USGS
science for a changing world

A Unified Approach To Probabilistic Risk Assessments for Earthquakes, Floods, Landslides, and Volcanoes

Proceedings of a Multidisciplinary Workshop held in Golden, Colorado
November 16-17, 1999
Sponsored by the U.S. Geological Survey Urban Hazards Initiative

Open-File Report 01-324

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

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

A Unified Approach To Probabilistic Risk Assessments for Earthquakes, Floods, Landslides, and Volcanoes

Proceedings of a Multidisciplinary Workshop held in Golden, Colorado
November 16-17, 1999
Sponsored by the U.S. Geological Survey Urban Hazards Initiative

Compiled by Aldo V. Vecchia

Open-File Report 01-324

Bismarck, North Dakota
2001

U.S. DEPARTMENT OF THE INTERIOR
GALE A. NORTON, Secretary

U.S. GEOLOGICAL SURVEY
CHARLES G. GROAT, Director

For additional information write to:

District Chief
U.S. Geological Survey
Water Resources Division
821 East Interstate Avenue
Bismarck, ND 58503

Copies of this report can be purchased from:

U.S. Geological Survey
Information Services
Box 25286
Denver, CO 80225-0286

CONTENTS

A Unified Approach To Probabilistic Risk Assessments for Earthquakes, Floods, Landslides, and Volcanoes

Proceedings of a Multidisciplinary Workshop held in Golden, Colorado
November 16-17, 1999
Sponsored by the U.S. Geological Survey Urban Hazards Initiative

Compiled by Aldo V. Vecchia

INTRODUCTION

In November 1999, the U.S. Geological Survey (USGS) Urban Hazards Initiative sponsored a workshop to foster communication among experts in probabilistic risk assessment for earthquakes, floods, landslides, and volcanoes and to determine if a more unified framework for assessing risk from the different hazards can be developed. The workshop was held November 16-17, 1999, at the USGS Building on the Colorado School of Mines Campus in Golden, Colo. The following scientists participated:

Aldo Vecchia[1]	U.S. Geological Survey	Bismarck, N. Dak.
Joseph Jones[2]	U.S. Geological Survey	Tacoma, Wash.
William Bakun	U.S. Geological Survey	Menlo Park, Calif.
Richard Bernknopf	U.S. Geological Survey	Menlo Park, Calif.
Allen Bradley	University of Iowa	Iowa City, Iowa
Charles Connor	Southwest Research Institute	San Antonio, Tex.
Roger Denlinger	U.S. Geological Survey	Vancouver, Wash.
Ute Dymon	Kent State University	Kent, Ohio
Art Frankel	U.S. Geological Survey	Golden, Colo.
Randall Jibson	U.S. Geological Survey	Golden, Colo.
Michael Karlinger	U.S. Geological Survey	Lakewood, Colo.
Upmanu Lall	Utah State University	Logan, Utah
Harry McWreath	U.S. Geological Survey	Fort Worth, Tex.
Manuel Nathenson	U.S. Geological Survey	Menlo Park, Calif.
Daniel O'Connell	Bureau of Reclamation	Lakewood, Colo.
David Perkins	U.S. Geological Survey	Golden, Colo.
William Roberds	Golder Associates	Redmond, Wash.
Brent Troutman	U.S. Geological Survey	Lakewood, Colo.
Rob Wesson	U.S. Geological Survey	Golden, Colo.

[1]Workshop coordinator.

[2]Representative of Urban Hazards Executive Committee.

The first day of the workshop consisted of a series of presentations aimed at describing the current tools and techniques used for hazard investigations and risk assessment. The purpose of these presentations was to increase awareness of the different approaches that can be used for assessing hazards and associated risk and to stimulate discussion

of the similarities and differences that exist among the different disciplines (earthquakes, floods, landslides, and volcanoes). The consensus of participants after the first day was that, although the language of probability theory is a common thread among hazard investigations, numerous differences exist among the specific approaches used for assessing risk. These differences exist among the different disciplines as well as among different applications within the same discipline. Imposing a rigid recipe for risk assessment in every application clearly is not advantageous. However, a great deal of progress can and should be made in unifying the products, such as hazards maps, the USGS provides for risk assessment and in improving the way hazards are communicated to our customers.

During the second day of the workshop, participants were divided into three panels with representatives from each discipline on each panel. Each panel was asked to discuss the similarities and differences in techniques among the disciplines and suggest areas in which the USGS can make its hazards investigations more unified and consistent among the disciplines. The panels also discussed emerging areas of research that may have significant consequences on the way in which risk assessment is done in the future.

A synopsis of the presentations given on the first day of the workshop is provided in the next section, and general results of the panel discussions on the second day of the workshop are summarized in a later section. A summary of the more technical aspects of hazard assessments is given in appendix A, abstracts of the presentations describing the current tools and techniques used for hazard investigations and risk assessment are given in appendix B, and selected papers are given in appendix C.

As indicated in the title, this workshop dealt with the general subject of probabilistic risk assessment for natural hazards. Risk assessment quantifies the potential economic and social impacts (loss of property, life, etc.) resulting from natural hazards. Hazard assessment, which is a necessary precursor to risk assessment, deals with evaluating the likelihood of occurrence of events, such as earthquakes or floods, that may have serious economic and social impacts.

Most USGS products currently being produced relate to hazard assessment. However, the customers of USGS hazard-assessment products are individuals, government entities, and cooperators who perform risk assessment. Therefore, the ultimate goal of this workshop is to make USGS hazard-assessment products more useful to risk assessors.

The term “hazard” is used in this report in two separate contexts. In the first context, hazard is a generic term for earthquakes, floods, landslides, and volcanoes. In the second context, which is standard in the technical literature on probabilistic risk assessment, hazard is synonymous with the probability of occurrence. Thus, for example, landslide hazard refers to the probability of occurrence of landslides, which are a particular type of hazard. Clarifying this terminology up front will hopefully avoid confusion on the part of the reader.

Much was accomplished from the workshop, and participants should be commended for an impressive effort. However, as expected, 2 days was not nearly enough for the group to explore all the issues involving risk assessment for earthquakes, floods, landslides, and volcanoes, much less recommend a master plan for unifying the disciplines. Future efforts should be made to build on this workshop and further enhance cooperation between the disciplines.

SYNOPSIS OF PRESENTATIONS AND RELATED DISCUSSION

The following synopsis was prepared by the workshop coordinator on the basis of notes from the presentations and related discussions on the first day of the workshop. Italicized type indicates comments by discussants and/or observations of the workshop coordinator that do not necessarily represent the view of the presenter.

Flood-Risk Assessment

The first day began with talks relating to flood-risk assessment by Allen Bradley of the University of Iowa, Brent Troutman of the USGS, and Upmanu Lall of Utah State University (now at Columbia University). Allen Bradley started by discussing flood-frequency distributions and recommended the partial-duration or threshold method along with estimation using L-moments to determine flood frequencies at sites with stationary record. *This recommendation is in contrast to federally-mandated methods (United States Water Resources Council, 1981) used by agencies such as the U.S. Army Corps*

of Engineers (COE) and the Federal Emergency Management Agency (FEMA). Many flood-hazard assessments of practical importance involve some sort of nonstationarity brought about by land-use modifications or flow regulation and for which too few data are available for standard statistical analyses. In these assessments, the effects of nonstationarity can be modeled through sound application of rainfall-runoff models and hydraulic-engineering practices. The extreme upper tail of a flood-frequency distribution is particularly important for design of high-risk structures. Rainfall-runoff models are the most reliable method for determining the upper tail because data records are not long enough to make this determination with any accuracy. *Paleoflood data also are useful for estimating the upper tail in cases when flow has not been modified and climatic conditions are assumed to be unchanged.*

Brent Troutman discussed the spatial aspects of flood-frequency analysis (Troutman and Karlinger, appendix C). He started by giving a summary of techniques for developing flood-frequency distributions at sites with limited or no data. Techniques such as those developed by Stedinger and Tasker (1985) for dealing with unequal sample sizes, cross correlation between sites, etc., when estimating flood frequencies are widely used both within and outside the USGS. The methodology of developing flood-hazard estimates at ungaged sites is widely referred to as flood-quantile regionalization. Flood-quantile regionalization equations depend on drainage-basin geometry, topography, soil properties, and climate. Much research is being conducted on relating flood quantiles to basin characteristics to develop more parsimonious and physically-based models.

Troutman also discussed the important problem of estimating regional flood probabilities. Regional flood-probability analysis deals with estimating the joint probability distribution of flood peaks at numerous sites (in contrast to regionalization techniques described earlier, which deal with estimating single-site flood probabilities). For example, an analysis of data from the Puget Sound area showed that a "one-hundred year flood" occurs at five sites in the Cascade Range an average of once every 24 years (indicating large floods occur nearly independently at these sites) and at five sites on the Olympic Peninsula an average of once every 40 years (indicating large floods are highly dependent at these sites). Important research problems include relating spatial correlation in flood peaks to drainage-network geometry and precipitation and using asymptotic theory to simplify the analysis for a large number of sites.

Upmanu Lall rounded out the flood-risk section by discussing the effect of low-frequency climate variability on evaluating changing flood risk. He presented convincing evidence that low-frequency variations such as the El Nino Southern Oscillation cause decadal and longer-scale fluctuations in precipitation patterns and, hence, floods. These variations occur on a global scale and affect some regions of the world much more than others. The signal is particularly strong in the Pacific Coast region of the United States and shows up in other regions of the country as well. This low-frequency variation presents far-reaching implications on the way in which flood-hazard assessment is performed. The long period of the fluctuations means historical flood records may not contain as much information as previously thought because annual flood peaks are long-range-dependent random variables. Hence, many historical records may be too short to accurately determine flood risk. *Many areas of the country are experiencing a much higher incidence of damaging floods (or droughts) in recent years than probability models predict, due in large part to inadequacy of established models for incorporating long-term climatic variability.*

Lall also discussed sophisticated time-series methods for analyzing low-frequency variability and estimating conditional flood probabilities. Further research is required to develop the link between low-frequency precipitation variability and streamflow. Although the link between precipitation and seasonal or annual streamflow volumes is strong, it is difficult to accurately determine the effect of low-frequency variability on peak flows.

Landslide-Risk Assessment

Presentations relating to landslide-risk assessment were made by William Roberds of Golder Associates, Richard Bernknopf of the USGS, and Randall Jibson of the USGS. William Roberds began by giving an overview of landslide-risk assessment and presenting several interesting case studies (Roberds, appendix C). He made the distinction between a landslide hazard, which deals with the number, timing, location, spatial extent, etc., of landslides and runout, and a landslide risk, which deals with casualties, property damage, and socioeconomic impacts of landslides. Valuable time and resources should be used to reduce uncertainties in the landslide hazard only if those uncertainties have consequences in terms of estimating risk. Thus, hazard and risk assessments for landslides (or any other hazard) should be considered in

tandem. *This is an important point often overlooked by investigators who deal primarily with hazard assessment and have limited experience or interest in risk assessment.*

Risk assessment for landslides is particularly difficult because of the highly complex physical processes involved and the sparseness of data for fitting models. A key tool for evaluating risk is probabilistic dynamic simulation, which can be used to simulate the probability of various outcomes from landslides (e.g., detachment of shallow zones or wedges, runout of rockfalls, etc.) given uncertainties in the inputs (e.g., soil properties, depth to phreatic surface, seismic loading, fracture geometry, etc.). The conditional landslide risk is the potential damage that results from a specific landslide event. The conditional risk often is assumed to be known with certainty. However, by stochastically simulating the model inputs, the "risks" associated with a given landslide event can be quantified in terms of a probability distribution. Sensitivity analysis also can be used to determine which inputs are producing the most uncertainty in the estimated risks (and, hence, require further data to reduce uncertainty) and which risk-management activities are most cost effective.

Richard Bernknopf talked about stochastic landslide-forecasting models being developed by USGS and Stanford University scientists, with particular emphasis on the economic aspects of risk assessment. He, in agreement with Roberds, emphasized that landslide-hazard and -risk assessments must be considered in tandem and that sound probabilistic methods must be used to quantify uncertainties in the outcomes of landslides before policy decisions can be made. He illustrated the concepts with three case studies--construction-induced landslides in Cincinnati, Ohio; earthquake-triggered landslides in Santa Cruz, Calif.; and rainfall-triggered debris flows in Oakland, Calif. He showed how Earth Science Information (ESI) could be used in each case to help forecast landslide occurrence and evaluate risk-mitigation alternatives (Bernknopf and others, 1993). He also quantified the economic utility of ESI for landslide-risk assessment. *The producers, such as the USGS, of ESI often may not fully evaluate the utility of the information for policy and decision makers. The USGS is primarily an earth-science information and research provider, and policy decisions usually are left to other organizations. However, we need to keep the needs of policy and decision makers in mind when developing ESI products.*

Randall Jibson's presentation dealt with modeling spatial properties of landslides and producing landslide-hazard maps (Jibson, appendix C). He began by pointing out that the temporal frequency of landslides is dependent on the magnitude-frequency distribution of other events, such as earthquakes and large storms, that trigger landslides. Therefore, landslide-hazard assessment generally focuses on conditional hazard assessment in which the spatial extent and magnitude of a landslide event is predicted given that an earthquake or rainfall event of a certain intensity occurs. With the data and models currently available, earthquake-triggered landslides provide the best opportunity for quantitative prediction.

Jibson also described a methodology for producing regional landslide-hazard maps for earthquake-triggered landslides. The maps show the probability of slope failure for a given level of earthquake-shaking intensity for each 10-meter pixel. The probability of slope failure for each pixel is a function of an index of dynamic-slope stability (critical acceleration) that is computed using detailed geologic maps, shear-strength data from direct field tests, and digital-elevation data. The probability of failure of a slope with a given critical acceleration is defined as the proportion of all pixels with that critical acceleration value that fail. Given a strong-motion record of an earthquake, the probability of failure is calculated by first computing the Newmark displacement, which depends on the critical acceleration and strong-motion record, and then expressing the failure probability as a function of Newmark displacement. An application of the methodology for the 1994 Northridge, Calif., earthquake shows an excellent agreement between the observed proportion of slope failures and the calculated failure probabilities using the Newmark displacement. The primary obstacle remaining for routinely applying the method is gathering the data required to perform a rigorous analysis of uncertainty in the model inputs.

Earthquake-Risk Assessment

Presentations relating to earthquake-risk assessment were made by Art Frankel, William Bakun, and David Perkins, all from the USGS. Art Frankel began with a summary of the methodology used to construct the USGS National Seismic Hazard Maps. The maps depict peak horizontal ground acceleration and spectral response at various periods corresponding to 10-, 5-, and 2-percent probabilities of exceedance in 50 years. Maps showing the probability of exceedance for given peak horizontal ground acceleration also are produced. Potential earthquake sources considered in the hazard calculations are spatially-smoothed historic seismicity, large background source zones based on geologic criteria, and specific fault sources. Logic trees are used to incorporate complex information involving different seismicity

models, fault recurrence models, Cascadia great earthquake scenarios, and ground-motion attenuation relations. Deaggregation plots are produced for selected cities to show the contribution to the hazard from potential earthquakes with given magnitudes and distances. The hazard maps and associated products are available in several media, including internet and CD-ROM.

William Bakun presented some recent results from the Bay Area Earthquake Hazard Program of the USGS (Michael and others, 1999). Scientists from the USGS, in cooperation with more than 100 scientists from other governmental, academic, and private organizations, determined a 70-percent ($\pm$10 percent) probability exists for a magnitude 6.7 or greater earthquake in the San Francisco Bay region by 2030. Because of the high probability, relatively short timeframe, and relatively small region, the result was particularly noteworthy, receiving much attention in the media. A number of innovative techniques, such as stochastic earthquake simulation models that calculate dynamic risk and account for interaction between different faults, were used to determine the probability. *Although making "predictions" is risky, in order for hazard assessments to affect future policy decisions and cause fundamental changes in public thinking, investigators must push the limits of their work toward higher probabilities, smaller time intervals, and smaller spatial regions.*

David Perkins discussed the likelihood distribution of the occurrence rate of N earthquake events observed over time T. He showed that the rate distribution under the Poisson assumption is independent of the uncertainty in the estimated times of historic events, and the uncertainty in rate is very large when the number of historic events is small. Asserting time dependence, as with the Weibull distribution, can narrow the likelihood distribution of the rate, but the narrowing is limited by the date uncertainty. Furthermore, when two event dates can be so close that they cannot be distinguished because of dating uncertainty, the number of events cannot be known for sure, and a truncated Poisson distribution may be mistakenly identified as a weakly periodic Weibull distribution. Because these two distributions can result in widely different estimated future occurrence probabilities (depending on the time elapsed since the last earthquake), it is important to develop methods to distinguish between the two. In particular, more precise identification and dating of historic events is required to establish the existence of time dependence. In any case, the rate uncertainty usually is much larger than usually is estimated from the standard deviation of the interoccurrence times, and this larger uncertainty has to be taken into account when making future occurrence probability statements.

A paper that describes the methodology used to estimate spatially smoothed earthquake probabilities is included in this report (Perkins, appendix C). Although the paper is not directly related to the oral presentation by Perkins, it provides a good overview of techniques for producing earthquake hazard maps.

Volcano-Risk Assessment

Presentations relating to volcano-risk assessment were made by Charles Connor of the Southwest Research Institute, Manuel Nathenson of the USGS, Roger Denlinger of the USGS, and Daniel O'Connell of the Bureau of Reclamation (BOR). Charles Connor began by giving an overview of current techniques used for volcanic-hazard and -risk assessments (Connor and others, appendix C). A full volcanic-hazard assessment requires integrating the effects of many potential "subhazards," including gas emissions, tephra fallout, debris avalanches, and lahars. Each type of volcanic hazard requires different physical models and different data inputs to evaluate the potential hazard. Connor focused on tephra fallout, but the same general methods and techniques can be applied to other volcanic hazards.

The magnitude-frequency relation for the temporal occurrence of volcanic eruptions is particularly difficult to estimate because volcanic eruptions occur relatively infrequently. Although some modern eruptions exist for which the timing and magnitude of the eruption are precisely known, geologic records of past eruptions often cannot be used to accurately determine the timing and frequency of eruptions. The difficulty is compounded by the inherent nonstationarity of eruptions through time, with long periods of dormancy interspersed with periods of unrest during which the probability of an eruption changes considerably.

Estimation of the conditional probability distribution of tephra accumulation depths given that an eruption occurs can be done using a combination of physically- and empirically-based simulation models. Connor described a modeling technique that has been successful in simulating tephra fallout for a variety of recorded volcanic events. Given model inputs such as eruption volume, column height, eruption velocity, windspeed and direction, and particle grain-size

distribution, the model computes the mass of tephra accumulated at any given spatial location surrounding the volcanic vent. The most important controlling model inputs are eruption volume, column height, and eruption velocity. These inputs are stochastically generated from probability distributions that are developed on the basis of past records of volcanic activity from the volcano being simulated or from similar volcanoes from other locations. By computing tephra-fallout realizations for numerous stochastically-generated inputs, a wide range of likely eruption scenarios are simulated and used to develop a probability distribution of tephra depth at any given spatial location. Hazard maps can be produced by contouring exceedance probabilities at different locations for a given tephra thickness. The practical usefulness of the methods was demonstrated through application of the technique for the Cerro Negro volcano in Nicaragua.

Manuel Nathenson focused on stochastic modeling of the temporal occurrence of volcanic eruptions in the Cascades (Nathenson, appendix C). Geologic histories of volcanic activity from the Medicine Lake volcano show a definite tendency for disparate time intervals between eruptions, with 2 long time intervals (greater than 1,500 years) and 13 short time intervals (less than 780 years). If the occurrence of volcanic eruptions follows a Poisson process, as often is assumed in practice, the time intervals between eruptions would follow an exponential distribution. However, the exponential distribution provides a poor fit to the Medicine Lake data, with the frequency of short time intervals greatly underestimated. The Weibull distribution, proposed by other volcanologists, and a mixed exponential distribution, proposed by Nathenson, both fit the Medicine Lake data much better than the exponential distribution. The mixed exponential distribution has a conceptual advantage over the Weibull distribution because the mixed distribution is analogous to a Poisson process but with two distinct mean recurrence rates.

The nonexponential distribution of times between eruptions has important implications for determining future eruption probabilities. The exponential distribution is the only distribution that has the so-called "lack-of-memory" property, in that the time until the next eruption takes place is independent of the time elapsed since the last eruption. For the Weibull and mixed exponential distributions, the time until the next eruption depends on the time elapsed since the last eruption--as the time elapsed since the last eruption increases, the expected time until the next eruption usually increases. The implications for computing future eruption probabilities can be significant, depending on the time elapsed since the last eruption. For example, the probability of an eruption in the next 30 years computed using the Weibull or mixed exponential distribution can be several times larger (if the last eruption was relatively recent) or smaller (if the last eruption was relatively distant) than the corresponding probability computed using the exponential distribution. *Clustering also can arise if the times between successive eruptions are nonstationary (the distribution changes through time) or autocorrelated (short intervals tend to follow short intervals and long intervals tend to follow long intervals). It is difficult to detect nonstationarity or autocorrelation from existing data because most data sets consist of only a few known events.*

Roger Denlinger and Daniel O'Connell discussed statistical estimation of debris-flow hazards. Denlinger began by describing the model and data uncertainties that are inherent in any debris-flow hazard assessment. Debris-flow hazard assessments, like other hazard assessments, must rely on sound geologic estimates of the time and magnitude of extreme events. Therefore, uncertainties in the age of debris-flow deposits (based on C14 dating) and uncertainties in the volume of deposits (based on estimated debris thickness and extent) must be quantified carefully. The uncertainty in volume estimates is particularly high and provides a large source of potential error in determining the risk from future extreme events. Therefore, physically-based flow-routing models need to be used to reduce uncertainties in volumes of deposits and provide the relation between runout distance and volume. Application of purely deterministic models is difficult because of the complexity of the processes involved and the uncertainty in the inputs to the physical model. Therefore, a combination of Bayesian statistical analysis and deterministic modeling was developed for estimating debris-flow hazards. The technique used is a modification of the methodology developed by O'Connell for flood-hazard estimation.

