Probabilistic Methodology for Estimation of Number and Economic Loss (Cost) of Future Landslides in the San Francisco Bay Region, California

Robert A. Crovelli[1] and Jeffrey A. Coe[2]

[1] Retired, U.S. Geological Survey, Denver Federal Center, MS 939, Denver, CO 80225
[2] U.S. Geological Survey, Denver Federal Center, MS 966, Denver, CO 80225

Open-File Report 2008-1116

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:

Crovelli, R.A. and Coe, J.A., 2008, Probabilistic methodology for estimation of number and economic loss (cost) of future landslides in the San Francisco Bay Region, California: U.S. Geological Survey Open-File Report 2008-1116, 27 p.

Probabilistic Methodology for Estimation of Number and Economic Loss (Cost) of Future Landslides in the San Francisco Bay Region, California

Robert A. Crovelli and Jeffrey A. Coe

ABSTRACT

The Probabilistic Landslide Assessment Cost Estimation System (PLACES) presented in this report estimates the number and economic loss (cost) of landslides during a specified future time in individual areas, and then calculates the sum of those estimates. The analytic probabilistic methodology is based upon conditional probability theory and laws of expectation and variance. The probabilistic methodology is expressed in the form of a Microsoft Excel computer spreadsheet program. Using historical records, the PLACES spreadsheet is used to estimate the number of future damaging landslides and total damage, as economic loss, from future landslides caused by rainstorms in 10 counties of the San Francisco Bay region in California. Estimates are made for any future 5-year period of time.

The estimated total number of future damaging landslides for the entire 10-county region during any future 5-year period of time is about 330. Santa Cruz County has the highest estimated number of damaging landslides (about 90), whereas Napa, San Francisco, and Solano Counties have the lowest estimated number of damaging landslides (5–6 each). Estimated direct costs from future damaging landslides for the entire 10-county region for any future 5-year period are about US $76 million (year 2000 dollars). San Mateo County has the highest estimated costs ($16.62 million), and Solano County has the lowest estimated costs (about $0.90 million). Estimated direct costs are also subdivided into public and private costs.

KEY WORDS: PLACES**,** probability, spreadsheet, landslide, cost, loss, risk, San Francisco Bay, California.

INTRODUCTION

Landslides occur nearly every year in the San Francisco Bay region of California. Most landslides occur during the late fall through early spring seasons, typically between December and April. During the fall through spring seasons of 1968–69, 1972–73, 1981–82, and 1997–98, landslides were widespread and caused extensive damage both to public and private property. Following these years, the U.S. Geological Survey (USGS) mapped locations of landslides that caused damage, and compiled the direct costs of damage to public and private property (Taylor and Brabb, 1972; Taylor and others, 1975; Creasy, 1988; Godt and others, 1999). The mapping and compilation were done for 10 counties in the region: Alameda, Contra Costa, Marin, Napa, San Francisco, Santa Clara, Santa Cruz, San Mateo, Solano, and Sonoma. Total numbers of, and costs from, damaging landslides in each of these counties are listed in table 1.

The length of the historical record used in this report (referred to as "past time" in table 1) is 39 years (1968/69–2006/07) for all counties except Santa Cruz, which is 35 years because data were not collected in Santa Cruz County in 1968–69 and 1972–73. Even though 39 and 35 years were used as the length of record, during this timeframe, we only had data available for the fall through spring seasons of 1968–69, 1972–73, 1981–82, and 1997–98. Because of this fact, the historical record used in this report is incomplete and all estimates of future landslide numbers and costs must be considered minimum estimates. This statement is true for several reasons, including the following: (1) some years between 1968 and present (September 2007) have had landslides that caused damage (for examples, see Brown, 1988) that were not recorded by the USGS, (2) there were undoubtedly some landslides that caused damage during the years when records were kept (that is, 1968–69, 1972–73, 1981–82, and 1997–98) that were missed by the various USGS compilers, and (3) historical records of costs from landslides triggered by earthquakes were not included in the study. Additional limitations of our analysis are that (1) we do not take into account any future increases or decreases in precipitation due to changing climatic conditions; we assume that precipitation conditions in the future will be similar to those reflected by the historical record, and (2) we do not explicitly account for future patterns of growth in public and private development that may affect future numbers and costs of damaging landslides.

To analyze the historical cost data, we have used a newly developed Probabilistic Landslide Assessment Cost Estimation System (PLACES) to estimate the mean (or expected) number of future damaging landslides and the mean economic losses from the landslides. Along with mean estimates, PLACES calculates, for any specified future time, prediction interval (low, high) estimates at any specified prediction probability level (percent) and exceedance probabilities at any specified loss exceedance level (dollars).

PLACES significantly expands on probability methods for landslide data that were previously described by Crovelli (2000). An application of the methods described by Crovelli (2000) using historical landslide data from Seattle was described by Coe and others (2004). PLACES expands on these previous studies primarily through the addition of methods to partition and aggregate landslide costs. New features include the concept of landslide clusters and landslides per cluster, costs of damage to public and private property, aggregation of totals under various degrees of correlation, and the inclusion of the complete historical data set from the San Francisco Bay region. The historical record of landslide costs in the San Francisco Bay region, which is unique because of the internal consistency of the data and the longevity of the compilation effort, serves as an ideal data set for an application of PLACES.

PLACES uses probabilistic methodology for analysis of a particular set of landslide random variables. A random variable is a variable that has a probability distribution, along with a mean and a standard deviation. The PLACES probabilistic methodology involves the following random variables and their relationships, which form an outline of the probabilistic methodology section of this report:

(1) Number of landslide clusters

(2) Recurrence interval of landslide clusters

(3) Number of landslides per landslide cluster

(4) Cost of landslides per landslide cluster

(5) Total number of landslides —
(5) is a function of (1) and (3).

(6) Total cost of landslides —
(6) is a function of (1) and (4).

(7) Fraction or percentage/100 (public and private)

(8) Fraction of total cost of landslides (public and private) —
(8) is a function of (6) and (7).

(9) Aggregation of total numbers of landslides —
(9) is a function of (5).

(10) Aggregation of total costs of landslides —
(10) is individually a function of (6) and then (8).

Application and discussion sections follow the detailed description of the probabilistic methodology.

PROBABILISTIC METHODOLOGY

PLACES was designed from probabilistic methodology to calculate estimates of the number and economic loss (cost) of landslides during a specified future time in individual areas, and then calculate the sum of those estimates. The analytic probabilistic methodology was developed by deriving the necessary mathematical equations based upon conditional probability theory and laws of expectation and variance. The derivations of the necessary equations are given in the following sections.

Number of Landslide Clusters

Landslide cluster: A group of one or more landslides that occurs within an individual water year.

Water year: The year-long period between July 1 and June 30 of the following year.

Discrete-time probability model for occurrence of landslide clusters: Binomial process where there is a series of water years and within each water year a landslide cluster may or may not occur.

