8,614,477,231
3	Near Calaveras Fault	5	Total	47,813,244,022	73,271,719,155
4	Near Calaveras Fault	5	Partial	6,663,904,959	10,202,508,717
5	Near Hayward Fault	5	Partial	6,277,125,872	9,612,402,149
6	Near Hayward Fault	10	Partial	6,277,125,872	9,612,402,149
7	Near Calaveras Fault	5	Total	48,738,370,116	74,682,875,457
8	GPL, Near Calaveras Fault	5	Partial	6,717,927,420	10,285,065,171
9	GPL, Near Hayward Fault	5	Partial	6,655,219,892	10,189,865,842
10	GPL, Near Hayward Fault	10	Partial	6,655,219,892	10,189,865,842

### **Scenario versus Risk Analysis**

Scenario analysis considers mitigation options that are explored as discrete cases, for example, an examination of the consequences of three floods of differing magnitudes over the same populated area. In attempting to mitigate for a flood of uncertain magnitude over an area of uncertain economic growth, a decision maker might first consider a range of individual scenarios. Aggregating the results of scenarios into probability distributions leads to risk analysis. The LUPM, as currently structured, is based on scenarios, but mathematical research and software development is underway to reformulate the LUPM for risk analysis.

## Conclusions

The LUPM is formulated as an axiomatic model based on Bernoulli random variables and definitions of risk and loss from financial portfolio theory. Algebraic equations for risk (expected loss) and variance of loss have been derived from the axioms, and incorporated into the LUPM software and (Version 1.0). Output from this software has been used to compute scenarios to illustrate risk and benefit for the mitigation of seismic hazard in the San Francisco Bay Area. Experience with software has shown that the quality of data for parameterizing the model and software is a concern.. Additional concerns are how well the model captures the variation of loss and the statistics of rare events. Also the Bernoulli model makes the simplifying assumption that outcomes at individual sites are statistically independent. A more realistic assumption would be to incorporate into the model the geospatial correlation of natural-hazard outcomes among nearby sites. A suggestion for future development is to use simulation to capture variation, the statistics of rare events, and the geospatial correlation of natural hazard outcomes. As currently formulated the Bernoulli model calculates discrete scenarios, but a more useful approach to natural-hazard portfolio planning is use risk analysis which aggregates a range of discrete scenarios. Simulation is also suggested as a future research direction for generalizing the LUPM from scenario to risk analysis.

## References

- Bernknopf, R.L., Dinitiz, L.B., Rabinovici, S.J.M., and Evans, A.M, 2001, A portfolio approach to evaluating natural-hazard mitigation policies: An application to lateral-spread ground failure in coastal California: *International Geology Review*, v. 43, p. 424-440.
- Bernknopf, R.L., Hearn P.P., Wein A.M., and Strong D., 2007, The effect of scientific and socioeconomic uncertainty on a natural-hazards policy choice *in* Oxley, L. and Kulasiri, D. (eds.) MODSIM 2007 International Congress on Modeling and Simulation, Modeling and Simulation Society of Australia and New Zealand, p. 1702-1708.  
[[http://www.mssanz.org.au/MODSIM07/papers/26\\_s32/TheEffectOfScie\\_s32\\_Bernknopf.pdf](http://www.mssanz.org.au/MODSIM07/papers/26_s32/TheEffectOfScie_s32_Bernknopf.pdf) last accessed July 7, 2008].
- Bernknopf, R., Rabinovici, S.J., Wood, N., and Dinitiz, L., 2006, The influence of hazard models on GIS-based regional risk assessments and mitigation policies: *International Journal of Risk Assessment and Management*, v. 6, nos. 4,5,6, p. 369-387.
- Dinitiz, L., Rabinovici, S.J., Hitchcock, C., Lettis, W., and Bernknopf, R., 2003a, Evaluation of natural-hazard mitigation alternatives using a geographic information system and financial portfolio theory: *Proceedings, Disaster Resistant California Conference, San Jose, CA*, p. 1-10.
- Dinitiz, L., Rabinovici, S.J., Kapla, R., Taketa, R., Wood, N., and Bernknopf, R., 2003b, An interactive GIS linking science to natural-hazard mitigation decisions: *Proceedings, Urban and Regional Information Systems Association (URISA) Annual Conference, Atlanta, GA, October, 2003*, p. 789-799.
- Feller, W., 1971, *An introduction to probability theory and its applications* (2nd ed.): New York, John Wiley and Sons, Inc. v. 1.
- Freund, J.E., and Walpole, R.E., 1987, *Mathematical statistics* (4th ed.): Englewood Cliffs, NJ, Prentice-Hall, Inc., 608 p.
- Intriligator, M.D., 1971, *Mathematical optimization and economic theory*: Englewood Cliffs, NJ, Prentice Hall, Inc., 508 p.
- Markowitz, H.M., 1959, *Portfolio selection; Efficient diversification of investments*: New Haven, Conn, Yale University Press, 344 p.
- Ross, S.M., 2002, *Simulation* (3rd ed.): Academic Press, San Diego, CA, An Elsevier Imprint, 274 p.
- Simons, G. and Johnson, N.L., 1971, On the convergence of binomial to poisson distributions: *The Annals of Mathematical Statistics*, v. 42, no. 5; 1735-1736.

Wein, A.M. and Bernknopf, R.L., 2007, Interdisciplinary approaches to regional risk reduction decision-making *in* Oxley, L., and Kulasiri, D. (eds.) MODSIM 2007: International Congress on Modeling and Simulation, Modeling and Simulation Society of Australia and New Zealand, December 2007, p. 1660-1666.  
[[http://www.mssanz.org.au/modsim07/papers/26\\_s32/Interdeciplinary\\_s32\\_Wein\\_.pdf](http://www.mssanz.org.au/modsim07/papers/26_s32/Interdeciplinary_s32_Wein_.pdf), last accessed July 1, 2008].