O'Connell described the Bayesian approach in more detail. Bayesian analysis provides a convenient and statistically rigorous way to incorporate data and model uncertainty into debris-flow hazard assessments and is particularly well-suited for dealing with incomplete information. For example, geologic data may provide an upper or lower bound on flow volume without providing an actual value. Such censored data can be incorporated easily in a Bayesian framework. Bayesian analysis also can use expert judgment to refine and improve estimates. Unknown parameters in the model are estimated by maximizing the posterior likelihood function, and, thus, the large body of theory for Bayesian likelihood estimation can be used to determine the statistical properties of the estimators and compute confidence or prediction intervals on model output. The Bayesian technique was demonstrated for estimating the debris-flow hazard for Mt. Rainier.

Incorporating uncertainties in model inputs was an integral part of the previous hazard assessments--for example, the probabilistic dynamic simulation discussed by William Roberds and the stochastic analysis of tephra fallout discussed by Charles Connor. Although these methods are similar to the Bayesian approach, the Bayesian formalism provides an alternative interpretation of model results.

OVERVIEW OF PANEL DISCUSSIONS AND RECOMMENDATIONS

The second day of the workshop began with a presentation by Ute Dymon on general aspects of hazard mapping (Dymon, appendix C). She raised many interesting points regarding the need for more unified standards for producing hazard maps, stimulating much discussion and helping set the stage for subsequent panel discussions.

Following Dymon's presentation, participants were divided into three panels with representatives from each of the disciplines on each panel. Each panel also had both USGS and non-USGS members. The panels were presented with four open-ended discussion topics:

1. What are the similarities and differences among earthquake-, flood-, landslide-, and volcano- (E-F-L-V-) hazard assessments?

2. Which USGS hazard-assessment products are beneficial for risk assessment, and how can products be improved or new products developed to support risk assessment?

3. Which research topics are most important for the future of E-F-L-V-hazard assessments?

4. How can hazard and risk assessments be made more unified among E-F-L-V?

The questions were somewhat vague and open-ended to encourage discussion and to allow each panel to take different directions and emphasize different themes. At the end of the day, all three panels met and discussed their findings with the entire group. The account presented here is a synthesis of findings from all three panels and was compiled by the workshop coordinator from meeting notes and written summaries provided by each panel leader.

Discussion Topic 1--What are the Similarities and Differences Among Earthquake-, Flood-, Landslide-, and Volcano-Hazard Assessments?

Several key similarities were noted by all three panels. For example, all hazard assessments must account for temporal and spatial nonstationarity, use a combination of physical and empirical models to evaluate hazard probabilities, evaluate the effects of model and data uncertainties, and use regionalization methods to develop smooth hazard maps. However, the specific methods for estimating nonstationarity, the types of physical and empirical models used, the degree of model and data uncertainty, and the specific regionalization methods used are different among the various hazards. These similarities and differences will be discussed separately for two distinct components of a hazard assessment--determination of a magnitude-frequency distribution for the temporal and spatial occurrence of source events and estimation of the hazard at a given location on the basis of the occurrence of source events at nearby locations.

The discussion in this section is kept to a nontechnical level. However, much constructive work of a more technical nature was done by the panels and should be highly enlightening for those with some background in probabilistic hazard assessment. Therefore, a detailed technical treatment of this discussion topic is presented in appendix A.

Temporal and Spatial Distribution of Source Events

Examples of source events are a severe rainstorm centered on a particular location, an earthquake with epicenter at a given location, or an eruption of a particular volcano. Workshop participants generally agreed source events are nonhomogeneous in space and nonstationary in time, and the probability of a source event is a complex function of both location and time. For rainstorms, in any given year, the probability of a source event depends on spatial differences in topography and atmospheric circulation patterns that change relatively slowly with time (here, atmospheric circulation

patterns refer to average annual climatic conditions, not day-to-day variability). Within time scales spanning several years, the probability of a source event may be relatively constant with respect to time. However, over longer time scales, long-term changes in climate circulation patterns (and even longer-term changes in topography) cause the probability of a source event to change. This longer-term variability with respect to time is an important consideration in flood-hazard assessment. Short-term variability in climate (on the order of a few years) is an important consideration for prediction but usually is considered part of the random "noise" for flood-hazard assessment. For earthquakes and volcanic eruptions, the probability of a source event depends on spatial variability in geologic and tectonic properties that change relatively slowly with time, and temporal changes occur over long time periods due to changes in stress with time, formation of new volcanic vents or geologic faults, etc.

A stochastic model for describing the complicated temporal and spatial structure of source events is a key component in any hazard analysis. However, fundamental differences exist in the types of stochastic models that are used among the various hazards. Volcanic eruptions are the least complicated events to model (in terms of probability of occurrence, not in terms of evaluating consequences such as ash deposition or debris flows) because eruptions generally occur at fixed and known locations. Furthermore, limited spatial interaction exists between locations so the probability of occurrence of an eruption can be evaluated separately for each location. Thus, the primary difficulty in evaluating eruption probabilities is in determining the temporal frequency distribution. Times between eruptions usually are assumed to be independent and identically distributed random variables (i.e., each time an eruption takes place, the volcano returns to the same state). However, the probability distribution of the waiting time until the next eruption depends on the time elapsed since the previous eruption. Estimating the distribution of the inter-eruption times is complicated by the sparseness of historic events and uncertainties in the dates and magnitudes of prehistoric events.

Earthquake probabilities are more complicated to model than eruption probabilities because source events can occur either along known faults, undiscovered faults, or essentially anywhere in a continuous earthquake source zone. Therefore, simplifying assumptions need to be made to allow earthquake probabilities to be computed at specific locations or over a specific region. The technical term for the stochastic process that embodies these assumptions is a space-time Poisson process. Loosely speaking, the occurrence of earthquakes in disjoint regions or disjoint time intervals are assumed to be independent events with a magnitude-frequency distribution that depends on location. Statistical smoothing procedures are required to estimate the spatial structure of the magnitude-frequency distribution and produce spatially continuous estimates of earthquake probabilities. The earthquake probabilities are complicated functions of three-dimensional geologic properties that often are subject to considerable uncertainty.

Rainfall probabilities are the most difficult to model because of the unlimited scope of potential source events that must be considered when evaluating flood hazards. Every rainstorm has a different temporal and spatial signature that defies classification. Thus, evaluation of flood hazards directly from rainfall events requires a time series of rainfall intensities from every location in a drainage basin. Furthermore, significant temporal memory exists in basin conditions, such as soil moisture or snow pack, so the flood hazard at any given time may depend on the integrated effect of precipitation from many individual storm events. For these reasons and others, a flood hazard rarely is calculated using rainfall source events (except for design of high-risk structures, in which an extreme flood may be evaluated using a “worse-case scenario” for precipitation). Rather, the “source event” for flood-hazard analysis usually is defined as the event that the discharge at a particular cross section of a river or stream exceeds a given value. Historical data from numerous stream gages operated by the USGS and other agencies are used to evaluate discharge-frequency relations at gaged locations and relate the frequencies to basin characteristics.

Hazard Assessments

Hazard assessments involve estimating the potential hazard intensity at a particular location given potential source events at neighboring locations. For example, the probability that a certain critical peak ground acceleration (the hazard intensity) is exceeded at a particular bridge location during the next 100 years depends on the potential for earthquakes from nearby faults or source zones. Other examples are estimating the probability of flood inundation for a particular location during a specified time interval or estimating the depth of tephra accumulation resulting from an eruption of a certain magnitude for locations surrounding a volcano.

A number of similarities exist in the ways in which hazard assessments are carried out for E-F-L-V. Workshop participants generally agreed physically- and empirically-based methods must be used in tandem to fully evaluate the hazards. Purely empirical methods are unreliable for extrapolating outside the range of recorded events, and physically-based models must be calibrated and verified using recorded data before the models can be used with confidence. Combining the best of both worlds is the most expedient approach for any hazard assessment. Relatively simple physical models (i.e., models that require few inputs and that are based on straightforward equations) often may be preferable to more complicated models. The modified Suzuki model described by Connor (appendix C) for predicting tephra accumulation is one such example. The required inputs (eruption volume, column height, eruption velocity, windspeed and direction, etc.) are relatively simple and yet account for much of the spatial and temporal variability of tephra accumulation as verified from calibrating the model to recorded events. Another example is the Newmark displacement analysis described by Jibson (appendix C) for estimating landslide susceptibility. The model is relatively simple and yet does an excellent job of predicting slope failures as verified by the Northridge earthquake data. Rainfall-runoff models described by Bradley (appendix B) for estimating extreme flood hazards, flow-routing models described by Denlinger and O'Connell (appendix B) for estimating volcanic-debris-flow hazards, and ground-motion attenuation models described by Frankel and others (appendix B) for estimating earthquake hazards are other examples of the utility of physically- or empirically-based models for hazard assessments.

Workshop participants also generally agreed evaluation of model uncertainty is a key component of any hazard assessment. Although most physically-based models are deterministic (the same inputs always produce the same output), knowledge about model inputs is almost always incomplete. Simplifying assumptions also are required in many cases to make the model tractable. Because small errors in estimated model parameters and/or minor deviations from model assumptions could result in large errors when using the model to extrapolate beyond the range of recorded events, the degree of uncertainty in the model and the effect of such uncertainty on evaluating the potential hazard must be quantified. Methods commonly used for evaluating model uncertainty include sensitivity analysis, stochastic simulation analysis, and Bayesian analysis. In a sensitivity analysis, model inputs are varied either individually or in tandem (within reasonable limits) to evaluate bounds on the model output. If the estimated hazard is not sensitive to uncertainties in the inputs, the model is robust and further data refinement is not required. However, if the estimated hazard is particularly sensitive to uncertainties in the inputs, further data refinement may be required (for example, field tests or higher-resolution geologic maps may be required). Sensitivity analysis may be used to provide a "worst-case" scenario for extremely-high-risk structures such as nuclear powerplants or dams.

In stochastic simulation analysis, model inputs are generated at random from probability distributions designed to cover the range of potential values of the inputs. The probabilistic dynamic-simulation technique described by Roberds (appendix C) is an example of this technique. By generating a large number of model realizations for different combinations of inputs, the model output can be evaluated probabilistically. Bayesian analysis is a statistical technique that, roughly speaking, combines both sensitivity analysis and stochastic simulation analysis to evaluate uncertainty in hazard estimates. The interpretation of model output provided by the Bayesian analysis is an alternate to that provided by classical statistical techniques such as maximum likelihood estimation or stochastic simulation analysis.

Discussion Topic 2--Which U.S. Geological Survey Products are Beneficial for Risk Assessment, and How Can Products Be Improved or New Products Developed to Support Risk Assessment?

Risk assessment deals with evaluating the potential consequences (casualties, loss of property, etc.) resulting from the occurrence of hazards and evaluating the methods for mitigating those consequences. The specific approach used in a particular risk-assessment study is highly dependent on the location and purpose of the study. For example, a flood-risk assessment would be handled differently by the Bureau of Reclamation for designing a dam than by a county or municipality for developing zoning laws or building a flood-mitigation project. Likewise, an earthquake-risk assessment would be handled differently by the Department of Energy for designing a nuclear reactor than by a county for developing building codes. Because of the complexity and diversity of risk-assessment studies, a full treatment of risk assessment was beyond the scope of this workshop. Another workshop focusing on risk assessment and building on the findings of this workshop would be a worthwhile undertaking.

The USGS needs to provide accurate, unbiased, and scientifically-defensible hazard maps at regional and national scales, provide digital access to maps and other hazard information products, conduct hazard education and research, and

promote standardized methods and tools for hazard assessment. Although all of these tasks are being conducted to various degrees, improvements can be made in a number of ways.

Maps produced by the USGS are used throughout the public and private sectors, not only for hazard assessments but for many other applications as well. However, a need exists for increased resolution, better accuracy, and better digital access of maps for making detailed hazard assessments. For example, a flood-hazard assessment at the detail needed by county or city agencies often requires digital-elevation data accurate to within 1 foot or less. Such accuracy is necessary for hydraulic flow-routing models and for mapping potential flood-inundation areas. Although emerging technologies are being developed to improve topographic-map resolution and accuracy, these technologies are at present too expensive to apply at a national level.

For earthquake-hazard assessments, high-resolution, accurate, three-dimensional geologic maps are needed for reducing uncertainty in hazard estimates. Such maps would help to better pinpoint potential earthquake source locations and to better account for small-scale variations in peak ground acceleration or liquefaction potential. Efforts are underway within the USGS to develop improved three-dimensional geologic maps. Cooperation between local, state, and Federal agencies; universities; private companies; and the USGS is critical to the success of these efforts. The Seattle Area Natural Hazards Project (http://Seattlehazards.usgs.gov) provides an example of one such cooperative effort.

For volcano-hazard assessments, uncertainties in hazard estimates due to uncertainties in eruption timing and volume, tephra dispersion, debris-flow dynamics, etc., tend to outweigh uncertainties caused by map resolution or accuracy. Reducing the former uncertainties requires more detailed age dating and volume estimation of past eruptions and improved models for eruption dynamics, tephra dispersion, debris flows, etc.

Landslide-hazard assessments are subject to the previously indicated uncertainties for trigger events such as earthquakes or floods, compounded by uncertain knowledge of highly variable small-scale processes such as pore pressure, shear strength, etc. Thus, in addition to improved topographic and geologic maps, landslide-hazard assessments require more detailed and accurate maps of shear strength and other engineering properties of soils and more research into the physical processes controlling landslides. The feasibility of developing detailed maps for landslide-hazard assessment has been demonstrated on a local scale (Jibson and others, 1998), but much more remains to be done to apply such detail on a regional or national scale.

Real-time hazard warning has been the focus of increasing efforts within the USGS in recent years, especially with respect to flood and volcano hazards. Because real-time hazard warning is related closely to probabilistic hazard assessments (where the probabilities are high and time scales are short), some of the USGS efforts in support of real-time hazard assessments are worthwhile mentioning. The national streamflow-gaging network is a widely recognized USGS program for flood-hazard warning and also provides much of the streamflow data used for developing probabilistic flood-hazard estimates. The earthquake- and volcano-hazards program of the USGS also provides seismic data for real-time warning of volcanic activity, and the proposed Advanced National Seismic System will provide seismic data for earthquake monitoring.

In addition to supplying information products such as maps, data, and educational publications that are used for hazard assessments and real-time hazard warnings, the USGS has many successful programs that deal directly with probabilistic hazard and risk assessments. The national seismic hazard mapping program (http://geohazards.cr.usgs.gov/eq) is one of the most widely visible of these programs. The national seismic hazard map is a comprehensive, visible, widely assessed, and useful product. The map is used on a routine basis by city and county agencies for developing building codes and mitigating risk from earthquakes. Interactive access to maps and related products is being improved continually, and products that better meet users needs are being developed.

The National Flood Frequency Program (NFFP) (http://water.usgs.gov/osw/nffp.html) provides equations and software for estimating flood-frequency distributions for ungaged basins in rural or urban areas of the United States and Puerto Rico. The frequency distributions are based on regional regression equations developed by the USGS district and regional offices and compiled by the Office of Surface Water. The equations and related documentation were discussed by Jennings and others (1994), and an effort is underway by the Office of Surface Water to update the equations, revise the computer programs, and provide interactive access to flood-frequency information.

Although the NFFP provides a nationally-coordinated program for determining flood-frequency distributions, no national program exists for converting flood frequencies to flood stages and producing flood-inundation maps for the entire United States. The panels generally agreed such an effort would be a substantial and worthwhile challenge for the next decade. The USGS is ideally suited for taking the lead in producing a national flood-inundation map. The map should be interactively available and continually updated and supported in a manner similar to that for the national earthquake-hazard map. Although a digital map with 1-foot vertical accuracy and 10-meter resolution should be the ultimate goal, a less precise map of certain river reaches could be produced first with updated versions being produced as more detail becomes available. A study in the USGS Washington District (http://wa.water.usgs.gov/reports/floodgis) is underway to produce accurate digital flood-inundation maps for the Puget Sound area. Studies such as this may lead to more extensive regional or national flood-mapping development.

A number of generic recommendations for product improvements were mentioned by all three panels. The USGS needs to continue efforts to make data more widely available on different media (internet, CD-ROM, etc.) and to make data more accessible to the public. Particular concern was shown among some panel members that USGS water-resources data were not as easy to access as data from other agencies such as the National Weather Service. The recent development of the National Water Information System (NWIS) web site for outside customers to obtain streamflow data should help alleviate some of these concerns. However, continued efforts are needed to improve digital access of data from the streamflow-gaging programs.

Another recommendation was for increased interaction between the USGS and its customers to better identify customer needs. Public opinion and customer feedback should be actively sought and continually evaluated to improve products. Scientists from the USGS need to assess the impact and value of their data, reports, and other publications and be flexible in altering format and content to meet demand.

Finally, a suggestion was made that more effort be placed toward common communication of hazards among disciplines, integration of probabilistic hazard maps for the different hazards, and calculation of risk from multiple hazards. These points will be revisited in discussion topic 4.

Discussion Topic 3--Which Research Topics are Most Important for the Future of Earthquake-, Flood-, Landslide-, and Volcano-Hazard Assessments?

Modeling Temporal and Spatial Dependence of Source Events

Temporal and spatial dependence of source events is an important consideration for estimation of hazard probabilities and development of hazard-prediction models. Thus, stochastic models for incorporating temporal and spatial dependence of source events need to be developed and tested. All of the hazard-probability estimates described earlier are based on one or more assumptions regarding independence of source events from year to year, independence of the waiting times between source events, or independence of source events among source locations. These assumptions often are born out of necessity rather than reality because temporal or spatial dependence introduces an exponential increase in the complexity of estimating the hazard probability. However, relatively recent mathematical developments such as scaling theory and nonlinear dynamics are being used to identify structure in the seemingly complex and random nature of source events.

For flood-hazard assessments, dynamic climate-simulation models are being used to simulate the effect of long-term climate variation and/or climate change on precipitation and temperature patterns on a global scale. However, more research is needed on connecting the global-scale models to basin-scale precipitation and temperature before the climate models can be used to develop dynamic flood-hazard estimates. More research also is needed on relating flood peaks to drainage-network geometry, overland- and subsurface-flow mechanisms, and stochastic properties of rainfall and temperature.

For earthquake-hazard assessments, dynamic models are being used to look for structure in the temporal and spatial occurrence of source events. However, models need to be improved or new models developed before significant improvements in hazard estimation can be gained over the nonhomogeneous Poisson models currently in wide use.

Reducing Spatial Uncertainty in Predicted Consequences of Source Events

Even if the location, magnitude, and timing of a source event is known, considerable uncertainty exists in the predicted consequences of the source event. The locations and severity of liquefaction or soil displacement resulting from a particular earthquake, the depth and extent of tephra accumulation or debris flow resulting from a volcanic eruption, and the location and timing of ice jams or breakout flows for certain floods are a few examples where considerable uncertainties come into play. Evaluation of the geophysical processes, data uncertainties, and model uncertainties affecting hazard assessment needs to continue to be a high research priority. Research on improving digital data bases for hazard assessment (for example, developing high-resolution, seamless, digital-elevation and three-dimensional geologic maps) also is needed to reduce data uncertainties and encourage more integrated hazard assessment among the various disciplines and agencies.

Development of Improved Short-Term Hazard Prediction Capability

Temporal nonstationarity or temporal dependence of source events introduces a dynamic aspect to probabilistic hazard assessment in that the probability of an event changes through time either deterministically (in the case of nonstationarity) or stochastically (in the case of temporal dependence). However, at present, the temporal and spatial resolution of stochastic models is not detailed enough to provide real-time hazard-prediction capability. Thus, for example, although probabilistic hazard assessments may show a high probability of an event occurring sometime during the next 10 years somewhere in a particular region, too much epistemic and stochastic uncertainty exists to predict exactly when or where the event will occur.

Future research is needed to develop stochastic models that can push the limits of the predictability of probabilistic hazard analysis to increasingly smaller time intervals and increasingly smaller spatial regions. For example, suppose a determination has been made using the best available stochastic techniques that a severe earthquake is likely (say, 80 percent or higher chance) somewhere along the Pacific Coast within the next 10 years. Although this information is valuable, it is not detailed enough to economically justify establishing seismic-monitoring networks and geologic field sites at every possible earthquake source location along the Pacific Coast to pinpoint exactly where the earthquake will occur and provide advance warning of the earthquake. However, as stochastic models are improved and the likely location of the earthquake is narrowed down to smaller and smaller regions, a point may occur when it is economically feasible to intensively monitor the potential source locations and provide advance warning of the event. Stochastic methods cannot replace deterministic methods for real-time hazard warning but can be used in tandem with deterministic methods to establish monitoring networks.

Discussion Topic 4--How Can Hazard and Risk Assessments Be Made More Unified Among Earthquakes, Floods, Landslides, and Volcanoes?

This is the ultimate question that brought about the workshop. Although the question is much too complex to answer in the limited timeframe available, workshop participants gave a number of valuable ideas and suggestions to provide a framework for a more unified treatment of hazard and risk assessments. Four important obstacles toward unification were identified: lack of a centralized national data base for hazard probability estimates, lack of a "location-free" measure of point-wise hazard magnitude for hydrologic hazards, lack of well-defined "regional hazard probabilities," and the highly localized and specialized nature of risk assessment. The following paragraphs indicate potential remedies for reducing or overcoming these obstacles.

Establish a National Flood Hazard Map

The USGS National Seismic Hazards Map provides a centralized national data base for evaluating seismic hazards. A comparable product needs to be developed for flood hazards before hydrologic and seismic hazards maps can be compared. Development of a national flood hazard map could be done in coordination with the National Streamflow Information Program (NSIP). The map should provide, in a GIS format, the flood stage corresponding to certain annual exceedance probabilities for any given river or stream cross section.