Random variable $N(t)$: Number of landslide clusters that occur during a time period of t water years in a particular area.

Range of $N(t)$: $\{0, 1, \ldots, t\}$

Assumptions: There are t independent water years. Within each water year a landslide cluster may or may not occur. The probability of a landslide cluster in a water year, denoted by p, remains constant from water year to water year.

Probability distribution of $N(t)$: Binomial distribution with parameters t and p.

Parameter t: Specified number of water years.

Parameter p: Probability of a landslide cluster in a water year.

Probability mass function: $P\{N(t) = n\} = C(t,n)\, p^n\, (1 - p)^{t-n}$

Mean or expected value of $N(t)$: $E[N(t)] = tp$

Standard deviation of $N(t)$: $S[N(t)] = [tp(1 - p)]^{1/2}$

Exceedance probability: Probability of one or more clusters during a time period of t water years.

Exceedance probability: $P\{N(t) \geq 1\} = 1 - (1 - p)^t$

Estimator of parameter p: $\underline{P} = N(t^*)/t^*$ where t^* denotes observed fixed time.

Recurrence Interval of Landslide Clusters

Random variable R: Recurrence interval is the number of water years from one landslide cluster until the next cluster.

Range of R: $\{1, 2, \ldots\}$

Assumptions: There is a series of independent water years after a landslide cluster occurs until the next cluster occurs. Within each water year a landslide cluster may or may not occur. The probability of a landslide cluster in a water year, denoted by p, remains constant from water year to water year.

Probability distribution of R: Geometric distribution with parameter p.

Parameter p: Probability of a landslide cluster in a water year.

Probability mass function: $P\{R = r\} = p(1 - p)^{r\text{-}1}$

Mean recurrence interval is the average time between landslide clusters.

Mean or expected value of R: $E[R] = 1/p$

Standard deviation of R: $S[R] = [(1 - p)/p^2]^{1/2}$

Exceedance probability: Probability of a recurrence interval being greater than r water years.

Exceedance probability: $P\{R > r\} = (1 - p)^r$

Estimator of parameter p: $\underline{P} = N(t^*)/t^*$ where t^* denotes observed fixed time.

Number of Landslides per Landslide Cluster

Random variable L: Number of landslides per landslide cluster.

Range of L: $\{1, 2, \ldots\}$

Mean or expected value of L: $E[L]$

Standard deviation of L: $S[L]$

Estimator of $E[L]$: Sample mean M_L, based on n observed landslide clusters.

$$M_L = \frac{\sum_{i=1}^{n} L_i}{n}$$

Estimator of $S[L]$: Sample standard deviation S_L, based on n observed landslide clusters.

$$S_L^2 = \frac{n\sum_{i=1}^{n} L_i^2 - \left(\sum_{i=1}^{n} L_i\right)^2}{n(n-1)}$$

Cost of Landslides per Landslide Cluster

Random variable X: Cost of landslides per landslide cluster.

Range of X: $(0, \infty)$

Mean or expected value of X: $E[X]$

Standard deviation of X: $S[X]$

Estimator of $E[X]$: Sample mean M_X, based on n observed landslide clusters.

$$M_X = \frac{\sum_{i=1}^{n} X_i}{n}$$

Estimator of $S[X]$: Sample standard deviation S_X, based on n observed landslide clusters.

$$S_X^2 = \frac{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}{n(n-1)}$$

Total Number of Landslides

Random variable $M(t)$: Total number of landslides from all of the landslide clusters during a time period of t water years in a particular area.

$$M(t) = \sum_{i=1}^{N(t)} L_i$$

where random variable L_i: Number of landslides from the ith landslide cluster.

Range of $M(t)$: $\{n, n+1, \ldots\}$

Assumptions: The L_i ($i = 1, 2, \ldots$) are independent and identically distributed random variables which are also independent of $N(t)$.

The random variable $M(t)$ is equal to the sum of a *random* number $N(t)$ of random variables L_i. The mean and standard deviation of $M(t)$ can be derived from the theory of conditional probability and conditional expectation (Ross, 2000).

The derivation of the formula for the mean of $M(t)$ is given in Ross (2000, p. 103–104).

Mean or expected value of $M(t)$: $E[M(t)] = E[N(t)]E[L]$

The derivation of the formula for the standard deviation of $M(t)$ is given in Ross (2000, p. 111–112).

Standard deviation of $M(t)$: $S[M(t)] = \{E[N(t)](S[L])^2 + (E[L])^2(S[N(t)])^2\}^{1/2}$

Total Cost of Landslides

Random variable $Y(t)$: Total cost of landslides from all of the landslide clusters during a time period of t water years in a particular area.

$$Y(t) = \sum_{i=1}^{N(t)} X_i$$

where random variable X_i: Cost of landslides from the ith landslide cluster.

Range of $Y(t)$: $(0, \infty)$

Assumptions: The X_i ($i = 1, 2, \ldots$) are independent and identically distributed random variables that are also independent of $N(t)$.

The random variable $Y(t)$ is equal to the sum of a *random* number $N(t)$ of random variables X_i. The mean and standard deviation of $Y(t)$ can be derived from the theory of conditional probability and conditional expectation (Ross, 2000).

The derivation of the formula for the mean of $Y(t)$ is given in Ross (2000, p. 103–104).

Mean or expected value of $Y(t)$: $\mu_Y = E[Y(t)] = E[N(t)]E[X]$

The derivation of the formula for the standard deviation of $Y(t)$ is given in Ross (2000, p. 111–112).

Standard deviation of $Y(t)$: $\sigma_Y = S[Y(t)] = \{E[N(t)](S[X])^2 + (E[X])^2(S[N(t)])^2\}^{1/2}$

Probability Distribution for Total Cost of Landslides

Crovelli (1992) showed that the lognormal probability distribution is a good approximate distribution for the type of random variable $Y(t)$. Hence, the fractiles (fractiles are the complement of percentiles) of $Y(t)$ can be approximated by using the lognormal distribution.

$Y(t)$ is a sum of positive random variables and, therefore, is also a positive random variable. It is well known that sums of random variables tend to have a bell-shaped distribution and, by the Central Limit Theorem, approach the normal distribution. The lognormal distribution is a positive bell-shaped distribution. We also have the following statement from the updated classical reference book on continuous probability distributions (Johnson and others, 1994, p. 239): “The two-parameter [lognormal] distribution is, in at least one important respect, a more realistic representation of distributions of characters like weight, height, and density than is the normal distribution. These quantities cannot take negative values, but a normal distribution ascribes positive probability to such events, while the two-parameter lognormal distribution does not. Furthermore, by taking σ small enough, it is possible to construct a lognormal distribution closely resembling any normal distribution. Hence, even if a normal distribution is felt to be really appropriate, it might be replaced by a suitable lognormal distribution.”