Wein, A.M., Journey, M. and Bernknopf, R.L., 2007, Scenario-based risk analysis within an analytic-deliberative framework for regional risk reduction planning *in* Oxley, L., and Kulasiri, D. (eds.) MODSIM 2007 International Congress on Modeling and Simulation: Modeling and Simulation Society of Australia and New Zealand, p. 1688-1695.  
[[http://www.mssanz.org.au/modsim07/papers/26\\_s32/Scenario-Based\\_s32\\_Wein\\_.pdf](http://www.mssanz.org.au/modsim07/papers/26_s32/Scenario-Based_s32_Wein_.pdf), last accessed July 7, 2008].

Wolfram, S., 1999, The mathematica book (4th ed.): Cambridge, Cambridge University Press, 1470 p.

## Appendix: Derivations

This appendix provides derivations of algebraic forms that can be given economic interpretations as summary statistics for the LUPM. The sum (27) follows by induction.

$$\left[ \sum_{i=1}^n x_i \right]^2 = \sum_{i=1}^n x_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^n (x_i * x_j) \quad (27)$$

It will be useful to have algebraic simplifications for polynomials in Bernoulli variables. For distinct and independent Bernoulli random variables  $B_p$  and  $B_q$ ,  $B_p * B_q$  has the same distribution as  $B_{p*q}$ . To see why, observe table 5 of outcomes and probabilities. The matrix on the left suggests a Bernoulli random variable. Collecting terms in the matrix on the right shows that:

$$\Pr(B_p * B_q = 0) = 1 - p * q$$

and

$$\Pr(B_p * B_q = 1) = p * q.$$

**Table 5.** Product of distinct Bernoulli random variables.

		<b>Bq</b>	
		0	1
<b>Bp</b>	0	0	0
	1	0	1

	<b>Probabilities</b>	
	1-q	q
1-p	(1-p)*(1-q)	(1-p)*q
p	p*(1-q)	p*q

Similarly, the square of a Bernoulli random variable  $B^2(1, p)$  has the same distribution as  $B(1, p)$ .

Table 6 shows all combinations of hypothetical outcomes. Only two outcomes (0 and 1) are actually possible, and there are only two ways of getting these outcomes. The X's indicate

outcomes that cannot occur because these would imply that the random variable would simultaneously take on the values of 0 and 1. The matrix on the right gives the probabilities of obtaining the two outcomes that are actually possible.

**Table 6.** The square of a Bernoulli random variable.

		<b>Bp</b>	
		<b>0</b>	<b>1</b>
<b>Bp</b>	<b>0</b>	<b>0</b>	<b>X</b>
	<b>1</b>	<b>X</b>	<b>1</b>

			<b>Probabilities</b>	
			<b>1-p</b>	<b>p</b>
<b>1-p</b>	<b>1-p</b>	<b>X</b>		
<b>p</b>	<b>X</b>	<b>p</b>		

The LUPM models hazards as Bernoulli variables in two stages. The first stage (28) indicates the occurrence, and the second (29) indicates consequences.

$$D = B(1, P_H) \tag{28}$$

$$S = \sum_{i=1}^n z_i * B(1, p_i) \tag{29}$$

In (29) the Bernoulli variables are independent, and the  $z_i$  initially are undefined. By using the asterisk to indicate one random outcome followed by another, the combination of the hazard event and its consequences can be written:

$$\Gamma = D * S. \tag{30}$$

Equation (30) is interpreted as saying that the hazard event has consequences only when the event actually occurs. The following computations give  $E(\Gamma)$  and  $Var(\Gamma)$ . Using (31)

$$Var(\Gamma) = E(\Gamma^2) - [E(\Gamma)]^2 \tag{31}$$

$$\begin{aligned}
E(\Gamma) &= \Pr(D=1) * E(S) \\
&= P_H * \left[ \sum_{i=1}^n p_i * z_i \right]
\end{aligned} \tag{32}$$

The second expectation on the right is:

$$E(\Gamma^2) = \Pr(D=1) * E(S^2), \tag{33}$$

$$\begin{aligned}
S^2 &= \left[ \sum_{i=1}^n z_i * B(1, p_i) \right]^2 \\
&= \left[ \sum_{i=1}^n z_i^2 * B(1, p_i) \right] + \left[ \sum_{\substack{i,j=1 \\ i \neq j}}^n z_i * z_j * B(1, p_i * p_j) \right]
\end{aligned} \tag{34}$$

Thus (33) becomes:

$$E(\Gamma^2) = P_H * \left[ \sum_{i=1}^n z_i^2 * p_i \right] + P_H * \left[ \sum_{\substack{i,j=1 \\ i \neq j}}^n z_i * z_j * p_i * p_j \right]. \tag{35}$$

Letting  $q_i = P_H * p_i$ :

$$Var(\Gamma) = \left[ \sum_{i=1}^n q_i * (1 - q_i) * z_i^2 \right] + P_H * (1 - P_H) * \left[ \sum_{\substack{i,j \\ i \neq j}}^n p_i * p_j * z_i * z_j \right]. \tag{36}$$

Letting the  $p_i$  in (36) be the conditional failure probabilities from (5), and setting the  $z_i$  to correspond to the appropriate combination of  $\phi_i$ ,  $\bar{\phi}_i$ ,  $\Delta\phi_i$ , or  $\bar{V}_i$  gives the conditional expectations (19), (21) and (23). Letting  $P_H = 1$  gives the unconditional expectations (20), (22), and (24).