Develop a "Location-Free" Measure of Flood Magnitude

A key to developing the national seismic hazard map is that various location-free measures of seismic hazard magnitude exist. For example, peak horizontal ground acceleration is a universally recognized measure of ground motion that can be used to estimate potential damages given localized soil properties and infrastructure. Spatially smooth hazard maps depicting annual probability of exceedance for given peak horizontal ground acceleration can thus be produced. However, there is no comparable hazard magnitude for floods. Consequently, flood hazards maps usually show the flood plain (a contour of the inundated areas) corresponding to a given exceedance probability. There is no measure of the degree of flood magnitude at different locations within the flood plain--all locations essentially are treated as having the same “risk” for flood-insurance purposes. However, the severity of a flood at a given location is clearly dependent on much more than whether or not that location is inundated. Using inundation depth as the measure of point-wise flood magnitude may be more representative of spatially varying risk, but inundation depth is a poor indicator of flood severity. For example, a high-velocity mountain stream cresting 5 feet above flood stage might produce catastrophic damage, whereas in a low-velocity stream such a crest might produce limited damage. Much discussion and research is needed to develop standardized measures of flood magnitudes, if such measures are indeed feasible. Until such time, spatially smooth flood probability maps will be of limited use.

Develop a Common Data Base and Base Map for all Hazards

A single data base needs to be established and maintained for all hazards. The USGS hazards program is ideally suited for this task. This is a most important step that needs to be taken before significant progress can be made toward unification. The data base does not need to include all of the variables used in calculating hazard probabilities, but it should contain hazard magnitude-frequency data for each pixel of a standardized map.

Contour Probabilities, Not Magnitudes

Hazard maps obtained by fixing an annual exceedance probability and contouring the hazard magnitude are not particularly useful for risk assessors. For example, a "one-hundred-year flood" may be of little consequence at one location and may be highly damaging at another location. Maps obtained by fixing the hazard magnitude and contouring the hazard probabilities are more informative. Such maps already are being used in the earthquake-hazard mapping program and in various volcano and landslide investigations and should be produced on a routine basis in future hazard investigations. Another advantage of contouring probabilities rather than hazard magnitudes is that probabilities are dimensionless and, hence, easier to compare among different hazards.

Quantify Regional Versus Point-Wise Risk

Point-wise hazard probabilities are useful on a local scale for establishing building codes, evaluating economic benefits of local risk-mitigation alternatives, etc. They also can be used to calculate expected damages for a region because the expected damage (annualized cost) for a region is the sum of the expected damages over all the individual locations within the region. However, when considering regional hazard-response and risk-mitigation measures, expected damages are of limited use. For example, the expected annualized cost of earthquake damage in California would be the same if an average of one $5 million event occurred every year or if an average of one $500 million event occurred every 100 years. However, there would be a huge difference in the emergency response measures required for the two situations.

Consideration of true regional risk requires much more than point-wise hazard probabilities. The great flood of 1993 in the upper Mississippi River Basin provides a prime example. In that year, peak flows with annual exceedance probabilities less than 1 percent occurred at 7 of the 17 major tributaries to the upper Mississippi River that had streamflow gages (Parrett and others, 1993). If flood peaks at individual tributaries were statistically independent, the annual probability of a similar flood occurring in the future would be astronomically small (10^{-14}) and would be of no practical concern. However, because of spatial dependence of flood peaks, the likelihood of a regional flood such as in 1993 is not astronomically small--in fact, estimates place it as high as a 1-in-200 year event (considering possible climate nonstationarity, the likelihood could even be much higher).

Calculation of regional risk also is important for hazards other than floods. For example, consider estimating the regional earthquake risk for a large metropolitan area (analogous to the "river basin" in the previous example) on the basis of point-wise estimates of earthquake hazard for smaller, disjoint subregions (analogous to the "tributaries" in the previous example). Although the occurrence of earthquakes (source events) in disjoint subregions may be independent, the peak horizontal acceleration (the hazard) is not. The risk from a regional standpoint thus depends on the spatial distribution of peak horizontal acceleration. The key point is that, because all hazards are complex spatial processes, fully evaluating regional risk requires a measure of spatial extent and severity of the hazard that cannot be computed from point-wise hazard probabilities.

Routinely providing joint estimates of hazard probabilities (i.e., including spatial dependence) rather than point-wise estimates (i.e., ignoring spatial dependence) is not feasible due to the extreme data requirements of such an approach. However, it may be feasible to allow interactive input by risk assessors so that regional risk can be calculated for any given application. For example, suppose an earthquake risk assessment is being done for Los Angeles County, Calif. Several variables, such as number of deaths, dollar damages to infrastructure, number of hospital beds required to treat casualties, etc., may be of interest. If the county emergency management agency could provide the estimated value of each variable for a range of values of earthquake intensity (such as peak horizontal acceleration), and for each section (or other geographic unit), it would be relatively straightforward to produce a probability distribution for the total value of each variable over the entire county. The probability distribution for each variable would provide much more information than just the expected value.

References

Bernknopf, R.L., Brookshire, D.S., Soller, D.R., McKee, M.J., Sutter, J.F., Matti, J.C., and Campbell, R.H., 1993, Societal value of geologic maps: U.S. Geological Survey Circular 1111, 53 p.

Jennings, M.E., Thomas, W.O., Jr., and Riggs, H.C, comp., 1994, Nationwide summary of U.S. Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993: U.S. Geological Survey Water-Resources Investigations Report 94-4002, 196 p.

Jibson, R.W., Harp, E.L., and Michael, J.A., 1998, A method for producing digital probabilistic seismic landslide hazard maps--An example from the Los Angeles, California, area: U.S. Geological Survey Open-File Report 98-113, 17 p.

Michael, A.J., Ross, S.L., Schwartz, D.P., Hendley, J.W., II, and Stauffer, P.H., 1999, Major quake likely to strike between 2000 and 2030: U.S. Geological Survey Fact Sheet-152-99, 4 p.

Newmark, N.M., 1965, Effects of earthquakes on dams and embankments: Geotechnique, v. 15, no. 2, p. 139-160.

Parrett, Charles, Melcher, N.B., and James, R.W., Jr., 1993, Flood discharges in the upper Mississippi River Basin, 1993: U.S. Geological Survey Circular 1120-A, 14 p.

Stedinger, J.R., and Tasker, G.D., 1985, Regional hydrologic analysis--Ordinary, weighted, and generalized least squares compared: Water Resources Research, v. 21, no. 9, p. 1421-1432.

Suzuki, T., 1983, A theoretical model for the dispersion of tephra, *in* Shimozuru, D., and Yokoyama, I., eds., Arc volcanism: Physics and Tectonics, Terra Scientific Publishing, Tokyo, p. 95-113.

United States Water Resources Council, 1981, Guidelines for determining flood flow frequency: Bulletin 17B of the Hydrology Committee, 28 p.

APPENDIX A--Technical Aspects of Hazard Assessments

True integration of hazard assessments among the various natural hazards requires a general mathematical framework for evaluation of hazard probabilities. This appendix describes some of the similarities and differences that exist among the currently accepted methods for probabilistic hazard assessment. Previous sections of this report were intended for a general audience and were not targeted specifically for experienced probabilists. This appendix is geared more toward those interested in the probability theory of hazard assessments.

The complicated temporal and spatial structure of a source event must be simplified before the probability can be estimated with any reliability. The first step in the simplification usually involves expressing the probability of a source event in one of two forms:

$$P(m, x, t) = f(m, x)g(t) \tag{A1}$$

or

$$P(m, x, t) = f(m)g(x, t) \tag{A2}$$

where

$P(m, x, t)$ is the probability of a magnitude m^+ (m or greater) event at location x in year t,

x is a particular location in a two-dimensional region,

f is a magnitude-frequency distribution that may or may not depend on location, and

g is a temporal occurrence rate function that may or may not depend on location.

Although more complicated models can be constructed, these two are the most commonly-used and analytically-tractable models. The first model usually is assumed for floods and the second for earthquakes or volcanoes. Model 1 usually is preferable when events occur relatively often (e.g., floods) and the time scales considered are relatively short (e.g., decades), and model 2 usually is preferable when events occur relatively infrequently (e.g., major volcanic eruptions or earthquakes) and the time scales considered are relatively long (e.g., centuries or much longer).

Both models require further simplification because data may not be adequate to estimate the probabilities for all locations and all times of interest. For model 1, the simplification generally proceeds as follows. Averaging $P(m, x, t)$ over a T-year time interval gives

$$P^*(m, x, T) = f(m, x)g^*(T)$$

where

$g^*(T)$ is the average of $g(t)$ over the T-year interval.

Without loss of generality, $g^*(T)$ can be set equal to one, in which case $f(m, x)$ is interpreted as the probability of an event of magnitude m^+ occurring at location x during a randomly chosen year. The resulting magnitude-frequency distribution can be estimated by pooling data from all T years at location x and using standard statistical estimation techniques.

To estimate $g(t)$, the probability, P^*, usually is fixed at a particular value depending on the application being considered. For example, for determining flood-plain delineations, P^* might be set equal to 0.01 (a "one-hundred-year return period"), and for design of high-risk structures, P^* might be set equal to 0.001. Given a value for P^*, the quantile function can be defined as

$$Q(P^*, x) = f^{-1}(P^*)$$

where

$Q(P^*, x)$ is interpreted as the magnitude that has a probability P^* of being exceeded at location x in a randomly selected year.

To estimate $g(t)$, note that

$$P[Q(P^*, x), x, t] = P^* g(t).$$

Thus, an unbiased estimate of $g(t)$ can be obtained by dividing the proportion (over all locations) of events in year t that exceed $Q(P^*, x)$ and dividing by P^*. If the estimated values of $g(t)$ do not change significantly from year to year (as is often the case), temporal stationarity is indicated and events occur in time according to a Poisson process and have an annual exceedance probability P^*. Otherwise, a nonstationary Poisson process is indicated.

For model 2, the simplification generally proceeds as follows. Averaging equation A2 over a T-year time interval and R locations gives

$$P^*(m, R, T) = f(m) g^*(R, T)$$

where, as before, g^* can be assumed to equal one, in which case $f(m)$ is interpreted as the probability of an event of magnitude m^+ occurring at a randomly chosen location in a randomly chosen year. The dependence on m usually is expressed as an empirical scaling law, such as the Gutenberg-Richter law for earthquake magnitudes, and unknown parameters are estimated by pooling data from a large number of locations and years.

Estimation of $g(x, t)$ for model 2 depends on whether a spatially discrete or spatially continuous process is being modeled. In a spatially discrete process, events occur at a finite set of specified locations. For example, volcanic eruptions from known volcanoes or earthquakes from known epicenter locations may be modeled as a discrete process. In a spatially continuous process, events occur anywhere in a continuous region. Earthquakes from source zones constitute a spatially continuous process. In a discrete process, $g(x, t)$ can be estimated separately for each fixed location given a long record of past events. For example, if the times between successive events are independent, identically-distributed, exponential random variables, the occurrences of events follow a temporally stationary Poisson process and have a nonhomogeneous annual exceedance probability equal to $f(m)g(x)$. In that case, $g(x, t)$ is the same for every t and is estimated from the inter-event times at location x. If the times between successive events are not exponentially distributed (as is often the case), another distribution, such as a Weibull or mixed exponential distribution, may be more appropriate. In that case, $g(x, t)$ depends on t because the probability of an event changes depending on the time elapsed since the most recent event. However, computation of $g(x, t)$ is straightforward given the inter-event time distribution and the time elapsed since the most recent event.

Estimation of $g(x, t)$ for a spatially continuous process is not as straightforward because $P(m, x, t)$ is equal to zero at any fixed point x. Thus, $P(m, x, t)$ must be defined with respect to a small neighborhood about location x. To obtain a reasonable number of historical events to estimate P, a large neighborhood may be necessary and $g(x, t)$ would not necessarily be constant over the neighborhood. Therefore, regionalization methods described in the next few paragraphs are necessary to obtain a reasonable estimate of $g(x, t)$.

Hazard assessments often require estimates of exceedance probabilities at locations for which little or no historical data exist. In that case, regionalization methods are required to express the spatial dependence of $f(m, x)$ in model 1 or $g(x, t)$ in model 2 as a smooth function of known physical properties or location. First, consider regionalization of $f(m, x)$ in model 1. The approach usually used is to fix P^* (the annual occurrence probability) and regionalize the quantile function, $Q(P^*, x)$. For example, in the flood-regionalization methods described by Troutman (appendix C), the "source event" is an annual peak discharge exceeding a given flood discharge at location x (a river cross section) as opposed to a rainfall event of a certain magnitude. In that case, $Q(P^*, x)$ is expressed in terms of a regression equation with known drainage-network properties, such as drainage area, as independent variables.

Next, consider regionalization for model 2. The model usually is applied to a particular spatial region, R, consisting of one or more seismic source zones or volcanic fields:

$$P(m, x, t) = f(m) g(x, t) \text{ for } x \text{ in } R.$$

The magnitude-frequency distribution is interpreted as the probability of an event of magnitude m or greater occurring during a randomly selected year in a randomly chosen neighborhood of unit area (such as a circle) somewhere in R. Given a reliable regional estimate of $f(m)$, the focus then is on estimating the temporal and spatial structure of $g(x, t)$. Because events generally cluster in both space and time, $g(x, t)$ cannot be held constant over the whole of region R or over any substantially long time interval. However, $g(x, t)$ must satisfy certain consistency properties to calculate the probability of an event occurring *somewhere* in any subregion of R (consisting of a union of many small neighborhoods) *sometime* during a specified time interval. The most analytically-tractable approach (and perhaps the only analytically-tractable approach at this time) is to assume events occur according to a nonhomogeneous

Poisson process and have a rate function $g(x, t)$ that doesn't depend on magnitude (magnitude 5 earthquakes, for example, may cluster differently from magnitude 7 earthquakes, but magnitude-dependent clustering would require a much more complicated model). The nonhomogenous Poisson model supposes that the occurrences of events in disjoint spatial regions or disjoint temporal intervals are independent and that the probability of exactly n events in region X during time interval T is

$$\text{Prob}[N(X, T) = n] = \exp\{-g^*(X, T)\}g^*(X, T)^n/n! \quad \text{(A3)}$$

where

$N(X, T)$ is the number of events in spatial region X and temporal interval T, and

$g^*(X, T)$ is the integral of $g(x, t)$ over region X and interval T.

Various nonparametric smoothing methods have been used to estimate $g(x, t)$. These methods depend on a smoothing function called a kernel. For example, the national earthquake hazard maps use Gaussian smoothing to obtain regionally smooth estimates of $g(x, t)$ (Perkins, appendix C). Smoothing techniques invariably depend on a smoothing parameter that controls the degree of smoothness of the resulting estimate. Selection of the optimal smoothing parameter is one area of ongoing research.

After an adequate model has been developed for estimating the probability of a source event at any given location, an overall hazard probability is computed for location x given potential source events at all locations that may affect the hazard at location x. At this point, a diversion occurs between the approach used for floods and the approach used for earthquakes, volcanoes, and landslides. Because floods are spatially continuous processes, the spatial Poisson assumption (independence between events in disjoint regions) clearly is not appropriate for a flood-hazard assessment. A spatial Poisson model is a point process--although the set of possible source locations is continuous, in any given time interval and any given region, only a finite number of events can occur. However, rainfall may exceed a critical level for producing floods at infinitely many spatial locations during any time interval, no matter how small, and in any spatial region, no matter how small. Therefore, because of the complex structure of precipitation processes, a flood-hazard assessment rarely is done by starting with a probability model for "source events" (i.e., precipitation events) and then estimating the flood hazard on the basis of the source events. [Precipitation-runoff models usually are used to help evaluate the flood hazard for extremely unlikely floods (such as floods with an annual exceedance probability less than 0.001).]

In a flood-hazard assessment, a "source event of magnitude m" at location x in a drainage network (where x now denotes a particular cross section of a river channel) is the event that the annual maximum discharge at x exceeds m. The probability of a source event in year t is defined as

$$P(m, x, t) = \text{Prob}[Y(x, t) > m]$$

where

$Y(x, t)$ is the maximum instantaneous flow at location x in year t (in units of volume/time, such as cubic feet per second).

Temporal nonstationarity (for example, due to climate change) may be approximated by assuming $Y(x, t)$ has the same distribution for all t except for a temporal scaling factor,

$$\text{Prob}[s(t)Y(x, t) > m] = f(m, x),$$

where

$s(t)$ is the temporal scaling factor, which may be a deterministic function or a stochastic process depending on the application.

The flood probability then can be expressed as

$$P(m, x, t) = f(s(t)m, x) = f(m, x)\left[\frac{f(s(t)m, x)}{f(m, x)}\right].$$

Comparing this equation with model 1 (eq. A1) shows the two are equivalent if the ratio $\left[\frac{f(s(t)m, x)}{f(m, x)}\right]$ depends only on $s(t)$ and not on x or m. Thus, model 1 places rather severe constraints on the probability distributions $f(m, x)$. However,

if $g(t)$ in model 1 is allowed to depend on m, model 1 places a much less severe constraint on the distributions $f(m, x)$, requiring only that the tail behavior of $f(m, x)$ is the same for every location.

The flood hazard in year t at location x usually is expressed as the probability of exceeding a certain stage (water level) at location x sometime in year t:

$$H(w, x, t) = \text{Prob}[W(x, t) > w]$$

where

$W(x, t)$ is the maximum water level in year t at location x, and

w is a specified water level.

In most investigations, discharge, $Y(x, t)$, and stage, $W(x, t)$, have a one-to-one relation called a discharge-stage rating curve (the curve can change through time in response to changing channel morphology) in which case the hazard H can be calculated directly from $P(m, x, t)$. However, for very high discharges, considerable uncertainty may exist in the rating curve, and hydraulic flow-routing models may be required to estimate the stage corresponding to those discharges. Hydraulic flow-routing models also are required to estimate the effects of drainage modifications such as dams or levees. Uncertainty in the estimated stage when using hydraulic flow-routing models generally is small and can be ignored in most investigations. However, in some cases involving floods for very flat-slope, low-velocity channels, floods in coastal regions, or floods affected by ice, the same peak discharge may produce significantly different peak stages depending on downstream conditions. No method is universally accepted for dealing with such situations in calculating the flood hazard. Sometimes a single stage is selected to represent the "expected" stage for a given peak discharge, and other times a conservative stage is selected to represent the highest stage that could be attained for a given peak discharge. In any case, expert judgment is required to estimate a reasonable discharge-stage rating curve.

In earthquake-, volcano-, and landslide-hazard assessments, the hazard at location x (now a point in a two-dimensional region) is defined as the probability of exceeding a given value of a particular quantitative measure of hazard intensity given potential source events from neighboring locations. The quantitative measure of hazard intensity depends on the application--the measure might be peak horizontal ground acceleration for earthquakes, depth of tephra accumulation for volcanoes, or displacement for earthquake-triggered landslides. For a generic year, let $Z(x|y)$ be the value of a hazard measure (such as peak horizontal acceleration) at location x given a source event (such as an earthquake) at location y, and let

$$Z(x) = \max\{Z(x|y);\ y \text{ in } R\}$$

where

R is a region of influence surrounding x beyond which source events are assumed to have negligible effect.

Source events in disjoint subsets of R are assumed to be independent. Therefore, R can be partitioned into small subsets or pixels, and the hazard in any given year can be expressed as follows:

$$H(z, x) = \text{Prob}[Z(x) > z] = 1 - \text{Prob}[Z(x) \le z] = 1 - \prod_{i=1}^{n} \{\text{Prob}[Z_i(x) \le z]\} \qquad \text{(A4)}$$

where

$z > 0$ is a particular value of hazard intensity (such as 0.5 g),

$R = \bigcup_{i=1}^{n} R_i$ is a partition of R into n pixels, and

$Z_i(x)$ is the maximum of $Z(x|y)$ over the ith pixel.

The pixels are chosen small enough so that the probability of more than one source event in any given pixel can be ignored, in which case $Z_i(x)$ is either zero (if no source events occur in R_i) or greater than zero (if exactly one source event occurs in R_i). Therefore,

$$\text{Prob}[Z_i(x) \le z] = \text{Prob}[N_i = 0] + \text{Prob}[Z_i(x) \le z | N_i = 1]\text{Prob}[N_i = 1]$$

where

N_i is the number of source events occurring in R_i in any given year.

Using the nonhomogeneous Poisson model (eq. A3) for the distribution of N_i and noting the dependence on the year, t, yields

$$\text{Prob}[Z_i(x) \le z] = \exp\{-g^*(R_i, t)\}[1 + p_i(z, x)g^*(R_i, t)]$$

where

$$p_i(z, x) = \text{Prob}[Z_i(x) < z | N_i = 1] .$$

Because $p_i(z, x)g^*(R_i, t)$ is much smaller than one, a further approximation gives

$$\text{Prob}[Z_i(x) < z] = \exp\{-g^*(R_i, t)[1 - p_i(z, x)]\} . \quad \text{(A5)}$$

The probability contribution for the ith pixel, $p_i(z, x)$, can be calculated given the magnitude-frequency distribution of source events and the attenuation relation between a source event in the ith pixel and the hazard intensity at x. For earthquakes, the attenuation relation is based on an empirical equation relating ground motion at location x to a seismic event of a given magnitude at a given distance from x, taking into account geologic properties of the intervening area. For tephra accumulation from volcanic eruptions or soil movement from earthquake-induced landslides, the attenuation relation may be computed using techniques such as those described by Connor and others, Jibson, and Roberds (appendix C).