As derived in Crovelli (1992), the characterizing parameters of the lognormal distribution, namely μ and σ, can be calculated from the mean μ_Y and standard deviation σ_Y of a lognormal random variable Y.

The mean μ_Y and standard deviation σ_Y of a lognormal random variable Y with characterizing parameters μ and σ are the following well-known formulas (Johnson and others, 1994, p. 212).

$$\mu_Y = e^{\mu+\sigma^2/2}$$

$$\sigma_Y^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

Solving the two equations for the lognormal characterizing parameters μ and σ, we get

$$\mu = \ln\left(\frac{\mu_Y^2}{\sqrt{\mu_Y^2 + \sigma_Y^2}}\right)$$

$$\sigma = \sqrt{\ln(\sigma_Y^2 / \mu_Y^2 + 1)}$$

The details of this derivation are given as Theorem 1 in the Appendix.

If one knows the lognormal characterizing parameters, the lognormal fractiles can be calculated from the formula

$$F100\,\alpha = e^{\mu + z_\alpha \sigma} \qquad 0 \le \alpha \le 1$$

where Z is a standard normal random variable and $P\{Z > z_\alpha\} = \alpha$.

For example, two fractiles of interest in this report are

$$F95 = e^{\mu - 1.645\sigma} \quad and \quad F5 = e^{\mu + 1.645\sigma}$$

There is a 95% chance of exceeding *F*95, and a 5% chance of exceeding *F*5. Together, the low value of *F*95 and the high value of *F*5 form a range of values that is a 90% prediction interval for *Y*(*t*), the total costs from landslides during a specified time (at a 90% prediction level).

The reverse problem would be to find the probability of exceeding a specified amount in economic loss due to landslides in a particular area during a specified time. That is, given y_α, find α such that

$P\{Y(t) > y_\alpha\} = \alpha$

Normalizing

$$z_\alpha = \frac{\ln y_\alpha - \mu}{\sigma}$$

Now, from z_α, find α such that $P\{Z > z_\alpha\} = \alpha$.

This probabilistic methodology would also apply in the case of the probability distribution for total number of landslides.

Fraction of Total Cost of Landslides

Public and private costs represent fractions of total cost of landslides.

Random variable *Z*(*t*): Fraction of total cost of landslides during a time period of *t* water years in a particular area.

$Z(t) = F * Y(t)$

where random variable *F*: Fraction or percentage/100.

The random variable *Z*(*t*) is equal to the product of a random fraction *F* and the random variable *Y*(*t*).

Range of *F*: (0, 1)

Range of *Z*(*t*): $(0, \infty)$

Assumption: *F* and *Y*(*t*) are assumed to be independent.

Mean or expected value of *F*: $E[F]$

Standard deviation of F: $S[F]$

Estimator of $E[F]$: Weighted mean M_F, based on n observed landslide clusters with fractions F_i and (weights) costs per cluster X_i ($i = 1, 2, \ldots, n$).

$$M_F = \frac{\sum_{i=1}^{n} F_i X_i}{\sum_{i=1}^{n} X_i}$$

Estimator of $S[F]$: Weighted standard deviation S_F, based on n observed landslide clusters.

$$S_F^2 = \frac{\sum_{i=1}^{n} F_i^2 X_i}{\sum_{i=1}^{n} X_i} - M_F^2$$

The derivation of the following formulas for the mean and standard deviation of $Z(t)$ are given as Theorem 2 in the Appendix.

Mean or expected value of $Z(t)$: $E[Z(t)] = E[F]E[Y(t)]$

Standard deviation of $Z(t)$:

$$S[Z(t)] = \{(S[F])^2(S[Y(t)])^2 + (E[Y(t)])^2(S[F])^2 + (E[F])^2(S[Y(t)])^2\}^{1/2}$$

Aggregation of Total Costs of Landslides

Random variable $W(t)$: Aggregation of total costs of landslides during a time period of t water years in k areas.

$$W(t) = \sum_{i=1}^{k} Y_i(t)$$

where random variable $Y_i(t) = Y_i$: Total cost of landslides in the ith area ($i = 1, 2, \ldots, k$).

Range of $W(t)$: $(0, \infty)$

The random variable $W(t)$ is equal to the sum of a *fixed* number k of random variables $Y_i(t)$.

Mean or expected value of $W(t)$:

$$E[W(t)] = \sum_{i=1}^{k} E[Y_i(t)]$$

Variance of $W(t)$:

$$V[W(t)] = \sum_{i=1}^{k} V[Y_i] + 2\sum_{i<j} Cov(Y_i, Y_j)$$

where covariance of Y_i and Y_j:

$$Cov(Y_i, Y_j) = E[(Y_i - E[Y_i])(Y_j - E[Y_j])]$$

Correlation coefficient of Y_i and Y_j:

$$\rho_{ij} = \frac{Cov(Y_i, Y_j)}{\sigma_i \sigma_j}$$

where $\sigma_i = S[Y_i]$: Standard deviation of Y_i

Number of distinct correlation coefficients ($i<j$):

$m = k(k - 1)/2$; for example, $k = 10$, then $m = 45$.

Variance of $W(t)$:

$$V[W(t)] = \sum_{i=1}^{k} \sigma_i^2 + 2\sum_{i<j} \rho_{ij}\sigma_i\sigma_j$$

Weighted-average correlation coefficient:

$$\rho_{wa} = \frac{\sum_{i<j} \rho_{ij}\sigma_i\sigma_j}{\sum_{i<j} \sigma_i\sigma_j}$$

If $\sigma_i = \sigma$ for all $i = 1, 2, \ldots, k$, then we get the average correlation coefficient:

$$\rho_{wa} = \frac{\sum_{i<j} \rho_{ij}}{k(k-1)/2}$$

Variance of $W(t)$:

$$V[W(t)] = \sum_{i=1}^{k} \sigma_i^2 + 2\rho_{wa} \sum_{i<j} \sigma_i \sigma_j$$

An algebraic relationship:

$$\left(\sum_{i=1}^{k} \sigma_i \right)^2 = \sum_{i=1}^{k} \sigma_i^2 + 2\sum_{i<j} \sigma_i \sigma_j$$

The final general case of variance of *W*(*t*):

$$V[W(t)] = \sum_{i=1}^{k} \sigma_i^2 + \rho_{wa} \left[\left(\sum_{i=1}^{k} \sigma_i \right)^2 - \sum_{i=1}^{k} \sigma_i^2 \right]$$

Note that: $-1 \le \rho_{wa} \le 1$

The special cases are (a) uncorrelation or independence, (b) perfect positive correlation, and (c) perfect negative correlation.

(a) $\rho_{wa} = 0 \Rightarrow$ uncorrelation or independence

(b) $\rho_{wa} = 1 \Rightarrow$ perfect positive correlation

(c) $\rho_{wa} = -1 \Rightarrow$ perfect negative correlation

$$(a) \quad V[W(t)] = \sum_{i=1}^{k} \sigma_i^2$$

$$(b) \quad V[W(t)] = \sum_{i=1}^{k} \sigma_i^2 + \left[\left(\sum_{i=1}^{k} \sigma_i \right)^2 - \sum_{i=1}^{k} \sigma_i^2 \right] = \left(\sum_{i=1}^{k} \sigma_i \right)^2$$

$$(c) \quad V[W(t)] = \sum_{i=1}^{k} \sigma_i^2 - \left[\left(\sum_{i=1}^{k} \sigma_i \right)^2 - \sum_{i=1}^{k} \sigma_i^2 \right] = 2\sum_{i=1}^{k} \sigma_i^2 - \left(\sum_{i=1}^{k} \sigma_i \right)^2$$

Standard deviation of *W*(*t*):

$$S[W(t)] = \{V[W(t)]\}^{1/2}$$

If $\sigma_i = \sigma$ for all $i = 1, 2, \ldots, k$, then

$$(a)\quad V[W(t)] = \sum_{i=1}^{k} \sigma_i^2 = \sum_{i=1}^{k} \sigma^2 = k\sigma^2$$

$$S[W(t)] = \sqrt{k}\sigma$$

$$(b)\quad V[W(t)] = \left(\sum_{i=1}^{k} \sigma_i \right)^2 = \left(\sum_{i=1}^{k} \sigma \right)^2 = (k\sigma)^2 = k^2\sigma^2$$

$$S[W(t)] = k\sigma$$

$$Ratios:\quad R_V = \frac{k^2\sigma^2}{k\sigma^2} = k \quad \& \quad R_S = \frac{k\sigma}{\sqrt{k}\sigma} = \sqrt{k}$$

$$For\ k = 25,\quad R_V = 25 \quad \& \quad R_S = 5$$

$$For\ k = 100,\quad R_V = 100 \quad \& \quad R_S = 10$$

For k equal to 25 and 100, respectively, the variance of a sum of random variables with equal variances in the case of perfect positive correlation is 25 and 100 times greater than the variance of a sum of random variables with equal variances in the case of uncorrelation or independence.

For k equal to 25 and 100, respectively, the standard deviation of a sum of random variables with equal standard deviations in the case of perfect positive correlation is 5 and 10 times greater than the standard deviation of a sum of random variables with equal standard deviations in the case of uncorrelation or independence.

Remarks concerning ρ_{wa}:

1) Estimation of ρ_{wa} directly can save considerable effort compared to the estimation of the m individual ρ_{ij}, especially when m is large.

2) Interpretation of ρ_{wa} as a measure of the degree of weighted-average correlation of all of the random variables is helpful, especially when the value of ρ_{wa} can be thought of as a decimal fraction (or percentage) lying somewhere between independence (0) and perfect positive correlation (1) inclusively.

3) When $\rho_{wa} = 1$, the $V[W(t)]$ is larger than in the case when $\rho_{wa} = 0$. On the other hand, when $\rho_{wa} = -1$, the $V[W(t)]$ is smaller than in the case when $\rho_{wa} = 0$.

Also, we have:

In the case of perfect positive correlation, the means, standard deviations, and fractiles are additive.

This aggregation method is also used in the aggregation of total numbers of landslides where $M(t)$ would replace $Y(t)$.

SPREADSHEET AND APPLICATION

The probabilistic methodology presented herein was used to construct a Microsoft Excel computer spreadsheet program. Using historical records, the PLACES spreadsheet was applied to estimate the number of future landslides, and total damage, as economic loss (cost), from future landslides caused by rainstorms in 10 counties of the San Francisco Bay region in California. Tables 1 and 2 contain summaries of historical cost data. These historical data were used in PLACES to estimate mean recurrence intervals and probabilities of future damaging landslide clusters (table 3); numbers of future landslides in a specified amount of time with confidence intervals and exceedance probabilities (table 4); and estimates of future total costs, and their aggregation, with prediction intervals and exceedance probabilities (table 5). The future specified time of 5 years used in the spreadsheet is a reasonable value selected for purely illustrative purposes. Table 6 contains historical cost data subdivided into public and private costs. PLACES estimates of public costs due to future landslides, and their aggregation, with prediction intervals and exceedance probabilities are given in table 7, whereas estimates of private costs due to future landslides are given in table 8. Recall that the historical record of landslides in the San Francisco Bay region is incomplete; therefore, all estimates of future numbers and costs (tables 3–5, 7–8) must be considered minimum estimates.

DISCUSSION

In the spreadsheet as it exists, the estimates of future total numbers and total costs of landslides in the San Francisco Bay region are based upon historical records from the region. When estimates of the future are based upon historical records, this could be referred to as a "historical" scenario. An assumption of the historical scenario is that the future will be similar to the past. In general, this assumption may or may not be acceptable in various applications. The problem is that everything in the physical world is continuously changing — nothing stays the same. Our information about the physical world is also continuously changing. If we have information that suggests the future will be different from the past in a certain "direction," for example, more landslides in the future than in the past, then we might want to modify the historical values of key parameters. That is, we might want to attempt some "what-if" scenarios by changing the historical values of certain key parameters in the PLACES model; for example, the probability of a landslide cluster might be increased. One of the most useful features of the PLACES spreadsheet is that we can modify the historical value of a key parameter and instantly see the effect on future estimates.

The probabilistic methodology and computer spreadsheet could be modified easily to become applicable to other types of hazards and even other types of disciplines. Therefore, there could be modifications of the PLACES system itself, besides modifications of values of parameters within the system. Two cases of modification of two different aspects of the system itself are given below.

Case 1: Continuous-Time Model

It would be a very simple procedure to modify the PLACES system itself in the case where $N(t)$ has a Poisson distribution (continuous-time probability model) instead of a binomial distribution (discrete-time probability model), which would require that R have an exponential distribution instead of a geometric distribution while the rest of the system would remain exactly the same.

Number of Point Events

Point event: An event that occurs at some time point in continuous time, where time is not discretized into one-year increments, as in the case of water years; for example, an earthquake.

Continuous-time probability model for occurrence of point events: Poisson process where there is a series of random point events in continuous time.

Random variable $N(t)$: Number of point events that occur during time t in a particular area.

Range of $N(t)$: $\{0, 1, 2, \ldots\}$

Assumptions: The process has independent increments; that is, the numbers of point events that occur in disjoint time intervals are independent. The process has stationary increments; that is, the distribution of the number of point events that occur in any interval of time depends only on the length of the time interval.

Probability distribution of $N(t)$: Poisson distribution with parameters t and λ.

Parameter t: Specified time interval.

Parameter λ: Rate of occurrence of point events.