The hazard at location x in year t is computed by combining equations A4 and A5 to obtain

$$H(z, x, t) = 1 - \exp\left\{- \sum_{i=1}^{n} g^*(R_i, t)[1 - p_i(z, x)]\right\} .$$

APPENDIX B--Abstracts of Presentations

FLOOD HAZARD ASSESSMENT IN ENGINEERING PRACTICE

Allen Bradley, University of Iowa, Iowa City, Iowa

Flood hazard assessment using probabilistic methods is the basis for many planning and design decisions, including the management of land in flood-prone areas (e.g., the 100-year flood plain) and the design of hydraulic structures (e.g., levees). The concepts used in probability-based flood design are based on traditional probabilistic models of floods, which assume that floods are generated by a random stationary process. This presentation describes the probability models commonly used to describe flood magnitudes. Although there has been considerable debate in the literature over the choice of probability distributions and the estimation of parameters from flow records, a standard statistical method for flood quantile estimation routinely is used for engineering design within the United States. Still, the majority of problems of practical importance require estimates of the changes in flood quantiles associated with land-use modifications (urbanization) or flow regulation (by flood-control reservoirs). Streamflow measurements are seldom available at these sites for statistical analysis; even if they were, nonstationarity due to changes in the underlying flood distribution would make past observations inconsistent with future conditions (for which the predictions are needed). This presentation also describes some of the common rainfall-runoff based approaches used for flood-hazard assessment and approaches for estimating the upper bound to flood discharges needed for the design of high-hazard structures.

CHARACTERISTICS OF FLOODS AS A SPATIAL PROCESS

Brent Troutman, U.S. Geological Survey, Lakewood, Colo.

River flow extremes constitute a complex random process in both space and time. In this presentation, some basic characteristics of flood flows considered as a spatial process are reviewed and discussed. Of particular interest are both marginal and joint distributional properties of annual peak flows at a collection of points in some large region. Two widely used regionalization methods for modeling the dependence of the marginal distribution of annual flood peaks on drainage area--the index-flood method and regional quantile regression--are compared. In addition, problems associated with characterizing spatial correlation of flows are discussed, and generalized least squares as used by the USGS to account for spatial correlation in quantile regression is briefly introduced. Finally the problem of estimating regional flood probabilities, with the goal of assessing frequency of occurrence of flow extremes somewhere in a large region, is discussed. Examples are presented showing how storm rainfall and network properties affect this frequency and how this frequency may be estimated using annual peak-flow data.

LOW-FREQUENCY CLIMATE VARIABILITY AND CHANGING FLOOD RISK

Upmanu Lall, Utah State University, Logan, Utah

Flood-frequency analysis, as traditionally practiced, is marked by an assumption that annual maximum floods conform to a stationary, independent, identically-distributed, random process. These assumptions are at odds with the growing recognition that low-frequency climate variations such as the El Nino Southern Oscillation systematically change the probabilities of seasonal and extreme daily rainfall and, hence, of floods in different regions of the world. Ongoing work on the diagnosis of such changes in flood series will be presented, focusing on two case studies. The implications of such analyses will be discussed in the context of (a) interpretation of the historical flood record, (b) situations in which it may be useful to estimate conditional flood risk, and (c) possible statistical methods for estimating flood risk.

LANDSLIDE RISK ASSESSMENT

William Roberds, Golder Associates, Redmond, Wash.

Landslides can occur in different forms, ranging from individual rock falls to large creep failures, depending on site conditions (e.g., topography, geologic structure, shear strength, pore pressures, and loads), some of which in turn are affected by various processes and events (e.g., earthquakes, precipitation, erosion, or excavation). Such landslides can have significant consequences (e.g., casualties, property damage, and socioeconomic impacts such as loss of service), depending on: (a) their characteristics (e.g., number, timing, location, size, mobility, runout, etc.), which constitute the "hazard"; and (b) the "vulnerability" of people, structures and infrastructure to those landslide characteristics, which in turn is a function of the population and development in the potentially affected area and their

characteristics (e.g., location, occupancy, awareness, mobility, damage susceptibility, etc.). Such undesirable consequences can be prevented or at least reduced through various actions designed to: (a) prevent landslides from occurring (e.g., by installing drains or support, scaling or flattening the slope); (b) reduce their severity if they do occur (e.g., by installing debris barriers); and/or (c) reduce the vulnerability of potentially affected people/structures/infrastructure (e.g., controlling development, warning/evacuating people, strengthening structures, maintaining contingency plans). However, some of these actions can be very expensive to implement (in financial terms as well as in worker safety and socioeconomic terms) and may prove to be either ineffective or unnecessary. The objective for any site or, even more importantly, a set of sites should be to identify and implement those actions (if any) which are most cost-effective, appropriately trading off: (a) the various "costs" of implementing the actions; (b) the "benefits" of reduced consequences of landslides; and (c) the change in value (e.g., for development) of potentially affected land. Such tradeoffs, and thus preferences amongst the alternatives, obviously may vary amongst the different stakeholders, e.g., with regulators focusing on risk reduction and developers focusing on costs and land values.

However, for any action (including no action), the occurrence of landslides at a site and the characteristics and consequences if they do occur cannot generally be predicted with certainty beforehand, due to inherent uncertainties in: (a) the site conditions/processes/events and in the vulnerability of people/structures/infrastructure, a combination of stochastic variability (temporal and/or spatial, which is natural and not reducible) and ignorance (which is reducible through more information); and (b) the relationships of instability, landslide characteristics and consequences to the site conditions/processes/events and the vulnerability of people/structures/infrastructure, a combination of simplifications and approximations.

The uncertainties in whether one or more landslides will occur at a site, and in the consequences if they do occur ("risks"), need to be adequately considered when deciding on possible actions. In fact, in some jurisdictions, acceptable levels of risk are specified, and compliance must be adequately demonstrated. Such risks can be assessed in many ways, differing widely in effort/cost, accuracy and defensibility. Fundamentally, each method is a combination of the following attributes: (a) site-specific or averaged over a set of sites; (b) qualitative or quantitative; and (c) comparative or absolute. Although qualitative and/or comparative methods may be adequate for ranking or screening slopes and possible actions for slopes, a site-specific, absolute, quantitative method is generally needed for making optimal, defensible decisions on significant actions. However, such quantitative methods can vary significantly in level of detail, ranging from: (a) direct assessments of the frequency of landslides and their consequences; to (b) detailed models of various modes of instability, runout and consequences which in turn rely on direct assessments of specific input parameters. Direct assessments (at any level) can be: (a) statistically driven (if a representative data set is available); (b) based on judgment (consistent with available data); or (c) a combination of the two. Similarly, the models can be: (a) empirically-based (which requires substantial data from similar situations relating outcomes to observable factors, e.g., geologic units); (b) theoretically-based (which requires the assessment of engineering parameters, e.g., peak pore pressure); or (c) a combination of the two. Care must be taken in assessing parameter values to ensure that: (a) the appropriate characteristic of that parameter (e.g., random value, average value or extreme value over a particular area or time period) is being assessed (i.e., consistent with how it will be used); (b) the assessments are consistent with all available data; and (c) any correlations (e.g., in one parameter at different times or locations, or between different parameters at the same time and location) are appropriately considered.

The appropriate method to use for risk assessment on a particular project depends on a variety of factors, primarily the "value" of increased accuracy in making decisions and increased defensibility in justifying those decisions. Other factors include: (a) the availability of relevant engineering and empirical data, and the cost associated with acquiring additional data; and (b) the availability of tools and expertise (e.g., regarding stability, runout and damage models, and objective and subjective probability assessments), and the cost associated with better tools and expertise.

STOCHASTIC LANDSLIDE FORECASTING MODELS: CASE STUDIES IN POLICY ANALYSIS

R. L. Bernknopf, U.S. Geological Survey and Stanford University, Menlo Park, Calif.

Policy initiatives to reduce landslide losses are analyzed in a decision framework where benefits and costs are identified and evaluated from society's perspective to promote public safety. In addition, there is a second objective to ensure that policies are implemented in an efficient manner. Cost-benefit analysis can help decide the economic value of a mitigation project. Three landslide-hazard models have been developed that utilize binary-choice statistical regression to make predictions of the probability of a physical-state change. The models are based on Earth Science Information (ESI) for (1) construction-induced landslides in Cincinnati, Ohio; (2) earthquake-triggered landslides in Santa Cruz, Calif.; and (3) rainfall-triggered debris flows in Oakland, Calif.

In the Cincinnati case, mitigation rules based on the Uniform Building Code were evaluated that are spatially selective. In the second case, landslides triggered by the 1989 Loma Prieta earthquake caused significant losses. By examining different structures of information for estimating the probability of a landslide, the economic impact of two styles of ESI for earthquake-hazard mitigation is

compared. In the third case, a space-time probability model is applied for rainfall-triggered debris-flow events. The model is demonstrated in the Oakland-Berkeley hills in the San Francisco Bay Region. This model provides the basis for hazard announcements and planning emergency response for a future hazardous event.

The utility associated with the ESI is to forecast the probability of the occurrence of landslides and to assist in identifying the "best" mitigation strategy a, where $a \in A$ mitigation options. The payoff $u_s(a, \theta)$ is based on the environment s, where $s \in S$ states of the environment, the action a, and θ which is an index of informational uncertainty, where $\theta \in \Theta$ information structures. In θ, $p_{s(k)\gamma_\theta}$ is the conditional probability of a landslide in location k, $k \in K$ locations, predicted by hazard assessment $\gamma = \gamma(g, h, t)$, where $g \in G$ geologic, $h \in H$ hydrologic, $t \in T$ topographic attributes, respectively. The conditional probability of hazard is combined with values at risk and applied in a policy analysis aimed at optimizing the allocation of resources.

A MODELING PROCEDURE TO PRODUCE PROBABILISTIC SEISMIC LANDSLIDE HAZARD MAPS

Randall W. Jibson, U.S. Geological Survey, Golden, Colo.

Analysis of the distribution of the more than 11,000 landslides triggered by the 1994 Northridge, Calif., earthquake has facilitated developing modeling procedures that estimate the probability of slope failure as a function of dynamic slope stability (critical acceleration) and earthquake-shaking intensity. Combining data sets that describe the geology, topography, ground shaking, and slope-material properties in a dynamic slope-stability model based on Newmark's permanent-deformation (sliding-block) analysis yields estimates of coseismic landslide displacement from the Northridge earthquake in each 10-meter grid cell. The modeled displacements then are compared with the digital inventory of landslides triggered by the Northridge earthquake to construct a probability curve that estimates probability of failure $[P(f)]$ as a function of Newmark displacement $(D_n, \text{in cm})$:

$$P(f) = 0.335[1 - \exp(-0.048 D_n^{1.565})] .$$

This equation was calibrated using data from six 7-1/2-minute quadrangles near the 1994 epicenter. Once calibrated, the probability function can be applied to estimate the spatial variability in failure probability in any ground-shaking conditions of interest. The resulting digital hazard maps can be updated and revised with additional data that become available, and custom maps that model any ground-shaking conditions of interest can be produced when needed.

This method uses a deterministic, physical model of slope failure to make empirically-calibrated estimates of failure probability. For the model to be rigorously probabilistic, the variability and uncertainty of the model parameters must be quantified. Gathering the types of data that will be necessary to make this quantification will be a huge challenge and should be a primary focus for the future.

U.S. GEOLOGICAL SURVEY NATIONAL SEISMIC-HAZARD MAPS AND ASSOCIATED PRODUCTS

A. D. Frankel, C. S. Mueller, E. V. Leyendecker, S. C. Harmsen, R. L. Wesson, T. P. Barnhard, F. W. Klein, D. M. Perkins, N. C. Dickman, S. L. Hanson, and M. G. Hopper, U.S. Geological Survey, Golden, Colo.

The U.S. Geological Survey recently completed new probabilistic seismic-hazard maps for the United States. These hazard maps form the basis of the probabilistic component of the design maps used in the 1997 edition of the NEHRP Recommended Provisions for Seismic Regulations for New Buildings prepared by the Building Seismic Safety Council and published by the Federal Emergency Management Agency. The maps depict peak horizontal ground acceleration and spectral response at 0.2, 0.3, and 1.0 sec periods, with 10 percent, 5 percent, and 2 percent probabilities of exceedance in 50 years, corresponding to return times of about 500, 1,000, and 2,500 years, respectively. We outline the methodology used to construct the hazard maps. There are three basic components to the maps. First, we use spatially-smoothed historic seismicity as one portion of the hazard calculation. In this model, we apply the general observation that moderate and large earthquakes tend to occur near areas of previous small or moderate events, with some notable exceptions. Second, we consider large background source zones based on broad geologic criteria to quantify hazard in areas with little or no historic seismicity, but with the potential for generating large events. Third, we include the hazard from specific fault sources. We use about 450 faults in the western United States (WUS) and derive recurrence times from either geologic slip rates or the dating of prehistoric earthquakes from trenching of faults or other paleoseismic methods. Recurrence estimates for large earthquakes in New Madrid and Charleston, S. C., were taken from recent paleoliquefaction studies. We used logic trees to incorporate different seismicity models, fault recurrence models, Cascadia great earthquake scenarios, and ground-motion attenuation relations. We present deaggregation plots showing the contribution to hazard at selected cities from potential earthquakes with various magnitudes and distances. In addition to paper and electronic versions of the hazard maps, we have produced CD-ROM's with programs to look up

seismic-hazard and seismic-design values by latitude and longitude or zip code. Our website (http://geohazards.cr.usgs.gov/eq) contains a comprehensive set of seismic-hazard related products, including seismic-hazard maps, earthquake catalogs, fault-parameter table, GIS export files, hazard look up by zip code, gridded hazard values, spatial deaggregation of hazard for 100 cities, and interactive hazard maps.

STATE-OF-THE-ART-AND-SCIENCE IN PROBABILISTIC VOLCANIC-HAZARD ASSESSMENT

Charles Connor, Southwest Research Institute, San Antonio, Tex.

The last 15 years has seen a dramatic improvement in forecasting of volcanic eruptions, in large part driven by the efforts of U.S. Geological Survey scientists. Volcanologists have developed a comprehensive understanding of the nature of volcanic phenomena and have used this understanding to mitigate hazards. For the first time, physical and empirical models have been introduced as tools in volcanic-hazard assessments, giving volcanologists the potential to quantify eruption hazards using measurable parameters. Better approaches to hazard mapping and the communication of hazards are emerging. These improvements mean that, in the United States, volcanic eruptions are less likely to produce unexpected and disastrous results than they were in the past.

Nevertheless, the demands society places on volcanic-hazards assessments are increasing at least as fast, as evinced by population growth around volcanoes and the construction of facilities that require extremely low geologic risk. Like many natural phenomena, volcanic eruptions are difficult to evaluate in terms of the relative significance: the probability of volcanic eruptions is often very low compared to the scale of human experience, and the consequences of volcanic eruptions are often comparatively high. Furthermore, volcanic eruptions are complex phenomena, producing many and varied hazards: ranging from long-term gas emissions to devastating volcanic-debris avalanches and lahars. Unlike many natural hazards, volcanic eruptions are very often preceded by periods of unrest, often lasting months or years. These volcano crises require flexible hazards and risk assessments.

Uncertainty exists in forecasts of eruption timing, magnitude, and consequences. It is critical to express this uncertainty in probabilistic volcanic-hazard assessments. There are numerous approaches to illustrating uncertainty, but two which are particularly successful are hazard curves, which show the likelihood a given event will exceed a given magnitude, and hazard maps, which contour hazards. Probabilistic volcanic-hazards assessments utilize data from geologic studies, numerical simulations, and monitoring activities. For example, a tephra-hazard curve for the city of Leon, Nicaragua, illustrates the probability of tephra accumulation to a given thickness, facilitating disaster planning. Hazards can be represented as annual probability, conditional probability, or screening-distance value. Annual probabilities include the likelihood of a volcanic eruption within a given year and the range of consequences of volcanic eruptions, such as the amount of tephra that might accumulate in a particular area. Conditional probabilities assume, for instance, that the likelihood of a volcanic eruption is unity. Thus, the probable consequences of an eruption may be evaluated independently and cast as a probability tree. Screening distance values are worst-case scenarios that bound the upper limit of potential consequences of volcanic activity based on reasonable, conservative assumptions.

PROBABILITIES OF VOLCANIC ERUPTIONS AND APPLICATIONS INVOLVING RECENT U.S. GEOLOGICAL SURVEY HAZARDS ASSESSMENTS

Manuel Nathenson, U.S. Geological Survey, Menlo Park, Calif.

An underlying assumption of U.S. Geological Survey hazards assessments for the Cascades is that the probability of volcanic eruptions may be treated as a Poisson process. Time histories for some volcanoes match this assumption well. In order to calculate an annual or 30-year probability for an eruption, the relation for the recurrence rate is used. For a Poisson process, this relation is obtained from the exponential distribution for the probability that an eruption will occur in a time T less than or equal to the interval time t:

$$P\{T \le t\} = 1 - e^{-\mu t} \approx \mu t, \text{ for } \mu t \text{ small,}$$

where

μ is the mean occurrence rate of events per year.

Since occurrence rates are small in the Cascades, the approximate relation shown is normally used. For the case of lava flows from isolated vents covering an area a in a volcanic field of total area A, a factor $p = \frac{a}{A}$ can be factored in as μtp to account for the

probability of areal coverage. This analysis assumes that the occurrence of vents are homogeneous in space within the defined area of the volcanic field.

The properties of a Poisson process include the characteristics that the conditional probability of waiting till an eruption occurs does not depend on the time that we have already waited only on the time that is in the future. For some volcanoes, the time history contains disparate time intervals between eruptions, with some being short and others being much longer. In this case, other probability distributions are a more accurate representation of the data, and the conditional probability for these distributions does depend on the time since the last eruption. The Weibull distribution introduced by Bebbington and Lai (1996) has mixed results in dealing with these disparate intervals. An alternate distribution is the mixed exponential

$$P\{T \le t\} = 1 - p_1 e^{-\mu_1 t} - p_2 e^{-\mu_2 t}$$

where

p_1 is the fraction of short intervals,

μ_1 is the average occurrence rate for the short intervals, and

p_2 and μ_2 are the same parameters for the long intervals.

The basic notion embodied in this relation is that there are two states, one involving short intervals and a second involving long intervals. The probability of an eruption occurring in each of these states is governed by an exponential distribution. The mixed-exponential distribution appears to match the available data reasonably well and resolves a conceptual problem for volcanoes with disparate eruption-time intervals.

Some examples of time histories with disparate eruption-time intervals are Mount Rainier, Mount St. Helens, and Medicine Lake volcano. The Weibull and mixed-exponential distributions agree much better with the Medicine Lake data set than does the exponential distribution. The conditional probability that there will be an eruption in the next 30 years is 2.5 percent for the Weibull and mixed-exponential distributions compared to the estimate using the exponential distribution of 3.7 percent. Estimates at other times since the last eruption differ from that for the exponential by larger factors.

ROBUST ESTIMATES OF DEBRIS-FLOW HAZARDS ON VOLCANOES

Roger P. Denlinger, U.S. Geological Survey, Vancouver, Wash., and Daniel R. H. O'Connell, Bureau of Reclamation, Lakewood, Colo.

Debris flows are common on volcanoes and pose a significant hazard that persists long after eruptive activity has ceased. Hazard estimates must incorporate both model uncertainties, which relate deposits to the magnitude of the flow that created them, and data uncertainties, which relate here to the ability to estimate deposit age and volume. While the uncertainty in age of debris-flow deposits commonly is constrained by C14 dates on organic matter, the uncertainty in volume is usually greater than the volume estimate. Currently, we rely on crude estimates of volume based principally on thickness of debris-flow deposits on flood plains. These crude estimates of volume could be better constrained by physical modeling, which would route flows through a given drainage to determine what volume corresponds to each deposit as well as provide the relation between runout distance and volume.

For hazard-assessment applications, the critical issue is not the estimation of a complete record of all flows but rather the frequency and potential magnitudes of flows large enough to exceed the safe conveyance capacity of the system. The bounds on volume and age in the geologic record are direct indicators of the frequency of the largest actual flows. A Bayesian methodology developed for flood-hazard estimation has been modified for use with debris flows and is applicable to nonvolcanic debris-flow hazards as well. The Bayesian approach naturally incorporates both nonexceedance and exceedance information over time using likelihood functions in a unified probabilistic flow-frequency analysis. The method includes both frequency-model and data-estimation uncertainties to derive Bayesian estimates of the annual probability that a debris flow of a given size will occur on a particular volcano. Application of this methodology to the debris-flow record for Mt. Rainier, Wash., indicates that debris flows large enough to be hazardous to the town of Ortiz are likely to occur more frequently than previous analyses would indicate.

OBSTACLES ON THE ROAD TO CONSISTENT PROBABILISTIC HAZARD ASSESSMENT

Daniel R. H. O'Connell, Bureau of Reclamation, Lakewood, Colo.

Probabilistic hazard-estimation goals require extrapolation beyond the limits of historical observation. Extrapolation requires either physical understanding of the system or extending the observational record. In earthquake-hazard estimation, there are physical bases for using the truncated Gutenberg-Richter magnitude-frequency relation for low-probability events. While comparable physical bases for magnitude frequency have not been found for floods, landslides, and volcanic eruptions, the geologic record contains abundant information about the occurrence, or the lack of occurrence, of large magnitude/low-probability events. However, there is incomplete understanding of the physics governing seismic wave, water, and sediment-routing processes during extreme events for all these systems. This makes probabilistic hazard assessment difficult.

Determinism and bias are other factors that undermine the goal of probabilistic hazard assessment. Characteristic earthquake models based on fault mapping are a deterministic approach relative to using distributed or "smeared" source zones to estimate seismic hazards. Virtually all extreme peak-discharge estimates for floods are derived from indirect estimates that often exhibit substantial biases toward extreme values that suggest a consistently conservative bias. Extreme event probabilities based solely on short observational records produce substantial biases when the physical processes representing extreme behavior are completely absent from the record.

Natural variability and "nonstationarity" are serious concerns when making hazard predictions. Probabilistic hazard assessment is viewed by the public, and most users, as making hazard predictions. Does process variability represent nonstationarity or random phase periodicity? Earthquake prediction has not worked well so far for earthquake-hazard estimation. Time-dependent hazard assessment may become viable for some small-dimensional systems representing integrated processes like lake-surface elevations. However, for large-dimensional point processes like earthquakes and floods, the stochastic assumption is probably inescapable for probabilistic hazard assessment in the next few years. What's needed to bridge the knowledge gap is substantial new basic science research and physical understanding that requires collecting detailed, long-term geological and geophysical data. Thus, the U.S. Geological Survey cannot simply be a group of applied consultants. Advances in probabilistic hazard assessment require a renewed commitment to extract new information from the geologic record. Without a strong U.S. Geological Survey commitment to sustained data collection, detailed field investigations, and long-term basic science, applied hazard-assessment efforts will run out of fuel.