Probability mass function: $P\{N(t) = n\} = \mathrm{e}^{-\lambda t} (\lambda t)^n / n!$

Mean or expected value of $N(t)$: $E[N(t)] = \lambda t$

Standard deviation of $N(t)$: $S[N(t)] = (\lambda t)^{1/2}$

Exceedance probability: Probability of one or more point events during time t.

Exceedance probability: $P\{N(t) \geq 1\} = 1 - \mathrm{e}^{-\lambda t}$

Estimator of parameter λ: $\underline{\lambda} = N(t^*)/t^*$ where t^* denotes observed fixed time.

Recurrence Interval of Point Events

Random variable R: Recurrence interval is the elapsed time between point events.

Range of R: $(0, \infty)$

Assumption: Probability model for occurrence of the point events is the Poisson process.

Probability distribution of R: Exponential distribution with parameter λ.

Parameter λ: Rate of occurrence of point events.

Probability density function: $f(r) = \lambda t e^{-\lambda tr}$

Mean recurrence interval is the average time between point events.

Mean or expected value of R: $E[R] = 1/\lambda$

Standard deviation of R: $S[R] = 1/\lambda$

Exceedance probability: Probability of a recurrence interval being greater than time r.

Exceedance probability: $P\{R > r\} = e^{-\lambda r}$

Estimator of parameter λ: $\underline{\lambda} = N(t^*)/t^*$ where t^* denotes observed fixed time.

Case 2: Cluster-of-One Model

In the degenerate case where the cluster always consists of only one single event so that the "cluster" would now become the "event," we have

Random variable $L = 1$

Mean or expected value of L: $E[L] = 1$

Standard deviation of L: $S[L] = 0$

Random variable $M(t) = N(t)$

Mean or expected value of $M(t)$: $E[M(t)] = E[N(t)]$

Standard deviation of $M(t)$: $S[M(t)] = S[N(t)]$

The rest of the system would remain exactly the same.

This case would be appropriate for the following situations:

1. Modeling occurrence of seasonal floods (a series of single floods) by the binomial process (discrete-time probability model).
2. Modeling occurrence of earthquakes (a series of single earthquakes) by the Poisson process (continuous-time probability model).

Reminder: Philosophy of Probability Models

We include this section as a reminder for readers who apply the methodology (or use the results) presented in this report. It is very important to distinguish between hazardous processes themselves and our descriptions or models of these processes. At the scale of geologic and atmospheric hazards (for example, landslides, earthquakes, floods, tsunamis, volcanoes, and storms), nature is deterministic: every hazardous event has a cause. We cannot predict exactly when a hazardous event will occur because of the limitations to our knowledge of nature. A probability model is a mathematical model that incorporates *our* uncertainty. Probability models are used for purposes of description and prediction of physical processes in nature. Randomness is an assumption of probability models, not an inherent quality of natural processes. Hazards do not occur at random in nature, but they do occur at random in the models. In summary, hazardous processes are deterministic; but, because of our limitations when studying hazards, we resort to probability models that incorporate our uncertainty.

SUMMARY

This report presents probabilistic methodology (PLACES) that can be used to assess numbers and costs of future landslides based on historical data. The probabilistic methodology is expressed in the form of a Microsoft Excel computer spreadsheet program. Useful features of the methodology and spreadsheet are the following: (1) aggregation of totals under various degrees of correlation, (2) flexibility to modify parameter values within the system using the PLACES spreadsheet and instantly see the effect on future estimates, and (3) flexibility to modify the system itself using alternative probability models — two important cases are included in the report.

The PLACES spreadsheet is used to estimate the number of future damaging landslides, and total damage, as economic loss, from future landslides caused by rainstorms in 10 counties of the San Francisco Bay region in California. Estimates are made for any future 5-year period of time. The estimated total number of future damaging landslides for the entire 10 county region during any given 5-year period of time is about 330. Santa Cruz County has the highest estimated number of damaging landslides (about 90), whereas Napa, San Francisco, and Solano Counties have the lowest estimated number of damaging landslides (5–6 each). Estimated direct costs from future damaging landslides for the entire 10-county region for any future 5-year period are about US $76 million (year 2000 dollars). San Mateo County has the highest estimated costs ($16.62 million), whereas Solano County has the lowest estimated costs (about $0.90 million). Estimated direct costs are also subdivided into public and private costs.

ACKNOWLEDGMENTS

The authors wish to thank reviewers Dave Perkins and Brent Troutman for their helpful comments that improved the description of the probabilistic methodology.

REFERENCES

Brown, W.M., III, 1988, Historical setting of the storm: Perspectives on population, development, and damaging rainstorms in the San Francisco Bay region, *in* Ellen, S.D., and Wieczorek, G.F., eds., Landslides floods, and marine effects of the storm of January 3–5, 1982 in the San Francisco Bay region, California: U.S. Geological Survey Professional Paper 1434, p. 7–15.

Coe, J.A., Michael, J.A., Crovelli, R.A., and Savage, W.Z., Laprade, W.T., and Nashem, W.D., 2004, Probabilistic assessment of precipitation-triggered landslides using historical records of landslide occurrence, Seattle, Washington: Environmental and Engineering Geoscience, v. 10, p. 103–122.

Crovelli, R.A., 1992, Probabilistic methodology for estimation of undiscovered petroleum resources in play analysis of the United States: Nonrenewable Resources, v. 1, no. 2, p. 153–162.

Crovelli, R.A., 2000, Probability models for estimation of number and costs of landslides: U.S. Geological Survey Open-File Report 2000–249, 23 p., accessed March 7, 2008 at http://pubs.usgs.gov/of/2000/ofr-00-0249/

Godt, J.W. (editor), Arnal, C.H., Baum R.L., Brien, D.L., Coe, J.A., De Mouthe, J., Ellis, W.L., Graymer, R.W., Harp, E.L., Hillhouse, J.W., Houdre, N., Howell, D.G., Jayko, A.S., Lajoie, K.R., Morrissey, M.M., Ramsey, D.W., Savage, W.Z., Schuster, R.L, Wieczorek, G.F., and Wilson, R.C., 1999, Maps showing locations of damaging landslides caused by El Niño rainstorms, winter season 1997–98, San Francisco Bay region, California: U.S. Geological Survey Miscellaneous Field Studies Maps MF–2325–A–J.

Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994, Continuous univariate distributions: New York, John Wiley & Sons, Inc., 2nd ed., 756 p.

Creasy, C.L., 1988, Landslide damage: A costly outcome of the storm, in Ellen, S.D., and Wieczorek, G.F., eds., Landslides, floods, and marine effects of the storm of January 3–5, 1982, in the San Francisco Bay region, California: U.S. Geological Survey Professional Paper 1434, p. 195–203.

Ross, S.M., 2000, Introduction to probability models: San Diego, Harcourt, Academic Press, Inc., 7th ed., 693 p.

Taylor, F.A., and Brabb, E.E., 1972, Maps showing distribution and cost by counties of structurally damaging landslides in the San Francisco Bay region, California, winter of 1968–69: U.S. Geological Survey Miscellaneous Field Studies Map MF–327, scales 1:500,000 and 1:1,000,000.

Taylor, F.A., Nilsen, T.H., and Dean, R.M., 1975, Distribution and cost of landslides that have damaged manmade structures during the rainy season of 1972–1973 in the San Francisco Bay region, California: U.S. Geological Survey Miscellaneous Field Studies Map MF–679, scales 1:500,000 and 1:1,000,000.

U.S. Department of Labor, 1997, Bureau of Labor Statistics Handbook of Methods, variously paged.

APPENDIX

Calculation of the characterizing parameters of a lognormal distribution from its mean and standard deviation.

Theorem 1

If Y has a lognormal distribution with characterizing parameters μ and σ, mean μ_Y, and standard deviation σ_Y, then μ and σ can be calculated from μ_Y and σ_Y by the following formulas:

$$\mu = \ln\left(\frac{\mu_Y^2}{\sqrt{\mu_Y^2 + \sigma_Y^2}}\right)$$

$$\sigma = \sqrt{\ln(\sigma_Y^2 / \mu_Y^2 + 1)}$$

Proof

The mean μ_Y and variance σ_Y^2 of a lognormal random variable Y with characterizing parameters μ and σ are the following well-known formulas (Johnson and others, 1994, p. 212).

$$\mu_Y = e^{\mu + \sigma^2/2}$$

$$\sigma_Y^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

We now solve these two equations for the lognormal characterizing parameters μ and σ.

Solving for σ

$$\sigma_Y^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

$$= (e^{\mu+\sigma^2/2})^2 (e^{\sigma^2} - 1)$$

$$= \mu_Y^2 (e^{\sigma^2} - 1)$$

$$\frac{\sigma_Y^2}{\mu_Y^2} + 1 = e^{\sigma^2}$$

$$\sigma^2 = \ln(\frac{\sigma_Y^2}{\mu_Y^2} + 1)$$

$$\sigma = \sqrt{\ln(\frac{\sigma_Y^2}{\mu_Y^2} + 1)}$$

Solving for μ

$$\mu_Y = e^{\mu+\sigma^2/2}$$

$$\ln \mu_Y = \mu + \sigma^2/2$$

$$\mu = \ln \mu_Y - \sigma^2/2$$

$$= \ln \mu_Y - (1/2)\ln(\frac{\sigma_Y^2}{\mu_Y^2} + 1)$$

$$= \ln \mu_Y - \ln \sqrt{\frac{\mu_Y^2 + \sigma_Y^2}{\mu_Y^2}}$$

$$= \ln\left(\frac{\mu_Y}{\sqrt{\mu_Y^2 + \sigma_Y^2}/\mu_Y}\right)$$

$$= \ln \left(\frac{\mu_Y^2}{\sqrt{\mu_Y^2 + \sigma_Y^2}} \right)$$

Mean and standard deviation of a product of two independent random variables

Theorem 2

If X and Y are independent random variables with means $E[X]$ and $E[Y]$ and standard deviations $S[X]$ and $S[Y]$, then

$E[XY] = E[X]E[Y]$

$S[XY] = \{(S[X])^2(S[Y])^2 + (E[Y])^2(S[X])^2 + (E[X])^2(S[Y])^2\}^{1/2}$

Proof

From Ross (2000, p. 50) we have the following theorem:

If X and Y are independent, then for any functions g and h

$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$

Therefore

$E[XY] = E[X]E[Y]$

and

$E[X^2Y^2] = E[X^2]E[Y^2]$

From Ross (2000, p. 45) we have

Variance of X: $V[X] = E[X^2] - (E[X])^2$

Variance of Y: $V[Y] = E[Y^2] - (E[Y])^2$

We now consider

Variance of XY: $V[XY] = E[(XY)^2] - (E[XY])^2$

$= E[X^2Y^2] - (E[X]E[Y])^2$

$= E[X^2]E[Y^2] - (E[X])^2 (E[Y])^2$

$= E[X^2]E[Y^2] - (E[X])^2 E[Y^2] - (E[Y])^2 E[X^2] + (E[X])^2(E[Y])^2$

$+ (E[X])^2E[Y^2] + (E[Y])^2E[X^2] - (E[X])^2(E[Y])^2 - (E[X])^2(E[Y])^2$

$= \{E[X^2] - (E[X])^2\}\{E[Y^2] - (E[Y])^2\} + (E[X])^2E[Y^2]$

$+ (E[Y])^2E[X^2] - (E[X])^2(E[Y])^2 - (E[X])^2(E[Y])^2$

$= V[X]V[Y] + (E[Y])^2\{E[X^2] - (E[X])^2\} + (E[X])^2\{E[Y^2] - (E[Y])^2\}$

$= V[X]V[Y] + (E[Y])^2 V[X] + (E[X])^2 V[Y]$

$S[XY] = \{(S[X])^2(S[Y])^2 + (E[Y])^2(S[X])^2 + (E[X])^2(S[Y])^2\}^{1/2}$

TABLES

Table 1. Summary of recorded numbers and costs of landslides in San Francisco Bay region of California. Sources of data for this table, and tables 2 and 6, are Taylor and Brabb (1972), Taylor and others (1975), Creasy (1988), and Godt and others (1999). Numbers of damaging landslides are taken from published text when available or, if written values are unavailable, from counted landslide locations on published maps. Costs were converted to August 2000 dollars using the Consumer Price Index (CPI) for shelter and guidelines described by the U.S. Department of Labor (1997). The percent change from each period to August 2000 was determined using the formula $(((CPI_{August, 2000} - CPI_{previous\ period})/CPI_{previous\ period})*100)$. CPI values used were 30.5 for March 1969; 37.5 for March 1973; 97.0 for February 1982; and 222.9 for August 2000. Percent change values to August 2000 were 630.8% from March 1969; 494.4% from March 1973; 129.8% from February 1982; and 17.8% from February 1998. Although no data were recorded for Napa County in 1982, we assume that the number of landslides and costs were zero based on a statement by LaVopa Creasey (1988) that the county had "sustained relatively few landslides." Designations used in this table: no. of yrs., number of years; 10^6$, million dollars.

Table 1. Probabilistic Landslide Assessment Cost Estimation System (PLACES)				
	Summary of recorded numbers and costs of landslides.			
San Francisco Bay region county	Past Time (no. of yrs)	No. of Landslides	Recorded Costs (10^6 $)	Cost per Landslide (10^6 $)
Alameda	39	256	73.338	0.28648
Contra Costa	39	444	95.825	0.21582
Marin	39	442	71.347	0.16142
Napa	39	45	15.867	0.35260
San Francisco	39	39	9.632	0.24697
San Mateo	39	356	129.636	0.36415
Santa Clara	39	72	25.065	0.34813
Santa Cruz	35	635	77.999	0.12283
Solano	39	51	7.014	0.13753
Sonoma	39	195	77.457	0.39722
Total		2535	583.18	

Table 2. Recorded numbers and costs of landslides per landslide cluster. See caption of table 1 for additional information regarding data sources. Designations used in this table: no. of yrs., number of years; S.D., standard deviation.