APPENDIX C--Papers

Characteristics of Floods as a Spatial Process

by
Brent M. Troutman and Michael R. Karlinger
U.S. Geological Survey

Presented at the Workshop on Probabilistic Risk Assessment for
Earthquakes, Floods, Landslides, and Volcanoes

November 16-17, 1999

Abstract

River flow extremes constitute a complex random process in both space and time. In this paper, some basic characteristics of flood flows considered as a spatial process are reviewed and discussed. Of particular interest are both marginal and joint distributional properties of annual peak flows at a collection of points in some large region. Two widely used regionalization methods for modeling the dependence of the marginal distribution of annual flood peaks on drainage area -- the index-flood method and regional quantile regression -- are compared. In addition, problems associated with characterizing spatial correlation of flows are discussed, and generalized least squares as used by USGS to account for spatial correlation in quantile regression is briefly introduced. Finally the problem of estimating regional flood probabilities, with the goal of assessing frequency of occurrence of flow extremes somewhere in a large region, is discussed. An example is presented showing how this frequency may be estimated using annual peak flow data from the state of Washington.

Introduction

A large part of flood-frequency analysis is concerned with the space-time process of annual peak flows in a region. In this paper we shall consider in particular properties of peak flows as a spatial process. Let x stand for a point in a region R of interest, and denote by Q(x) the peak flow at x in a given year. There are two factors that are responsible for much of the interesting spatial behavior that we see in Q(x). The first is channelization of flow by the drainage network. This channelization results in a high degree of dependence of the distribution of Q(x) on the drainage area, A(x), associated with the point x. The second factor is the spatial coherence of precipitation patterns. Both of these factors result in a high degree of spatial correlation in the Q(x) process. We shall look at a sampling of some approaches to modeling these two factors and at some implications for regionalization and determination of regional flood potential.

Properties of Annual Peaks

Let us first consider the marginal distribution of the annual peak random variable Q(x) at a point x. Hydrologists often choose to express properties of this distribution in terms of the quantile function $q_\alpha(x)$, defined by

$$\alpha = \mathrm{Prob}[Q(x) > q_\alpha(x)], \tag{1}$$

where $0 \le \alpha \le 1$ is an exceedance probability. If $T = 1/\alpha$ (taking $0 < \alpha \le 1/2$), $q_\alpha(x)$ is the so-called T-year flood. Considered as a spatial process, the most prominent feature of Q(x) is that it is highly nonstationary, and the main reason for this arises from the dependence of the distribution of Q(x) on drainage area, A(x). Much of the work in *regionalization* has been, in general terms, directed toward understanding this dependence on area and trying to find a way that data from basins with different areas may be compared or combined for purposes of sta-

tistical analysis. We shall illustrate the problems here by considering two approaches that have been used to look at how $q_\alpha(x)$ depends on A(x): (1) the index-flood approach, and (2) the quantile regression approach. In all that follows we shall refrain from showing the dependence on x when there is no chance of confusion.

Rather than considering flood dependence only on area, we can be more general and look as well at dependence on other physiographic and climatic characteristics, such as main channel slope or mean annual precipitation, that are believed to be related to flood magnitude; let us denote these variables by $Z_1, Z_2, \ldots, Z_p$, where it is understood that the Z_i depend on x. The major assumption for the index-flood approach is that there exists a scale factor, say $\sigma(x)$, that is a function of the Z_i,

$$\sigma(x) = f(Z_1, Z_2, \ldots Z_p), \tag{2}$$

and such that

$$Y(x) = \frac{Q(x)}{\sigma(x)} \tag{3}$$

has a distribution that is the same for all points x in the region. The advantage of this assumption in regionalization studies is that we may pool Y(x) from different sites for estimation of the common probability distribution, thereby effectively increasing sample size over a site-specific estimation approach. (For further discussion of this point, see Hosking and Wallis (1997).) An implication of the index-flood assumption is that the flood quantile may be factored as

$$q_\alpha(x) = \sigma(x) y_\alpha, \tag{4}$$

where y_α is the quantile of Y(x), or, alternatively, that scaling of flood quantiles by a "reference flow," such as the median,

$$\frac{q_\alpha(x)}{q_{0.5}(x)}, \tag{5}$$

gives a quantity that is constant for all basins in the region, and in particular that does not change with area. How well do actual floods conform to the assumptions of the index-flood method? A number of studies have revealed systematic departures from these assumptions. For example, Dawdy (1961) presents data from several regions in the U.S. for which there is a tendency for the ratio in Eqn. (5) to decrease as a function of drainage area, indicating a violation of the index-flood assumption. Historically, one reason the quantile regression method was adopted by USGS over the index flood method is empirical evidence that index-flood assumptions do not hold.

For the quantile regression approach, it is assumed that flood quantiles may be represented as

$$q_\alpha(x) = \beta_0 Z_1^{\beta_1} \ldots Z_p^{\beta_p}. \tag{6}$$

This can be compared to the corresponding relation in Eqn. (4) for the index flood model. It is seen that these models are compatible only if the regression coefficients $\beta_1, \beta_2, \ldots, \beta_p$ do not change with α, and this does not seem in general to hold. Gupta and Dawdy (1995) point out that the exponent for drainage area tends to vary less with α in regions with floods dominated by seasonal snowmelt runoff.

Given that the at-site distribution of Q(x) can be described, there is still the problem of characterizing correlation between sites. Runoff correlation is caused both by spatial correlation of precipitation (for example, flows resulting from large frontal systems) and also by structure imposed by the channel network (for example, correlation between two stations is highly dependent on whether one is upstream from the other). This correlation has implications for parameter estimation in quantile regression, determination of regional flood probabilities (which is considered in more detail below), and geostatistical approaches to interpolation (kriging).

Consider parameter estimation in the quantile regression approach. Assume that only drainage area is included in the model, or that

$$q_\alpha(x) = \beta_0 A^{\beta_1}. \tag{7}$$

Log-linear regression may be performed, so that the regression model may be written

$$\log q_\alpha(x) = \log\beta_0 + \beta_1 \log A + e, \tag{8}$$

where e is the "model error" term. We do not know $q_\alpha(x)$ exactly, so we estimate it with $\hat{q}_\alpha(x)$ obtained using at-site data. The regression model in Eqn. (8) becomes

$$\log \hat{q}_\alpha(x) = \log\beta_0 + \beta_1 \log A + e + d, \tag{9}$$

where d is sampling error in the quantile estimate. It is unrealistic to assume that the error term e+d here obeys the usual regression (ordinary least-squares) assumptions of independence and equality of variances. Clearly Var(d) may be expected to change with record-length, which in general differs from site to site, and also there is dependence which comes from intersite correlation. Stedinger and Tasker (1985, 1986) have developed a generalized least-squares (GLS) technique, now used by USGS, to account for the unequal variances and correlation in regionalization studies.

One requirement of the GLS method is the need to parsimoniously characterize the cross-correlation of annual peaks, and there are problems with doing this. There is typically (see, for example, Tasker and Stedinger (1989) Figure 1) a high degree of scatter in plots of annual peak correlation versus distance between stations, indicating that second-order stationarity (correlation dependence only on distance) is not a good assumption. As mentioned before, when considering covariation of annual peaks at two stations, whether one station is upstream from the other would be expected to have a large influence on the correlation. An important open question is how to best characterize spatial flow correlation structure.

Regional Flood Probabilities

In the preceding section we discussed some issues in modeling $q_\alpha(x)$, i.e., the marginal distribution of Q(x), at single points x. Hydrologic regionalization methodologies use data from a number of stations in a region to make inferences about $q_\alpha(x)$, but in this section we wish to go further and consider random quantities that depend on the joint distribution of annual peaks at a number of sites: $Q(x_1), Q(x_2), \ldots, Q(x_n)$. Although there are many spatial functions of annual peaks that might be of interest, we shall consider the question of regional extremes, for which it is desired to look at the distribution of the spatial maximum

$$M_R = \max_{x \in R} Q(x), \tag{10}$$

where R denotes a region of interest; if n particular sites are under consideration, then

$$R = \{x_1, x_2, \ldots, x_n\}. \tag{11}$$

For example, one reason to look at the distribution of the regional maximum would be to answer the question: how often can a large flood be expected to occur *somewhere* in a large region R. This question may be of interest to, for instance, governmental emergency planners whose jurisdiction covers some large region.

Determining the frequency of occurrence of large floods depends on how one chooses to define "large flood." Let P denote the "regional flood probability," and assume that a large flood is taken to be flow of some fixed magnitude q_0. Then P is given in terms of the distribution of M_R by

$$P = \text{Prob}[Q(x) > q_0 \text{ for some } x \text{ in } R] = \text{Prob}[M_R > q_0]. \tag{12}$$

An alternative way to look at the question is to consider the distribution of

$$M'_R = \max_{x \in R} Y(x) \tag{13}$$

rather than M_R in Eqn. (10). Here, as in Eqn. (3), Y(x) stands for an annual peak that has been standardized so that the at-site distribution is the same for all sites x. Now the regional flood probability, for a fixed exceedance probability α, is

$$P = \text{Prob}[Q(x) > q_\alpha(x) \text{ for some x in R}] = \text{Prob}[M'_R > y_\alpha]. \quad (14)$$

Hence, a large flood in this second formulation is defined in terms of a fixed return period T $(=1/\alpha)$ rather than in terms of a fixed absolute magnitude q_0 as in Eqn. (12). P then represents the probability that a T-year flood occurs somewhere in R in a given year. We shall consider only the second formulation in the remainder of the paper; the regional flood probability P will be taken to be defined by Eqn. (14).

P depends on the spatial correlation of flows among the sites in R. If sites in R are perfectly correlated, then $P = \alpha$, or the regional flood probability is the same as the single-site exceedance probability. If on the other hand R consists of n independent sites, then $P = 1 - (1 - \alpha)^n$, which for small α is approximately equal to $n\alpha$. Thus, for example, for n=2 independent sites, the 100-yr. flood occurs about every 50 years on the average. In applications, spatial dependence will be somewhere between these two limiting cases of perfect correlation and no correlation. Clearly P will be determined by the relative location of the n points in Eqn. (11), with not only geographical distances but channel network connections playing a role.

Let us illustrate these ideas using peak flow data from the state of Washington. A set of 10 stations in the Puget Sound area was selected, 5 of which were located west of the Sound (on the Olympic Peninsula) and 5 of which were east of the Sound (receiving flow from the Cascade Range). Data for the 19 common years of record for this set of stations were used to estimate parameters. Peaks were found to obey a log-normality assumption well, so standardization was done using at-site mean $(\mu(x))$ and standard deviation $(\varphi(x))$ of log Q(x):

$$Y(x) = \frac{\log Q(x) - \mu(x)}{\varphi(x)}. \quad (15)$$

(If $\varphi(x)$ is taken to be constant, this is essentially equivalent to the standardization for the index-flood method, Eqn. (3).) Monte Carlo methods were then used to obtain the distribution of M'_R in Eqn. (13) using the estimated correlation matrix, and this distribution was in turn used to estimate the regional flood probability P from Eqn. (14). This was done for the 10 stations together and for each of the 5-station subsets. Results are given in Table 1 for $\alpha = 0.02$ (50-year flood) and $\alpha = 0.01$ (100-year flood). Values of the return period (1/P) in this table are rounded to the nearest integer. The effects on P of both the number of stations and the average correlation among the stations can be seen in this Table. For instance, for the 5 stations on the Olympic Peninsula, the average correlation was quite high (0.834), and the 100-year flood will occur on the average about every 40 years. On the other hand, for the 5 stations receiving flow from the Cascade Range, the average correlation was lower (0.377) and the 100-year flood will occur every 24 years. The 10 stations together will see a 100-year flood every 17 years, more frequently than for either of the 5-station subsets.

The fact that P in Eqn. (14) is given in terms of the distribution of the maximum of identically distributed random variables suggests that asymptotic extreme value theory may be of some use in looking at regional flood probabilities if the number of stations is large. It is well known, for example, that the maximum of independent standard normal random variables is asymptotically a two-parameter Gumbel. It is also known that, under certain weak conditions on autocorrelation, the maximum of a stationary sequence obeys the same law, and in fact there are additional results that allow extension to nonstationary sequences (e.g., Berman (1964)). If a two-parameter Gumbel distribution is assumed to apply for regional flood probabilities, an explicit expression for P is

$$P = 1 - \exp(-e^{\gamma - \beta y_\alpha}); \quad (16)$$

using this expression eliminates the need to use Monte Carlo methods to obtain P. Asymptotic dependence of the parameters γ and β on n is well-known for independent and stationary sequences, and it is also known that in passing to a continuous limit local effects come into play which lead to correction of γ by an “extremal index”

which depends on the correlation function (Pickands (1967)). Use of such an index would be appropriate if we wished to estimate regional flood probabilities for sites that are close together. An important question, then, in using the Gumbel distribution in Eqn. (16) to approximate P is to understand how the parameters γ and β depend on the region R of interest, and in particular how they are influenced by precipitation correlation and network effects. These are subjects of ongoing investigation.

Summary and Conclusions

Some approaches have been presented for looking at annual peaks Q(x) as a spatial process. In particular, we have looked at the index-flood method and regional regression as ways of characterizing how the distribution of Q(x) depends on drainage area and other basin characteristics, and at some consequences of spatial correlation in the Q(x) process for estimation in regional regression and determination of regional extremes. However, much remains to be done before a full understanding of these problems is at hand. Among the important open questions that need to be addressed in future research are: (1) Given the lack of spatial second-order stationarity in peak flows, what is the most useful way to characterize the spatial correlation structure? (2) How is the spatial correlation of flows related to that of precipitation and to network structure? We have proposed extreme value theory as one promising way of addressing the second question.

Table 1. Regional Flood Probabilities for Washington

	Average correlation	P 50-year flood	1/P 50-year flood	P 100-year flood	1/P 100-year flood
10 sites	0.496	0.107	9	0.059	17
5 sites (Cascade Range)	0.377	0.077	13	0.041	24
5 sites (Olympic Peninsula)	0.834	0.047	21	0.025	40

References

Berman, S.M., 1964, Limit theorems for the maximum term in stationary sequences, Ann. Math. Statist., 35, 502-516.

Dawdy, D.R., 1961, Variation of flood ratios with size of drainage area, U.S. Geological Survey Prof. Paper 424-C, p. C36-C37.

Gupta, V.K. and Dawdy, D.R., 1995, Physical interpretations of regional variations in the scaling exponents of flood quantiles, Hydrological Processes, 9, 347-361.

Hosking, J.R.M. and Wallis, J.R., 1997, Regional frequency analysis, Cambridge University Press, Cambridge.

Pickands, J., 1967, Maxima of stationary Gaussian processes, Z. Wahr. verw. Geb., 7, 190-223.

Stedinger, J.R. and Tasker, G.D., 1985, Regional hydrologic analysis 1. Ordinary, weighted, and generalized least squares compared, Water Resour. Res., 21(9), 1421-1432.

Stedinger, J.R. and Tasker, G.D., 1986, Regional hydrologic analysis 2. Mean-error estimators, estimation of sigma, and log-Pearson type III distributions, Water Resour. Res. 22(10), 1487-1499.

Tasker G.D. and Stedinger, J.R., 1989, An operational GLS model for hydrologic regression, J. Hydrol., 111, 361-375.

Landslide Risk Assessment

by Dr. William Roberds, Principal, Golder Associates Inc., Seattle Washington

Overview

Landslides can occur in different forms, ranging from individual rock falls to large creep failures, depending on site conditions (e.g., topography, geologic structure, shear strength, pore pressures, and loads), some of which in turn are affected by various processes and events (e.g., earthquakes, precipitation, erosion, or excavation). Such landslides can have significant consequences (e.g., casualties, property damage, and socio-economic impacts such as loss of service), depending on: a) their characteristics (e.g., number, timing, location, size, mobility, runout, etc.), which constitutes the "hazard"; and b) the "vulnerability" of people, structures and infrastructure to those landslide characteristics, which in turn is a function of the population and development in the potentially affected area and their characteristics (e.g., location, occupancy, awareness, mobility, damage susceptibility, etc.). Such undesirable consequences can be prevented or at least reduced through various actions designed to: a) prevent landslides from occurring (e.g., by installing drains or support, scaling or flattening the slope); b) reduce their severity if they do occur (e.g., by installing debris barriers); and/or c) reduce the vulnerability of potentially affected people/structures/infrastructure (e.g., controlling development, warning/evacuating people, strengthening structures, maintaining contingency plans). However, some of these actions can be very expensive to implement (in financial terms as well as in worker safety and socio-economic terms) and may prove to be either ineffective or unnecessary. The objective for any site or, even more importantly, a set of sites should be to identify and implement those actions (if any) which are most cost-effective, appropriately trading off: a) the various “costs” of implementing the actions; b) the "benefits" of reduced consequences of landslides; and c) the change in value (e.g., for development) of potentially affected land. Such tradeoffs, and thus preferences amongst the alternatives, obviously may vary amongst the different stakeholders, e.g., with regulators focusing on risk reduction and developers focusing on costs and land values.

However, for any action (including no action), the occurrence of landslides at a site and the characteristics and consequences if they do occur cannot generally be predicted with certainty beforehand, due to inherent uncertainties in: a) the site conditions/processes/events and in the vulnerability of people/structures/infrastructure, a combination of stochastic variability (temporal and/or spatial, which is natural and not reducible) and ignorance (which is reducible through more information); and b) the relationships of instability, landslide characteristics and consequences to the site conditions/processes/events and the vulnerability of people/structures/infrastructure, a combination of simplifications and approximations.

The uncertainties in whether one or more landslides will occur at a site, and in the consequences if they do occur (“risks”), need to be adequately considered when deciding on possible actions. In fact, in some jurisdictions, acceptable levels of risk are specified, and compliance must be adequately demonstrated. Such risks can be assessed in many ways, differing widely in effort/cost, accuracy and defensibility. Fundamentally, each method is a combination of the following attributes: a) site-specific or averaged over a set of sites; b) qualitative or quantitative; and c) comparative or absolute. Although qualitative and/or comparative methods may be adequate for ranking or screening slopes and possible actions for slopes, a site-specific, absolute, quantitative method is generally needed for making optimal, defensible decisions on significant actions. However, such quantitative methods can vary significantly in level of detail, ranging from: a) direct assessments of the frequency of landslides and their consequences; to b) detailed models of various modes of instability, runout and consequences which in turn rely on direct assessments of specific input parameters. Direct assessments (at any level) can be: a) statistically driven (if a representative data set is available); b) based on judgment (consistent with available data); or c) a combination of the two. Similarly, the models can be: a) empirically-based (which requires substantial data from similar situations relating

outcomes to observable factors, e.g., geologic units); b) theoretically-based (which requires the assessment of engineering parameters, e.g., peak pore pressure); or c) a combination of the two. Care must be taken in assessing parameter values to ensure that: a) the appropriate characteristic of that parameter (e.g., random value, average value or extreme value over a particular area or time period) is being assessed (i.e., consistent with how it will be used); b) the assessments are consistent with all available data; and c) any correlations (e.g., in one parameter at different times or locations, or between different parameters at the same time and location) are appropriately considered.

The appropriate method to use for risk assessment on a particular project depends on a variety of factors, primarily the "value" of increased accuracy in making decisions and increased defensibility in justifying those decisions. Other factors include: a) the availability of relevant engineering and empirical data, and the cost associated with acquiring additional data; and b) the availability of tools and expertise (e.g., regarding stability, runout and damage models, and objective and subjective probability assessments), and the cost associated with better tools and expertise.

The above concepts, which are discussed in more detail elsewhere (e.g., ref. Roberds and Ho, 1997), are summarized in Figure 1, which shows how decisions on the various types of actions (i.e., investigation, design, construction, land use controls, inspection/monitoring, maintenance/repair, and emergency response) will affect the risks as well as the other costs and benefits: a) the consequences of landslides are due to a combination of hazards and vulnerability; b) hazards result from changes in initial conditions, which in turn are due to natural processes and events as well as human actions (e.g., construction and repair); c) investigations are used to determine (with some uncertainty) initial conditions, and natural processes and events that may affect the system, as well as normal vulnerability; d) changes in conditions can be detected, triggering repairs (to reduce the hazard) or evacuations (to reduce vulnerability), if necessary; and e) the design and land use controls are intended to manage risks by reducing the hazards and vulnerability.

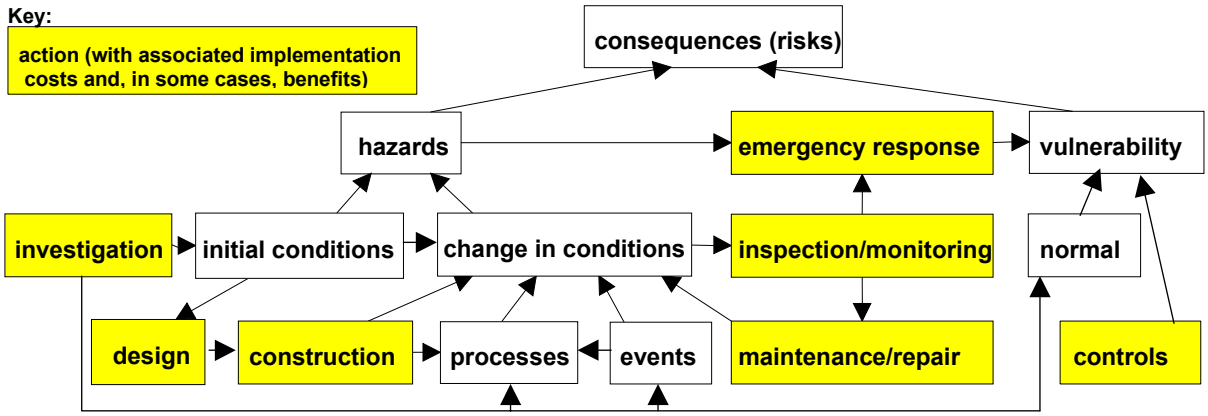

Figure 1. Assessment and Management of Landslide Risks

Risk Assessment Methods

A landslide risk assessment typically consists of the following steps, some of which may be simplified:

1. Identify area of interest (both upslope and downslope) and potential detachment modes (e.g., shallow colluvium failures, deep rotational failures, structurally controlled wedge failures, rock or boulder falls, etc.);
2. Identify potential detachment areas for each detachment mode, and evaluate likelihood of each, recognizing correlations amongst the areas and modes (e.g., one failure mode will typically

preclude another in the same area, and a failure mode which occurs in one area will be more likely in other nearby areas);

3. Identify potential movement modes for each detachment mode (e.g., sliding, rolling/bouncing, slurry), and evaluate probable extent for each from each detachment area (e.g., in terms of runout angles);
4. Identify vulnerable elements (e.g., population, property and critical infrastructure), and evaluate probable consequences for each movement mode for each detachment mode from each detachment area; and
5. Integrate over all vulnerable elements, movement modes, detachment modes and detachment areas (e.g., in GIS), and determine relative contribution of each factor.