Table 2. Recorded numbers and costs of landslides per landslide cluster.												
San Francisco Bay region county	No. of landslides per cluster.				Sample Mean	Sample S.D.	Cost of landslides per cluster (millions of $)				Sample Mean	Sample S.D.
	1968-69	1972-73	1982	1997-98			1968-69	1972-73	1982	1997-98		
Alameda	58	24	87	87	64.000	29.967	39.439	2.135	8.180	23.584	18.335	16.718
Contra Costa	70	110	145	119	111.000	31.102	37.872	10.029	16.118	31.806	23.956	13.047
Marin	66	153	197	26	110.500	78.335	7.710	18.215	42.430	2.992	17.837	17.587
Napa	1	8	20	16	11.250	8.461	10.801	0.780	2.967	1.319	3.967	4.650
San Francisco	9	8	17	5	9.750	5.123	0.972	2.913	0.917	4.830	2.408	1.862
San Mateo	70	54	191	41	89.000	69.027	26.302	21.371	17.173	64.790	32.409	21.907
Santa Clara	12	16	34	10	18.000	10.954	13.880	0.892	1.340	8.953	6.266	6.281
Santa Cruz			470	165	317.500	215.668			60.706	17.293	39.000	30.698
Solano	3	19	23	6	12.750	9.743	0.029	0.169	0.926	5.890	1.754	2.786
Sonoma	45	5	138	7	48.750	62.281	47.018	1.250	4.451	24.738	19.364	21.167
Total	334	397	1322	482								

Table 3. Numbers and recurrence intervals of landslide clusters. Designations used in this table: Prob., probability; distri., distribution; No., number; yrs., years; S.D., standard deviation; %, percent.

Table 3. Numbers and recurrence intervals of landslide clusters.								
	Geometric Distri.						Binomial Distri.	
San Francisco Bay region county	No. of Clusters	Prob. of Cluster	Recurrence Interval Mean (no. of yrs)	Recurrence Interval S.D. (no. of yrs)	Specified Time (no. of yrs)	Chance of one or more clusters (%)	No. of Clusters Mean	No. of Clusters S.D.
Alameda	4	0.103	9.750	9.236	5	41.787	0.513	0.678
Contra Costa	4	0.103	9.750	9.236	5	41.787	0.513	0.678
Marin	4	0.103	9.750	9.236	5	41.787	0.513	0.678
Napa	4	0.103	9.750	9.236	5	41.787	0.513	0.678
San Francisco	4	0.103	9.750	9.236	5	41.787	0.513	0.678
San Mateo	4	0.103	9.750	9.236	5	41.787	0.513	0.678
Santa Clara	4	0.103	9.750	9.236	5	41.787	0.513	0.678
Santa Cruz	2	0.057	17.500	16.993	5	25.487	0.286	0.519
Solano	4	0.103	9.750	9.236	5	41.787	0.513	0.678
Sonoma	4	0.103	9.750	9.236	5	41.787	0.513	0.678

Table 4. Future total numbers of landslides and their aggregation. The three aggregations assume respectively perfect positive correlation (p.p.c.), independence (indep.), and weighted-average correlation coefficient of 0.5 for illustrative purposes. Designations used in this table: No., number; S.D., standard deviation; %, percent; Aggre., Aggregation.

Table 4. Future total numbers of landslides and their aggregation.									
	Estimates of number of landslides during a specified time.								
San Francisco Bay region county	No. of Landslides Mean	S.D.	Lognormal Mu	Sigma	Prediction Level (%)	No. of Landslides Low	High	Specified No. (SN)	Chance Exceed SN (%)
Alameda	32.821	48.431	2.913	1.075	90	3	108	10	71.489
Contra Costa	56.923	78.527	3.509	1.032	90	6	183	10	87.868
Marin	56.667	93.628	3.379	1.147	90	4	194	10	82.592
Napa	5.769	9.745	1.078	1.161	90	0	20	10	14.587
San Francisco	5.000	7.564	1.014	1.091	90	0	17	10	11.884
San Mateo	45.641	78.031	3.137	1.169	90	3	158	10	76.240
Santa Clara	9.231	14.514	1.600	1.116	90	1	31	10	26.448
Santa Cruz	90.714	201.110	3.619	1.333	90	4	334	10	83.827
Solano	6.538	11.113	1.199	1.165	90	0	23	10	17.175
Sonoma	25.000	55.524	2.329	1.334	90	1	92	10	50.779
Aggre. (p.p.c.)	334.304	598.187	5.094	1.198	90	22	1160	100	65.845
Aggre. (indep.)	334.304	259.580	5.576	0.687	90	85	817	100	92.125
0.5	334.304	461.091	5.279	1.032	90	36	1072	100	74.314

Table 5. Future total costs of landslides and their aggregation. The three aggregations assume respectively perfect positive correlation (p.p.c.), independence (indep.), and weighted-average correlation coefficient of 0.5 for illustrative purposes. Designations used in this table: S.D., standard deviation; %, percent; Aggre., Aggregation; 10^6$, million dollars.

Table 5. Future total costs of landslides and their aggregation.									
	Estimates of total damage as economic loss during a specified time.								
San Francisco Bay region county	Total Costs Mean (millions of $)	S. D.	Lognormal Mu	Sigma	Prediction Level (%)	Total Costs Low (millions of $)	High	Specified Costs (SC) (10^6 $)	Chance Exceed SC (%)
Alameda	9.40231	17.26379	1.503	1.215	90	0.61000	33.15308	25	7.891
Contra Costa	12.28526	18.74614	1.907	1.097	90	1.10898	40.88935	25	11.581
Marin	9.14705	17.46521	1.445	1.239	90	0.55263	32.58918	25	7.622
Napa	2.03423	4.28147	-0.136	1.301	90	0.10276	7.41630	25	0.495
San Francisco	1.23487	2.10890	-0.472	1.168	90	0.09131	4.26416	25	0.079
San Mateo	16.62000	27.00948	2.164	1.137	90	1.34265	56.50378	25	17.682
Santa Clara	3.21346	6.18869	0.393	1.245	90	0.19112	11.47406	25	1.159
Santa Cruz	11.14271	26.05704	1.477	1.366	90	0.46294	41.46206	25	10.123
Solano	0.89923	2.32262	-1.125	1.427	90	0.03103	3.39707	25	0.117
Sonoma	9.93038	20.05839	1.483	1.275	90	0.54116	35.87101	25	8.665
Aggre. (p.p.c.)	75.90951	141.50172	3.580	1.224	90	5.03458	267.02005	100	20.123
Aggre. (indep.)	75.90951	53.21450	4.130	0.632	90	21.97078	175.84985	100	22.600
0.5	75.90951	106.89836	3.783	1.045	90	7.87301	245.34682	100	21.583

Table 6. Recorded percentages of public and private costs. See title of table 1 for additional information regarding data sources. Designations used in this table: %, percent; 10^6$, million dollars.