The types of assessments required for such a landslide risk assessment are summarized in Figure 2, differentiating between (a) direct assessments and (b) modeling of frequency-magnitude relationships (λ-m) or conditional probability distributions (p|m) at each level of detail. These assessments are then chained together to determine the landslide risks (λ-m of adverse consequences). Such assessments and analyses are discussed in more detail elsewhere (e.g., ref. Roberds and Ho, 1997).

<table>
<tr><td rowspan="15">consequences (λ-m_c), including casualties, repair costs and other losses</td><td colspan="4">a) directly from cleanup or accident statistics (e.g., along highways, railroads, or pipelines) in similar conditions</td></tr>
<tr><td rowspan="12">b1) hazards (λ-m_h), including undermining and debris runout</td><td colspan="3">a) directly from the time history of slope movements in similar conditions (e.g., rockfalls in talus slopes or road indentations, slip/debris flows from geologic mapping including air photo int3erpretation, large scale slope movement, or slope regression)</td></tr>
<tr><td rowspan="5">b1) detachment (λ-m_d), including various modes</td><td colspan="2">a) directly from inferred time -history of slips or other detachments in similar conditions from geologic mapping, including air photo interpretation</td></tr>
<tr><td rowspan="2">b1) triggers (λ-m_t), including natural (e.g., precipitation or earthquake), man-induced (e.g., excavation, loads, infiltration from leaks), or either (e.g., erosion or changes in vegetation)</td><td>a) directly from inferred time-history of triggers (e.g., by type)</td></tr>
<tr><td>b) various process models (e.g., seismic hazard or climate/weather)</td></tr>
<tr><td rowspan="2">b2) conditional detachment ($p_s|m_t$), given trigger</td><td>a) directly from statistics of detachments in similar geologic, hydrologic and geometric conditions, conditioned on specific triggers (e.g., by type and magnitude)</td></tr>
<tr><td>b) stability (and infiltration) models</td></tr>
<tr><td rowspan="3">b2) movement ($p_x|m_d$), including various modes</td><td colspan="2">a) directly based on the statistics of movement of detachments in similar conditions from geologic mapping, including air photo interpretation</td></tr>
<tr><td colspan="2">b1) mobility ($p_m|s$), from statistics of movements in similar geologic and hydrologic conditions (e.g., runout angles), or theoretical considerations (e.g., mass sliding, slurry flow, individual rock/boulder bouncing/rolling, or discontinuous displacement of discrete particles)</td></tr>
<tr><td colspan="2">b2) downstream topography (z|x), from survey/maps</td></tr>
<tr><td rowspan="5">b2) vulnerability ($p_c|m_h$), including people, property and services</td><td colspan="3">a) directly from complete hazard incident reports in similar conditions, with associated damage noted</td></tr>
<tr><td rowspan="2">b1) intersection ($p_i|m_h$)</td><td colspan="2">a) directly based on either complete hazard incident reports in similar conditions, noting the consequences</td></tr>
<tr><td colspan="2">b) geometric model (considering number, size, location, occupancy, warning, mobility, etc.)</td></tr>
<tr><td rowspan="2">b2) damage ($p_d|i,m_h$)</td><td colspan="2">a) directly based on complete damage or accident reports in similar conditions, noting the cause</td></tr>
<tr><td colspan="2">b) damage model (e.g., event tree or structural analysis)</td></tr>
</table>

Figure 2. Landslide Risk Factors and their Assessment

As noted above (particularly in Figure 2), models can be used to determine various aspects of landslide risk. For example, various probabilistic slope stability models have been developed in the past to quantify the probability of failure (detachment) for specific detachment modes (e.g., shallow colluvium failures, deep rotational failures, and structurally controlled wedges) (e.g., ref. Roberds and Ho, 1997). Recently, other models have been developed to do such probabilistic analysis better (e.g., for detachment) or to do what has not been done before (e.g., for infiltration and rock fall runout), including amongst others the following:

- Detachment of shallow zones (e.g., colluvium) and its timing can be modeled by probabilistic dynamic simulation (using GoldSim, ref. Kossik et al, 2000) as follows: 1) the relevant time independent properties (i.e., slope angle, depth to potential failure plane, average soil density, and average cohesion and friction angle on failure plane) are randomly sampled from their specified probability distributions (including correlations); 2) the relevant time dependent properties (i.e., normal depth to phreatic surface, changes in depth to phreatic surface due to storms, and seismic loading due to earthquakes) are randomly sampled at each time step from their specified probability distributions (including correlations); 3) the safety factor on the potential failure plane at each time step, and the minimum factor of safety over the specified time period, is determined; and 4) based on many such analyses, the probability of failure (i.e., detachment, defined as the minimum factor of safety less than 1.0) is determined. An example of the computer screen for such an analysis (including one time-history for the factor of safety) is presented in Figure 3.

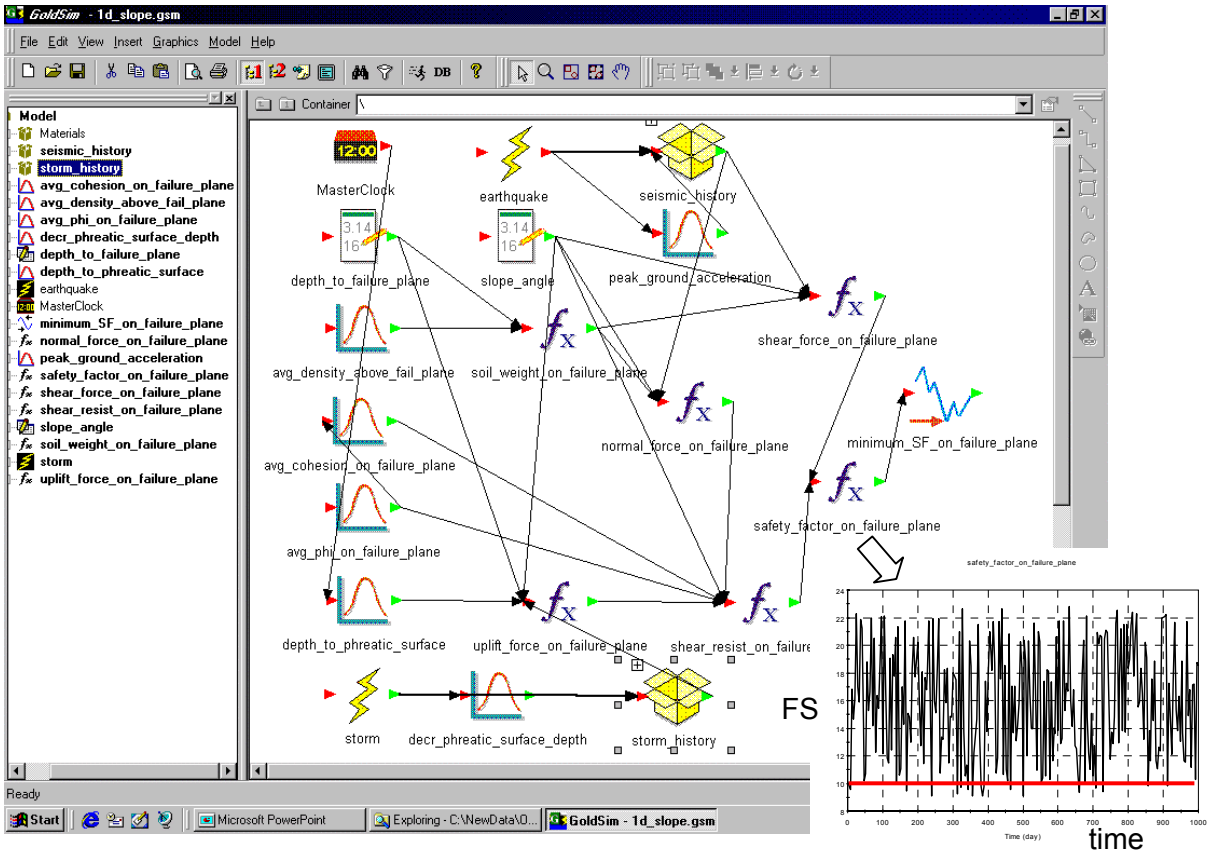

Figure 3. Example of Shallow Detachment Analysis (GoldSim)

- Detachment of structurally controlled wedges and their number, location and characteristics can be modeled probabilistically (using FracMan/RockBlock, ref. Dershowitz et al 1998) as follows: 1) the relevant properties (i.e., slope fracture geometry, and fracture strength characteristics) are

randomly sampled from their specified probability distributions; 2) the locations/orientations of fractures are simulated, conditioned on any information (e.g., from structural mapping); 3) all the kinematically possible wedges in the slope are identified and their stability analyzed (based on limiting equilibrium, with simulated strength values); and 4) based on many such analyses, the uncertainty in whether any wedges fail, and if so, the probable number, location and size of wedges, is determined. An example of the computer screen for such an analysis is presented in Figure 4.

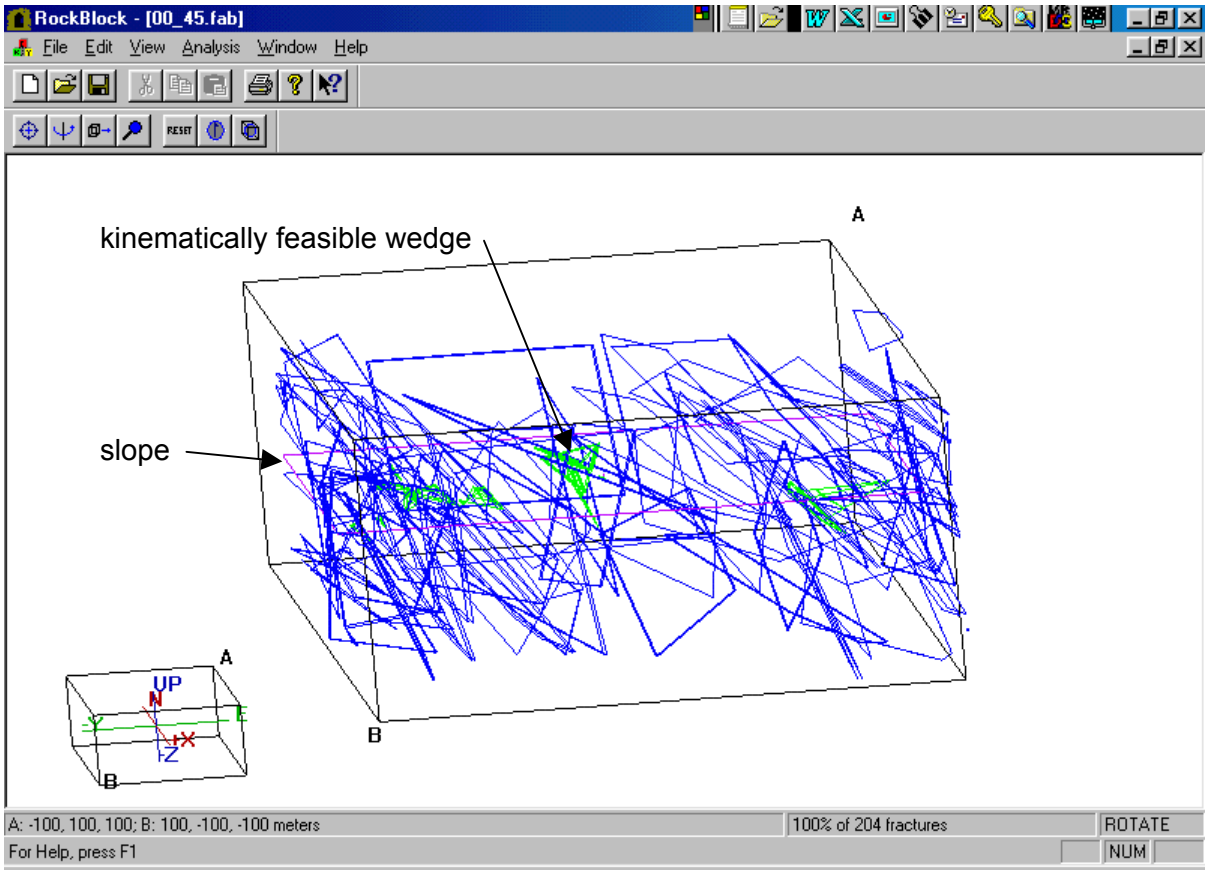

Figure 4. Example of Structurally-Controlled Wedge Detachment Analysis (FracMan/RockBlock)

- Infiltration and hence pore pressures of fractured rock can be modeled probabilistically (e.g., also using FracMan) as follows: 1) the relevant properties (i.e., fracture geometry, and fracture transmissivity) are randomly sampled from their specified probability distributions; 2) the locations/orientations of fractures are simulated, conditioned on any information (e.g., from structural mapping); 3) all the fracture connections are identified and the transmissivity of each potential "pipe" analyzed (based on simulated transmissivity values); and 4) based on many such analyses, the uncertainty in the overall transmissivity, is determined.

- Runout of rockfalls can be modeled by probabilistic dynamic simulation (e.g., using RockFal3, ref. Lee and Elliott, 1998) as follows: 1) the relevant properties (i.e., rock and slope geometry, and bounce characteristics) are randomly sampled from their specified probability distributions; 2) the path of the rock is dynamically simulated using those properties; and 3) based on many such analyses, the uncertainty in maximum runout, and the uncertainty in the bounce height and energy at any location (e.g., a fence), is determined. An example of the computer screen for such an analysis is presented in Figure 5.

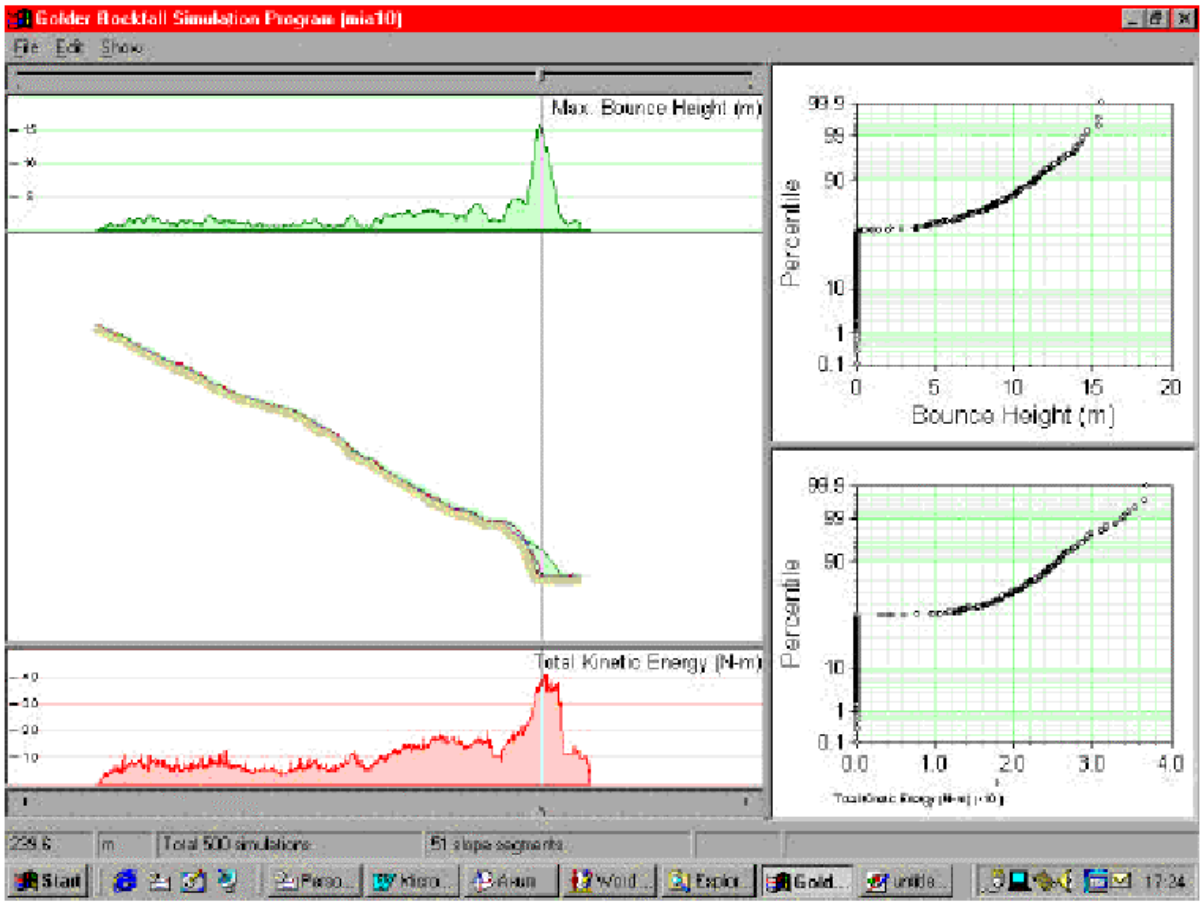

Figure 5. Example of Rock Fall Analysis (RockFal3)

Summary

Risk assessment, if appropriately done, adequately and efficiently quantifies the "risks" associated with possible landslides for a particular slope or set of slopes, in terms of the probability distributions for various types of unplanned and undesirable consequences (e.g., casualties, property damage, loss of service, etc.), considering all possibilities. Potential risk management activities focusing on the key risk factors can be identified and logically evaluated in terms of how much they reduce the risks, in conjunction with their costs and other benefits, so that the risks can be managed in an optimal way.

References

Dershowitz, W., G. Lee, J. Geier, T. Foxford, P. LaPointe, and A. Thomas, 1998. FracMan, Interactive Discrete Feature Data Analysis, Geometric Modeling, and Exploration Simulation: User Documentation (Version 2.6), Golder Associates, Redmond WA.

Lee, K., and G. Elliott, 1998. "Rockfall: Application of Computer Simulation to Design of Preventive Measures," paper presented at Planning, Design and Implementation of Debris Flow and Rockfall Hazards Mitigation Measures, 27 October 1998, Hong Kong.

Kossik, R., I. Miller, and S. Knopf, 2000. GoldSim, Graphical Simulation Environment: User's Guide (Version 6.03), Golder Associates, Redmond WA.

Roberds, W. and K. Ho, 1997. "A Quantitative Risk Assessment and Risk Management Methodology for Natural Terrain in Hong Kong," in First International Conference on Debris-Flow Hazards Mitigation: Mitigation, Prediction and Assessment, ASCE, San Francisco, CA, August 1997.

PROBABILISTIC LANDSLIDE HAZARD ASSESSMENT

Randall W. Jibson
U.S. Geological Survey

INTRODUCTION

For the purposes of predictive hazard assessment, landslides can be separated into two categories: those that are triggered by observable events such as earthquakes and rainfall, and those that occur with no discernable trigger. With our current state of knowledge, we have no way to meaningfully predict (time, location, magnitude) the latter. Hazard predictions of triggered landslides inherently involve conditional probabilities because the probability of a landslide being triggered is a function of the probability of the triggering event occurring. With the data and models currently available, earthquake-triggered landslides present the best opportunity for quantitative hazard prediction.

Landslides are one of the most damaging collateral hazards associated with earthquakes. Therefore, predicting where and in what shaking conditions earthquakes are likely to trigger landslides is a key element in regional seismic hazard assessment. The factors contributing to seismically triggered slope failure at a specific site are generally complex and difficult to assess with confidence; therefore, regional analysis of a large group of landslides triggered in a well documented earthquake is useful in estimating general conditions related to failure. The 1994 Northridge, California, earthquake (**M** 6.7) presents the ideal case for such an analysis because all of the data sets required for detailed regional analysis of slope failures are available. We summarize here a method (Jibson and others, 1998) to map the spatial distribution of probabilities of seismic slope failure in any set of shaking conditions of interest. The method is calibrated using data from the 1994 Northridge earthquake in the Oat Mountain, Newhall, Piru, Val Verde, Santa Susana, and Simi 7½′ quadrangles, near Los Angeles, California.

MODELING METHOD

We model the dynamic performance of slopes using the permanent-displacement analysis developed by Newmark (1965). Newmark's method models a landslide as a rigid block that slides on an inclined plane (Wilson and Keefer, 1983; Jibson, 1993). The block has a known critical (or yield) acceleration, a_c, which is simply the threshold ground acceleration required to overcome basal shear resistance and initiate sliding. The analysis calculates the cumulative permanent displacement of the block relative to its base as it is subjected to the effects of an earthquake acceleration-time history. An acceleration-time history is selected, and the parts of the record that exceed the critical acceleration are double-integrated to obtain the cumulative displacement of the block (Wilson and Keefer, 1983; Jibson, 1993). Newmark's method is based on a fairly simple model of rigid-body displacement, and thus it does not precisely predict measured landslide displacements in the field. Rather, Newmark displacement is a useful index of how a slope is likely to perform during seismic shaking.

Newmark (1965) showed that the critical acceleration of a landslide block is a simple function of the static factor of safety and the landslide geometry, expressed as

$$a_c = (FS - 1)\, g \sin\alpha, \tag{1}$$

where a_c is critical acceleration in terms of g, the acceleration of gravity; FS is the static factor of safety; and α is the angle from the horizontal that the center of mass of the landslide first moves, which can generally be approximated as the slope angle. Thus, conducting a Newmark analysis requires knowing the static factor of safety and slope angle and selecting an earthquake strong-motion record.

HAZARD-MAPPING METHODOLOGY

The Northridge earthquake is the first for which we have all of the data sets needed to conduct a detailed regional analysis of factors related to triggered landsliding. These data sets include (1) a comprehensive inventory of triggered landslides (Harp and Jibson, 1995, 1996), (2) about 200 strong-motion records of the main shock, (3) detailed geologic mapping of the region, (4) extensive data on the engineering properties of geologic units, and (5) high-resolution digital elevation models of the topography. These data sets have been digitized and rasterized at 10-m grid spacing in the ARC/INFO GIS platform. Combining these data sets in a dynamic model based on Newmark's sliding-block analysis yields estimates of coseismic landslide displacement in each grid cell from the Northridge earthquake. The modeled displacements are then compared with the digital inventory of Northridge landslides to construct a probability curve relating predicted displacement to probability of failure. Data layers consist of 10-m raster grids of the entire quadrangle. Figure 1 is a flowchart showing the sequential steps involved in the hazard-mapping procedure. The sections that follow describe each step in the process.

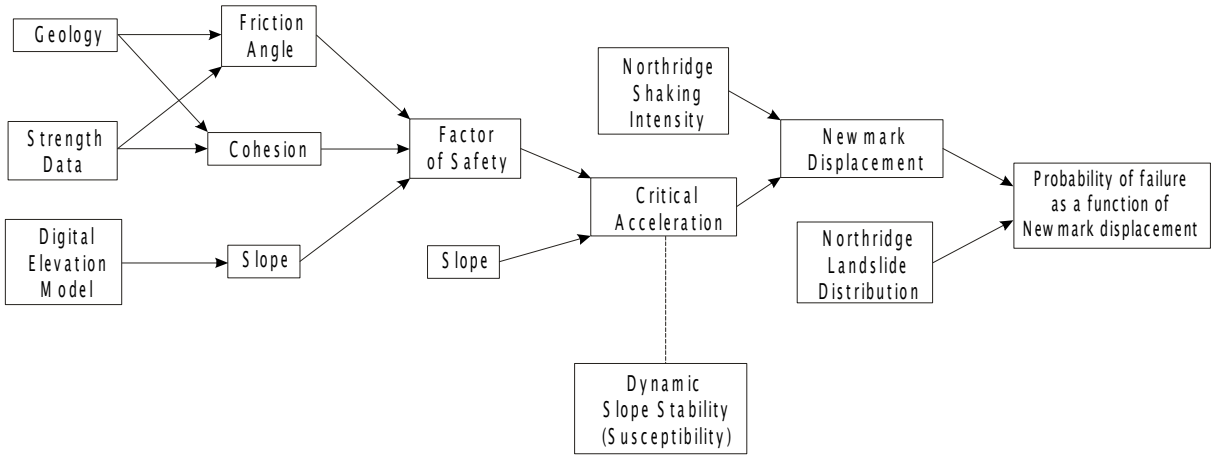

Figure 1. *Flowchart showing steps involved in producing seismic landslide hazard maps.*

Computing the Static Factor of Safety

The dynamic stability of a slope, in the context of Newmark's method, is related to its static stability (see eq. 1); therefore, the static factor of safety (ratio of resisting to driving forces) for each grid cell must be determined. For purposes of regional analysis, we use a relatively simple limit-equilibrium model of an infinite slope with both frictional and cohesive strength:

$$FS = \frac{c'}{\gamma t \sin\alpha} + \frac{\tan\phi'}{\tan\alpha} - \frac{m\gamma_w \tan\phi'}{\gamma \tan\alpha} \qquad (2)$$

where FS is the static factor of safety, ϕ' is the effective friction angle, c' is the effective cohesion, α is the slope angle, γ is the material unit weight, γ_w is the unit weight of water, t is the slope-normal thickness of the failure slab, and m is the proportion of the slab thickness that is saturated. No pore-water pressure is included for this calibration (m=0) because almost all of the Northridge failures occurred in dry conditions. For simplicity, the product γt is a constant 38.3 kPa (800 lbs/ft^2), reflecting a typical unit weight of 15.7 kN/m^3 (100 lbs/ft^3) and slab thickness of 2.4 m (8 ft), representative of Northridge slides. The factor of safety is calculated by inserting values from the friction, cohesion, and slope-angle grids into equation 2.

Geologic map. A 1:24,000-scale digital geologic map (Yerkes and Campbell, 1993, 1995a-h, 1997a-c) forms the basis for assigning material properties throughout the area. Representative values of the frictional and cohesive components of shear strength were assigned to each geologic unit.

Shear-strength data. Shear-strength values for geologic units were selected based on (1) compilation of numerous direct-shear test results from local consultants, (2) the judgment of several experienced geotechnical engineers and geologists in the region, and (3) the constraint that the computer slope model be statically stable. Jibson and others (1998) show the strengths used for each geologic unit and discuss the basis for assigning strengths.

Digital elevation model and slope map. A 10-m digital elevation model (DEM) was produced by high-resolution scanning of the 1:24,000-scale contour plates of the six quadrangles. We selected a 10-m scanning resolution to preserve the subtle topographic features in which many landslides occur; too many topographic irregularities are lost in the more commonly used 30-m DEMs. The slope map was produced by applying a simple algorithm to the DEM that compares the elevations of adjacent cells and computes the maximum slope.

Factor-of-safety grid. Combining these data layers in equation 2 yields a grid factor-of-safety values. Factors of safety range from just greater than 1.0, for steep slopes in weak material, to more than 8 for flatter slopes in strong material.

Computing Critical Acceleration

As indicated above, the critical acceleration of a slope is a function of its static factor of safety and the slope angle (see eq. 1). Therefore, a critical-acceleration grid is created using equation 1 to combine the slope angle with the calculated factors of safety.

Within the context of the Newmark analysis, critical acceleration uniquely describes the dynamic stability of a slope: for a given shaking level, any slopes having the same critical acceleration will yield the same Newmark displacement. The critical-acceleration grid thus portrays a measure of intrinsic slope properties independent of any ground-shaking scenario; thus, it is a map of seismic landslide susceptibility (Jibson and others, 1998).

Estimating Newmark Displacements

To facilitate using Newmark's method in regional analysis, Jibson and others (1998) developed a simplified Newmark method wherein an empirical regression equation is used to estimate Newmark displacement as a function of shaking intensity and critical acceleration:

$$\log D_n = 1.521 \log I_a - 1.993 \log a_c - 1.546, \qquad (3)$$

where D_n is Newmark displacement in centimeters; a_c is critical acceleration is g's; and I_a is earthquake shaking intensity in meters per second, defined by Arias (1970) as the integral of the squared acceleration values in a strong-motion record (Wilson, 1993).

We produced an Arias-intensity ground-shaking grid for the Northridge earthquake by plotting the average Arias intensity from the two horizontal components of each of 189 strong-motion recordings of the mainshock. We then used a simple kriging algorithm to interpolate values across a regularly spaced grid (Jibson and others, 1998).

Newmark displacements from the Northridge earthquake were estimated in each grid cell of the six quadrangles by using equation 3 to combine corresponding grid values of critical acceleration and Arias intensity. Predicted displacements range from 0 to more than 5,000 cm.

Estimating Probability of Failure

For the Newmark method to be useful in a predictive sense, modeled displacements must be quantitatively correlated with field performance. Therefore, we compared the predicted Newmark displacements with the actual inventory of Northridge landslides (Harp and Jibson, 1995, 1996). Newmark-displacement grid cells were grouped into bins, such that all cells having displacements between 0 and 1 cm were in the first bin; those having 1 to 2 cm of displacement were in the second bin, and so on. For displacements greater than about 10 cm, the number of cells in 1-cm bins became very small; therefore, broader ranges of displacement were grouped together. For each bin, the proportion of the cells in landslide source areas was calculated.

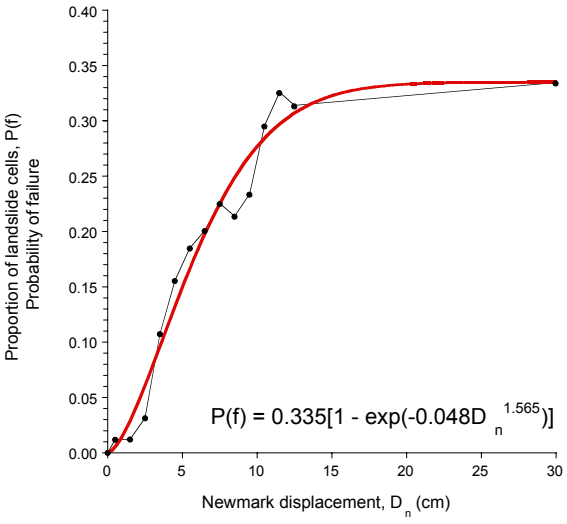

Figure 2. *Proportion of landslide cells as a function of Newmark displacement.*

Figure 2 shows, for each bin, the proportion of cells occupied by landslide source areas plotted as a function of Newmark displacement. The data clearly demonstrate the utility of Newmark's method to predict the spatial density of seismically triggered landslides: the proportion of landslide cells within each displacement bin increases monotonically with increasing Newmark displacement. This relation is critical in a predictive sense because the proportion of landslide cells in each displacement bin is a direct estimate of the probability that any cell in that displacement range will be occupied by a landslide. The proportion of landslide cells increases rapidly in the first few centimeters of Newmark displacement and then levels off abruptly in the 10- to 15-cm range at a proportion of about 35 percent. This maximum proportion of failed slopes is reasonable because, in documenting worldwide earthquakes, we have rarely seen more than 25-35 percent of slope areas fail, even on the most susceptible slopes in epicentral areas.

We fit the data in figure 2 with a Weibull (1939) curve, which was initially developed to model the failure of rock samples (Jaeger and Cook, 1969). The functional form produces an S-shaped curve that is apparent in the data:

$$P(f) = m[1 - \exp(-aD_n^{\,b})], \tag{4}$$

where $P(f)$ is the proportion of landslide cells, m is the maximum proportion of landslide cells indicated by the data, D_n is Newmark displacement in centimeters, and a and b are the regression constants. The expression inside the brackets takes the form of the original Weibull equation, which yields values from 0 to 1; the m outside the brackets scales this range to reflect the range represented by the data. The regression curve based on the Northridge data is

$$P(f) = 0.335[1 - \exp(-0.048\ D_n^{\,1.565})]. \tag{5}$$

The curve fits the data extremely well (R^2=96 percent), and prediction of the proportion of landslide cells ($P(f)$) can be used to directly estimate probability of slope failure as a function of Newmark displacement. Once calibrated, this equation can be used in any ground-shaking conditions to predict probability of slope failure as a function of Newmark displacement.

Producing Seismic Landslide Hazard Maps

Figure 2 and equation 5 relate Newmark displacements to probabilities of landslide occurrence. The curve thus forms the basis for producing seismic landslide hazard maps, which portray spatial variation in slope-failure probability in a specified set of ground-shaking conditions. Jibson and others (1998) show such a map for Northridge ground-shaking conditions.

Constructing a hazard (probability) map for other ground-shaking scenarios is equally straightforward, provided the ground shaking can be reasonably modeled. Such a procedure would involve the following:

A. Specify the ground-shaking conditions in terms of Arias intensity. This could be a uniform level of shaking (for example, representing a 50-year expected maximum shaking level) or shaking generated from a hypothetical earthquake of specified magnitude and location. Simple equations relating Arias intensity to other measures of ground shaking (peak ground acceleration, magnitude and distance, etc.) have been published elsewhere (Wilson and Keefer, 1983; Jibson, 1993; Wilson, 1993).

B. Combine the shaking intensities with the critical-acceleration grid using equation 3 to estimate Newmark displacements.

C. Estimate failure probabilities from the Newmark displacements using equation 5.

SUMMARY AND CONCLUSION

Analysis of data from the Northridge earthquake allows quantitative physical modeling of conditions leading to coseismic slope failure. The procedure described in this paper can be used to produce hazard maps showing the spatial distribution of slope-failure probability. Within the limitations discussed, such maps can find useful application in regional seismic hazard and risk assessment.

REFERENCES

Arias, A., 1970, A measure of earthquake intensity, *in* Hansen, R.J. (ed.) Seismic design for nuclear power plants: Cambridge, Massachusetts, Massachusetts Institute of Technology Press, p. 438-483.

Harp, E.L., and Jibson, R.W., 1995, Inventory of landslides triggered by the 1994 Northridge, California earthquake: U.S. Geological Survey Open-File Report 95-213, 17 p., 2 pl.

Harp, E.L., and Jibson, R.W., 1996, Landslides triggered by the 1994 Northridge, California earthquake: Bulletin of the Seismological Society of America, v. 86, no. 1B, p. S319-S332.

Jaeger, J.C., and Cook, N.G.W., 1969, Fundamentals of Rock Mechanics*:* London, Methuen and Company, 513 p.

Jibson, R.W., 1993, Predicting earthquake-induced landslide displacements using Newmark's sliding block analysis: Transportation Research Record, no. 1411, p. 9-17.

Jibson, R.W., Harp, E.L., and Michael, J.A., 1998, A method for producing digital probabilistic seismic landslide hazard maps: An example from southern California: U.S. Geological Survey Open-File Report 98-113, 17 p., 2 pl.
Newmark, N.M., 1965, Effects of earthquakes on dams and embankments: Geotechnique, v. 15, p. 139-160.
Weibull, W., 1939, A statistical theory of the strength of materials: Ingenioersvetenskapsakademien Stockholm Handlingar, v. 151.
Wilson, R.C., 1993, Relation of Arias intensity to magnitude and distance in California: U.S. Geological Survey Open-File Report 93-556, 42 p.
Wilson, R.C., and Keefer, D.K., 1983, Dynamic analysis of a slope failure from the 6 August 1979 Coyote Lake, California, earthquake: Bulletin of the Seismological Society of America, v. 73, p. 863-877.
Wilson, R.C., and Keefer, D.K., 1985, Predicting areal limits of earthquake-induced landsliding, *in* Ziony, J.I. (ed.), Evaluating Earthquake Hazards in the Los Angeles Region—An Earth-Science Perspective: U.S. Geological Survey Professional Paper 1360, p. 316-345.
Yerkes, R.F. and Campbell, R.H., 1993, Preliminary geologic map of the Oat Mountain 7.5' quadrangle, southern California: U.S. Geological Survey Open-File Report 93-525, 13 p.
Yerkes, R.F. and Campbell, R.H., 1995a, Preliminary geologic map of the Newhall 7.5' quadrangle, southern California: U.S. Geological Survey Open-File Report 95-503, 12 p.
Yerkes, R.F. and Campbell, R.H., 1995b, Preliminary geologic map of the Newhall 7.5' quadrangle, southern California: a digital database: U.S. Geological Survey Open-File Report 95-800.
Yerkes, R.F. and Campbell, R. H., 1995c, Preliminary geologic map of the Oat Mountain 7.5' quadrangle, southern California: a digital database: U.S. Geological Survey Open-File Report 95-89.
Yerkes, R.F. and Campbell, R.H., 1995d, Preliminary geologic map of the Piru 7.5' quadrangle, southern California: U.S. Geological Survey Open-File Report 95-511, 6 p.
Yerkes, R.F. and Campbell, R.H., 1995e, Preliminary geologic map of the Piru 7.5' quadrangle, southern California: a digital database: U.S. Geological Survey Open-File Report 95-801.
Yerkes, R.F. and Campbell, R.H., 1995f, Preliminary geologic map of the Simi 7.5' quadrangle, southern California: U.S. Geological Survey Open-File Report 95-828, 10 p.
Yerkes, R.F. and Campbell, R.H., 1995g, Preliminary geologic map of the Val Verde 7.5' quadrangle, southern California: U.S. Geological Survey Open-File Report 95-504, 9 p.
Yerkes, R.F. and Campbell, R.H., 1995h, Preliminary geologic map of the Val Verde 7.5' quadrangle, southern California: a digital database: U.S. Geological Survey Open-File Report 95-699.
Yerkes, R.F. and Campbell, R.H., 1997a, Preliminary geologic map of the Santa Susana 7.5' quadrangle, southern California: U.S. Geological Survey Open-File Report 97-258, 11 p.
Yerkes, R.F. and Campbell, R.H., 1997b, Preliminary geologic map of the Santa Susana 7.5' quadrangle, southern California: a digital database: U.S. Geological Survey Open-File Report 97-258.
Yerkes, R.F. and Campbell, R.H., 1997c, Preliminary geologic map of the Simi 7.5' quadrangle, southern California: a digital database: U.S. Geological Survey Open-File Report 97-259.

Fuzzy Sources, Maximum Likelihood, and the New Methodology

David M. Perkins
U.S. Geological Survey, MS 966
Federal Center, Box 25046
Denver, CO 80225
USA
perkins@ghtmail.cr.usgs.gov

Geophysicist, Central Geologic Hazards Team

Abstract

The new methodology (Frankel, and others, 1994; Frankel, 1995) is a logical generalization of existing developments in classical probabilistic hazard estimation—smoothing, allocation, and maximum likelihood.

Classic probabilistic ground motion hazard calculation uses, for areal seismic sources, polygonal zone sources, whose rates are determined by historical seismicity which has occurred within the zones. These zone sources usually have uniform areal seismicity rates up to their boundaries. The result is that hazard is nearly uniform throughout the interior, regardless of the actual historical pattern, and abruptly changes in the vicinity of the boundary.

Smoothing. To decrease this boundary abruptness, it is possible to provide for a "tapered" seismicity by using Gaussian smoothing of the source. This smoothing does not represent the uncertainty of the historical location, but rather the uncertainty of the location of future earthquakes in the vicinity of the historical seismicity.

Allocation. The more detailed and smaller the source zones are, the lower is the likelihood that enough historical seismicity exists for seismicity parameters ("a- and b-values") to be accurately determined for a single source. To avoid the large resulting uncertainty, the seismicity which has occurred in those zones which can be assumed to have the same b-value is combined into a superzone whose rate is such that seismicity parameters (particularly b-values) can be accurately determined. The resulting seismicity rate has to be "back-allocated" to the original source zones.

Maximum Likelihood. The allocation methods which can be envisioned are numerous, and all should work satisfactorily if the number of historical earthquakes per zone is very large. However, for the usual sparse sampling, the maximum likelihood method has performed best in simulated catalogs. This method has the unintuitive result that every earthquake provides the same amount of rate information for back-allocation, regardless of magnitude (above some minimum magnitude associated with some common percent level of completeness)!

Generalization. It is only a simple further step to use, as sources, the historical earthquake locations themselves, fuzzed with a Gaussian smoother. The resulting algorithm provides greater fidelity to the pattern information given by the historical seismicity, and avoids the subjectivity involved in drawing zone boundaries. The only adjustments remaining to the analyst are choice of minimum magnitude and the width of the Gaussian smoother.

Fuzzing the Classic Zone Source

The function of the classical zone source is to generalize the future location of seismicity similar to, and at the same rate as, that which has occurred at some historical rate in the zone. Zone sources are limited in size around some geological feature to the extent that recurrent seismicity is considered more likely to recur in the vicinity of that geological feature. (To the extent that fault information indicates seismicity whose recurrence is too low to have reasonable likelihood

of being experienced historically, these are modeled separately as features of a fault source.) Often seismicity itself is the sole basis for zone source delineation.

The classic zone source has been considered to be spatially uniform in future seismicity, regardless of magnitude. (Magnitude-based zone sources can be drawn, but I am not aware of their use in any analysis.) The uniformity of the future seismicity in the source leads to abrupt changes in hazard at the source boundaries. This abruptness is a function of ground motion or (equivalently) the return period of the map—the higher the ground motion value, the more closely the contour conforms to the zone source boundary.

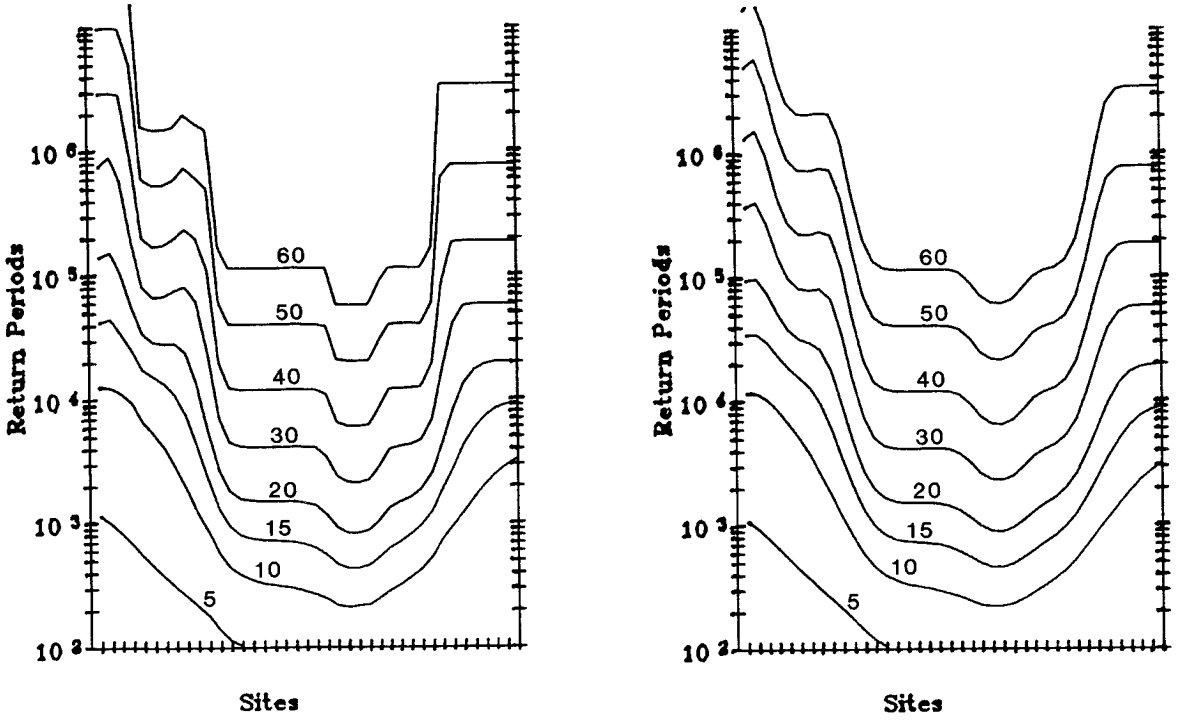

Figure 1. Return periods for various values of peak acceleration for a series of sites 10 km apart, traversing three seismic zone sources. On the left, no spatial smoothing. Note that return periods for high ground motion values change abruptly within 10-20 km. On the right, Gaussian smoothing using a standard deviation of 10 km.