Table 6. Recorded percentages of public and private costs.

	1968-69			1972-73			1982			1997-98			Percent Public		Percent Private	
San Francisco Bay region county	Cost of Slides 10^6 $	Percent Public (%)	Percent Private (%)	Cost of Slides 10^6 $	Percent Public (%)	Percent Private (%)	Cost of Slides 10^6 $	Percent Public (%)	Percent Private (%)	Cost of Slides 10^6 $	Percent Public (%)	Percent Private (%)	Weighted Mean (%)	Weighted S.D. (%)	Weighted Mean (%)	Weighted S.D. (%)
Alameda	39.439	8.20	91.30	2.135	75.40	24.60	8.180	47.20	52.80	23.584	50.50	49.50	28.11	21.91	71.62	21.67
Contra Costa	37.872	70.50	27.80	10.029	57.80	42.20	16.118	39.30	60.70	31.806	72.00	28.00	64.42	12.02	34.91	12.38
Marin	7.710	79.90	7.80	18.215	64.30	35.70	42.430	56.50	43.50	2.992	42.00	58.00	60.41	8.26	38.26	11.60
Napa	10.801	29.00	54.10	0.780	98.50	1.50				1.319	100.00	0.00	40.46	26.00	45.39	19.76
San Francisco	0.972	24.80	75.20	2.913	100.00	0.00	0.917	22.10	77.90	4.830	0.00	100.00	34.85	43.80	65.15	43.80
San Mateo	26.302	33.20	34.60	21.371	64.30	35.70	17.173	51.80	48.20	64.790	64.00	36.00	56.18	12.29	37.28	4.30
Santa Clara	13.880	55.40	25.90	0.892	50.30	49.70	1.340	44.60	55.40	8.953	95.00	5.00	68.79	19.70	20.86	14.05
Santa Cruz							60.706	29.40	70.60	17.293	51.00	49.00	34.19	8.97	65.81	8.97
Solano	0.029	100.00	0.00	0.169	31.50	68.50	0.926	53.40	46.60	5.890	100.00	0.00	92.20	18.41	7.80	18.41
Sonoma	47.018	39.40	0.00	1.250	95.20	4.80	4.451	94.20	5.80	24.738	3.00	97.00	31.82	24.26	31.39	44.97

Table 7. Future public costs of landslides and their aggregation. The three aggregations assume respectively perfect positive correlation (p.p.c.), independence (indep.), and weighted-average correlation coefficient of 0.5 for illustrative purposes. Designations used in this table: S.D., standard deviation; %, percent; Aggre., Aggregation; 10^6$, million dollars.

Table 7. Future public costs of landslides and their aggregation.

	Estimates of public damage as economic loss during a specified time.								
San Francisco Bay region county	Public Costs Mean (millions of $)	Public Costs S. D. (millions of $)	Lognormal Mu	Lognormal Sigma	Prediction Level (%)	Public Costs Low (millions of $)	Public Costs High (millions of $)	Specified Costs (SC) (10^6 $)	Chance Exceed SC (%)
Alameda	2.643	6.489	-0.003	1.396	90	0.10027	9.91233	10	4.935
Contra Costa	7.914	12.373	1.450	1.112	90	0.68464	26.56267	10	22.173
Marin	5.526	10.676	0.932	1.247	90	0.32675	19.74678	10	13.586
Napa	0.823	2.126	-1.213	1.427	90	0.02840	3.10946	10	0.689
San Francisco	0.430	1.298	-2.000	1.521	90	0.01110	1.65188	10	0.234
San Mateo	9.338	15.668	1.565	1.157	90	0.71262	32.07007	10	26.181
Santa Clara	2.210	4.473	-0.021	1.276	90	0.12005	7.98800	10	3.431
Santa Cruz	3.810	9.264	0.371	1.391	90	0.14713	14.26695	10	8.237
Solano	0.829	2.190	-1.226	1.441	90	0.02743	3.14101	10	0.717
Sonoma	3.160	8.381	0.109	1.443	90	0.10380	11.97883	10	6.429
Aggre. (p.p.c.)	36.684	72.938	2.802	1.265	90	2.26218	130.42799	50	19.016
Aggre. (indep.)	36.684	27.235	3.383	0.663	90	9.90424	87.59104	50	21.223
0.5	36.684	55.053	3.013	1.086	90	3.40886	121.38080	50	20.379

Table 8. Future private costs of landslides and their aggregation. The three aggregations assume respectively perfect positive correlation (p.p.c.), independence (indep.), and weighted-average correlation coefficient of 0.5 for illustrative purposes. Designations used in this table: S.D., standard deviation; %, percent; Aggre., Aggregation; 10^6$, million dollars.

Table 8. Future private costs of landslides and their aggregation.									
	Estimates of private damage as economic loss during a specified time.								
San Francisco Bay region county	Private Costs Mean (millions of $)	Private Costs S. D. (millions of $)	Lognormal Mu	Lognormal Sigma	Prediction Level (%)	Private Costs Low (millions of $)	Private Costs High (millions of $)	Specified Costs (SC) (10^6 $)	Chance Exceed SC (%)
Alameda	6.734	13.078	1.126	1.250	90	0.39444	24.09555	10	17.327
Contra Costa	4.288	7.108	0.795	1.149	90	0.33453	14.67181	10	9.488
Marin	3.500	7.063	0.441	1.274	90	0.19101	12.63872	10	7.200
Napa	0.923	2.157	-1.013	1.366	90	0.03842	3.43509	10	0.761
San Francisco	0.805	1.742	-1.087	1.318	90	0.03858	2.95050	10	0.507
San Mateo	6.196	10.162	1.171	1.143	90	0.49256	21.12787	10	16.104
Santa Clara	0.670	1.620	-1.362	1.387	90	0.02618	2.50790	10	0.412
Santa Cruz	7.333	17.336	1.050	1.373	90	0.29856	27.33716	10	18.077
Solano	0.070	0.493	-4.617	1.980	90	0.00038	0.25660	10	0.024
Sonoma	3.117	11.872	-0.234	1.656	90	0.05198	12.05697	10	6.278
Aggre. (p.p.c.)	33.637	72.630	2.649	1.317	90	1.86662	121.07818	50	16.868
Aggre. (indep.)	33.637	28.753	3.241	0.741	90	7.56231	86.45081	50	18.259
0.5	33.637	55.235	2.862	1.143	90	2.66769	114.74080	50	17.921