This abrupt change in hazard associated with the zone boundary is not warranted, given the uncertainty with which the zone boundary has been selected. To mitigate this excess dependence on boundary location, Bender and Perkins (1987) implemented a way to express boundary uncertainty in their hazard algorithm, using a Gaussian smoother. In effect, the smoothing is achieved by assuming that each little sub-source has a likelihood of future earthquake occurrence in its vicinity, expressible by a two-dimensional Gaussian "spot." The net effect for a zone source is a "fuzzy" boundary

What is zone boundary uncertainty?

This uncertainty which is being modeled is not the uncertainty with which a seismic epicenter has been determined, but rather the uncertainty of future location of seismicity in the vicinity of past seismicity. To the extent that a zone source is drawn solely on the basis of a cluster of historical seismicity, the epicentral uncertainty should play a role, but inasmuch as a cluster of epicenters may have greater locational certainty than an individual epicenter, the boundary uncertainty could be less than the epicentral uncertainty. Nor is this uncertainty the spatial correlation between all events in a catalog. This latter uncertainty is dominated by the correlation of the most-frequent, lower-magnitude events. The desired uncertainty is the expected distance of the *next* earthquake having magnitude *larger* than the minimum magnitude and occurring *in the vicinity* of the given catalog event and *independent* of the occurrence of that event. It is a measure of the spatial relevance of a catalog event for defining a zone for future events.

What is the role of small-magnitude events?

The zone source models future seismicity for magnitudes in a range running from some minimum to some maximum magnitude. The minimum magnitude is chosen to be large enough to be significant for building damage (or some other purpose). What is the role of epicenters having lower magnitudes than that minimum magnitude? The role is two-fold. To some extent they provide rate information, although it is problematic as to whether they should be allowed tohave as dominating an effect as their number would give them. They are incompletely reported and also may not be good indicators of the Gutenberg-Richter slope of the larger-magnitude events. Their more important role, because of their number, is that they provide more spatial information for drawing zone boundaries, than do the larger-magnitude, less frequent events.

Allocating Regional Seismicity to Constituent Zones

Usually, in the magnitude range of interest, the number of events observed in each zone source is far too low to reliably define a Gutenberg-Richter b-value, and often too low to define a rate, given an uncertain b-value. It has been shown with artificial data that the accuracy of such fits depends strongly on the number of events available in the data. A minimum of 40 events per zone is recommended (see, for example, Thenhaus and others, 1980; Bender, 1983). Accordingly, instead of zonal seismicity, regional seismicity is analyzed in order to have sufficient number of events to obtain a more accurate b-value. Given the regional b-value, zone rates are then determined. There are a large number of candidate methods for making these zonal rate determinations. In our national hazard map work, we tested a number of techniques to perform this “back allocation.” The techniques which seemed most intuitive, physically, or regression-wise, actually performed very poorly on simulated data, as determined by the use of a number of effectiveness measures. The simplest technique, maximum likelihood, although at first glance unintuitive, performed best even on simulated data which violated its assumptions.

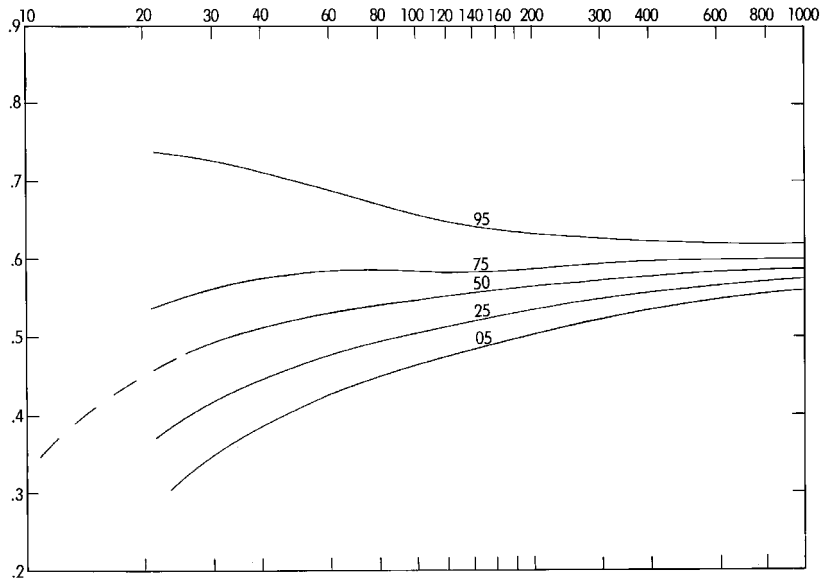

Figure 2. Fractiles of b-values obtained in simulated sets of Gutenberg-Richter data, as a function of set size.

Two Example Methods

In this section I want to contrast the performance of two of the candidate allocation methods on several hypothetical or simulated data arrays. The data showing number of events, by intensity category, for several zones, is presented in what I shall call a "tableau," to emphasize its hypothetical character. The columns of the tableau represent the number of observations in a particular intensity category for several different zones. The rows display the number of observations in each category for an individual zone. Beside each row is a calculated allocation fraction, the share each zone gets of the regional seismicity rate.

The first allocation method uses the ordinary least squares fit to the zone data, given a regional b-value. The allocation fraction for a zone is given by the relative number 10^a obtained for each zone, compared to the sum of all zone 10^a values. The table header abbreviation is "ava" for "average a." This method may be the "logical," intuitive method to use.

The second allocation method counts the observed number of events in the zone. The allocation fraction for a zone is given by the relative number of events obtained for each zone, compared to the total number of events for all zones. This result is maximum likelihood. (See equation A12 of Bender, 1983.) The table header abbreviation is ML for "maximum likelihood."

Likelihood Considerations

The following tableau has been designed to give the same allocations under the average a-value method. (The regional b-value is assumed to result in a factor of 3 decrease in rate with a unit increase in epicentral intensity.) Consider: If zone A were equivalent to zone B, we would have expected to observe 3 events of epicentral intensity VI. The chance of observing none, when we expect three is, under the Poisson assumption, $p(0) = e^{-3} = 0.05$. This is quite a low probability, arguing that the rate expected in that category could more likely be 2 (p=0.13) or 1 (p=.35). On the other hand, it is unlikely to observe 9 in category V if we expected 3, as we would if 1 were

the observed number in category VI. Thus zone A could not have 1/3 the rate of zone B, but might have 2/3 or less of the rate of zone B, and could not be of rate equal to B.

	Observed Number			Allocation Fraction	
Zone Name	V	VI	VII	ava	ML
Zone A	9	0	0	0.33	0.26
Zone B	9	3	0	0.33	0.35
Zone C	9	3	1	0.33	0.38

Similarly, comparing zone B and zone C, if we expect 1 in the intensity VII category, we could with the same probability have observed 0. On the other hand, zone C could have a slightly higher rate than B and still have observed the same number in the other categories. For zone index numbers governed by these considerations,
we might guess A : B : C = 2/3 : 1: 7/6 = 4 : 6 : 7, say.
Notice that the ML method conforms a good deal to these considerations. Hence, we might conclude that the "logical" method does not properly deal with likelihoods.

The Impact of an Unlikely Large Event

The next tableau has been designed to show the impact of one, relatively unlikely larger earthquake.

	Observed Number				Allocation Fraction	
Zone Name	V	VI	VII	VIII	ava	ML
Zone A	3	1	0	0	0.22	0.28
Zone B	3	1	1	0	0.32	0.36
Zone C	3	1	0	1	0.46	0.36

The isolated large event dominates the ava allocation, while having little or no effect on the ML method. It would seem desirable to provide a relatively lower weight to the higher magnitude events and greater weight to the lower magnitude events. This makes sense for the following reason. An ordinary least squares fit is appropriate if the variances of the observations being fit are the same for all values of the independent variable. If not, we should choose to equalize the variances by normalizing the data by the inverse of the square root of the variance (Draper and Smith, 1981, pp 108–1150.) Experiments have shown that for values of expected N less than 50, the apparent variance of log N is about proportional to $1/N$, as judged by the behavior of the 17th and 84th percentiles of the distribution of log N, when N has the Poisson distribution. The variance gets much larger than $1/N$ as N gets less than about 5, because of the rapidity with which logs of very small numbers approach negative infinity. The ML method does downweight the contribution of the rare, large events, compared to the ava, least squares fit.

Randomness in Sparse Zones of Equal Seismicity.

The final tableau shows the impact of random variation in distribution for zones having the same seismicity. Here I have selected 3 out of ten random draws from a simulation in which zones have the same rate and there is incompleteness in the lowest category. These 3 were selected so that the ML method would give identical weights (and in fact the true allocations should be equal). The result shows the excess variability given by the "ava" method.

	Observed Number				Allocation Fraction	
Zone Name	V	VI	VII	VIII	ava	ML
Zone A	1	3	0	0	0.26	0.33
Zone B	2	2	0	0	0.29	0.33
Zone C	1	2	0	1	0.45	0.33

These examples are selected from a much larger collection of examples and possible allocation methods in order to make it reasonable that the unintuitive (just-count-the-earthquakes) maximum likelihood method should perform better than the ordinary least squares fit method. Other methods can do almost as well, for instance, the minimum chi-square method. In discussing the relative merits of maximum likelihood estimators and minimum chi-square estimators, Kendall and Stuart (1973, p. 97–99) conclude that because both estimators have the same asymptotic properties, the choice between them must be decided on the basis of computational convenience and small-sample properties. They cite papers suggesting that maximum likelihood estimators should have the better small-sample properties. I have discovered in simulations that the same is true in simulations of sparse zone-category data. And as to computational convenience, in our application it would be difficult to find a simpler estimator than the maximum likelihood estimator.

The Optimality of the Maximum Likelihood Allocation Method

We have seen that the maximum likelihood allocation method is

1. Optimal for low-rate zones, and
2. Simple

It is useful to note that the method has three additional properties which make it desirable in the context of seismic allocation:

3. The method is insensitive to size error—it doesn't matter what intensity category the event falls in, the contribution is the same.
4. Because of the magnitude insensitivity, there is only a minor sensitivity to location error. If an event seems to fall in one zone instead of another, it moves only a contribution of 1/N of the total seismicity from one zone to the other.
5. The method is independent of b-value. In the ordinary least squares fit, the allocation weight given by a high-magnitude event depends upon the b-value.

Generalization to Epicenters as Sources

In the maximum likelihood allocation method for zones, each observed event carries a contribution of 1/N to the allocation given its zone (where N is the total number of events observed—aftershocks and other dependent events excluded). This is true even for incomplete events as long as all events of a given magnitude range have the same incompleteness, spatially. What then is to prevent us from allocating future events to each epicenter as a source, where this source can have all the magnitudes from the minimum to the maximum, and for which the relevance for future seismicity in the vicinity is given by a two-dimensional Gaussian uncertainty? This is equivalent to drawing a small zone source around each epicenter, obtaining a regional seismicity a- and b-value, and then back-allocating into the N costituent zones. We have already motivated the conclusion that the maximum likelihood method is optimal for back-

allocation. The only remaining issue is what the shape of the zone should be around each epicenter.

We can imagine a number of choices, a mini-square, a mini-rectangle oriented in some direction to indicate anisotropic uncertainty (perhaps along some regional structural trend), a mini-disc, an oriented mini-ellipsoid, etc. We have chosen in the new methodology an isotropic Gaussian "blob," which represents an un-oriented uncertainty.

Choice of Smoothing Parameter

What standard deviation should be used for the Gaussian "fuzzing" parameter? I have already suggested using some number reflecting the uncertainty of a zone boundary, if one were to be used in the absence of notable seismogenic geological structure. This would ordinarily be on the order of a few tens of kilometers.

Another possibility would be choosing the standard deviation so that the resulting hazard map is smooth, that is, not showing local small closed contours representing individual events. In the eastern United States, experiments have shown this can be achieved with standard deviations on the order of 30 to 60 km. Some estimate of this parameter can be made by looking at the characteristic spacing of the larger events in what would otherwise be designated as zones.

The smoothing effect depends on the number of events observed. For many events, a small smoothing parameter can be used. For few events, a larger parameter is required to obtain smoothness. Caution must be used, however. In our experiments, larger standard deviations, more than 100 km, have produced artifactual seismicity where no seismicity had been experienced and where it is unreasonable to expect it. In one case, seismicity in an arc far to the east and south of a site, combined to produce artificial seismicity in the vicinity of the site.

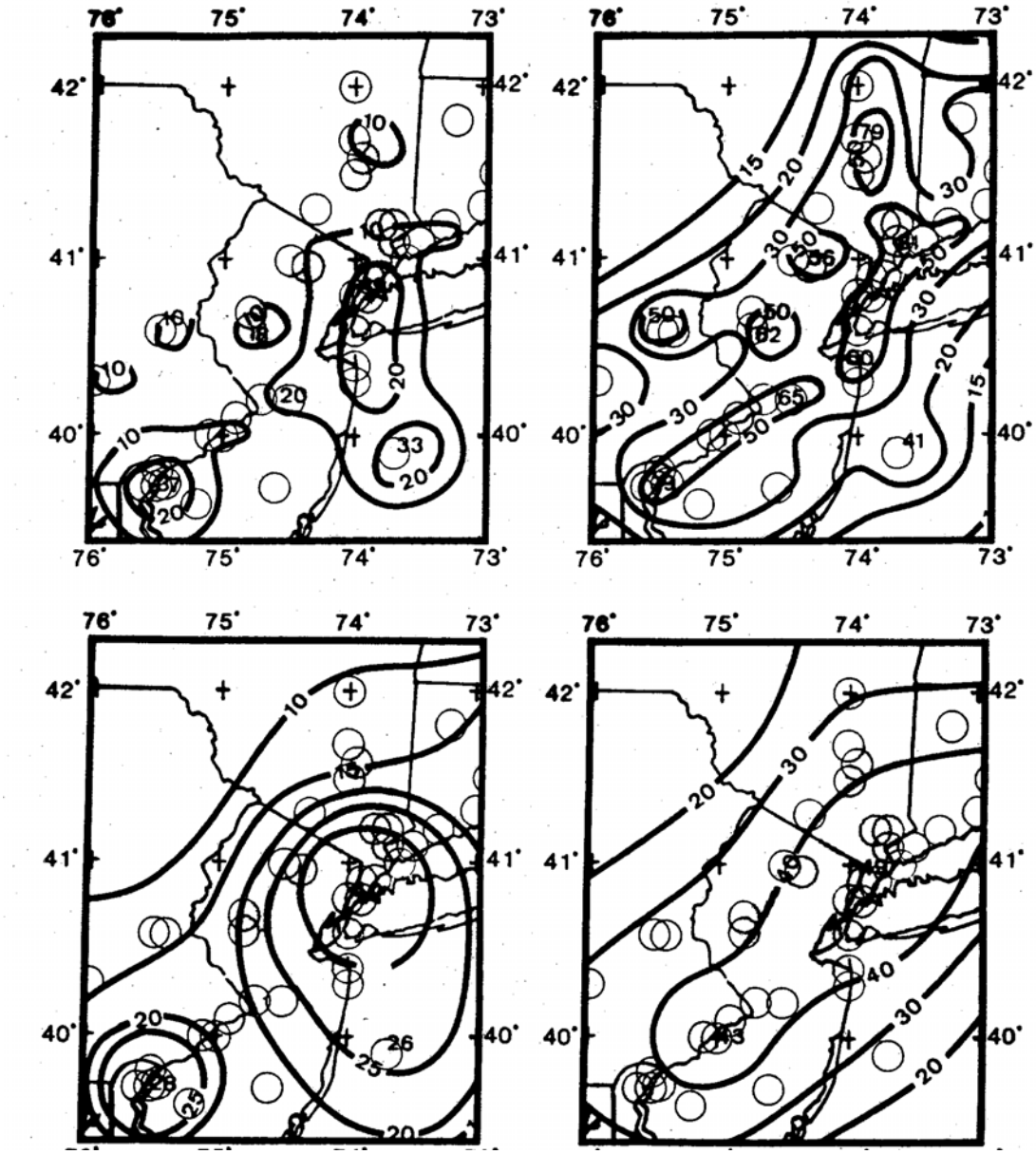

Figure 2. Hazard maps produced by (UL) using past earthquake locations and magnitudes, (UR) allowing Gutenberg Richter (GR) occurrences at past locations, (LL) using past earthquake magnitudes but smoothing past locations by 40 km standard deviation, (LR) allowing GR occurrences and smoothing past locations by 40 km.

Discussion

To criticize the new methodology, it is necessary to question the accepted basis for assigning fitted past seismicity to the models for future seismicity. Although one recurringly hears the comment that the next damaging earthquake will be where none have been experienced before, for the most part this is not so. To the extent that there could be surprise future damaging earthquakes, no one has yet found a good predictive method for anticipating them in the context of hazard mapping. Furthermore, no one has seriously proposed basing a hazard map on *absence* of seismicity.

It is helpful to recall why it is reasonable to use historical seismicity spatially to anticipate future seismicity.

1. Past seismicity tells us where earthquakes can happen. It is reasonable to design for recurrence. ("Railroads run on tracks, earthquakes occur on faults")
2. Some seismicity consists of anticipatory, long-term foreshocks.
3. There are many examples of major cities around the world which have experienced not only recurring events over several hundred years interval, there are also examples of areas which, though ordinarily not always active, once seismicity has started, seismicity has continued for several damaging events.
4. Even where seismicity has occurred and is declining, it is reasonable to expect adjacent areas to have received an increase of seismic stress. This implies that a "neighborhooding" (spatial-generalization) algorithm is suitable.

The new methodology has been shown to be merely an extrapolation of the existing classical zone source methodology—using the maximum likelihood allocation method as a basis for reducing each suitably selected epicenter as a mini-source indicator for future seismicity. (By suitably selected I mean the usual—dependent events are removed from the catalog; smaller events which do not have spatially similar completeness across a region are also removed.) It has been successfully used in the production of the latest national probabilistic ground motion hazard maps for the United State (Frankel, and other, 1996).

The method has a number of advantages over the usual zone source methodology

1. A significant amount subjectivity has been removed from the hazard estimation—the subjectivity of drawing zone source boundaries. In the eastern US, where there is considerable seismicity, as well as different candidates for zonation methodologies, the results of expert solicitation estimates of alternative zone sources for probabilistic hazard at nuclear sites has produced a wide range of hazard estimates for the same site, largely attributable to differences in zonation (see Perkins, and others, 1988). In studying a portion of this geographic area, we have found that different nuclear site hazard estimates obtained from a smoothing method do not vary as much as the estimates obtained by alternative expert zonation. (This suggests that the process of zonation "imports" unnecessary variability into the hazard estimation!)
2. The method also encourages the use of regional seismicity fits and back-allocation—a much more accurate method of b-value and rate estimation. The method particularly avoids fitting seismicity to sparse data sets, for which the usual ordinary least squares method is particularly badly suited. In regional fits, where the seismicity is ample, the difference in methods is not so important.
3. The method provides spatial non-uniformity in rate, which is more reasonable than the spatially uniform rate imposed by a zone source.

The classical zone source can still be used where there is considerable tectonic information to restrict spatial uncertainty of sources more than this new methodology would. It is also hoped that anisotropic smoothing can be assessed as a mediating method between the classic and the new methodology.

References

Bender, Bernice, 1983, Maximum Likelihood Estimation of B-values for Magnitude Grouped Data, Bulletin of the Seismological Society of America, v. 73, n. 3, p. 831–851.

Bender, Bernice K., and Perkins, David M., 1987, Seisrisk III, A Computer Program for Seismic Hazard Analysis: U.S. Geological Survey Bulletin 1772, 48 p., including program listing.

Draper, N.R., and Smith, H., 1981, Applied Regression Analysis, 2nd ed., John Wiley and Sons, New York, 709 p.

Frankel, Arthur, Thenhaus, Paul, Perkins, David, and Leyendecker, E.V., 1994, Ground Motion Mapping—Past, Present and Future, *in* Proceedings of the Seminar on New Developments in Earthquake Ground Motion Estimation and Implications for Engineering Design Practice, ATC 35-1, Applied Technology Council, Los Angeles, CA, January 26, 1994, p. 7-1–7-40.

Frankel, A. (1995). Mapping Seismic Hazard in the Central and Eastern United States, Seism. Res. Letts, v.,66, no. 4, pp. 8-21.

Frankel, A., Mueller, C., Barnhard, T., Perkins, D., Leyendecker, E., Dickman, N., Hanson, S., and Hopper, M., 1996, National seismic-hazard maps: documentation June 1996, USGS Open-file Report 96-532, 110 p.

Kendall, Maurice G., and Stuart, Alan, 1973, The Advanced Theory of Statistics, v. 2, Inference and Relationship, Hafner Publishing Company, New York, 723 p.

Perkins, David M., Bender, Bernice K., and Thenhaus, Paul C., Review of Seismicity Owners Group-Electric Power Research Institute Seismic Hazard Methodology, Report to the U.S. Nuclear Regulatory Commission, July 25, 1988, 83 p. (Appears as Appendix B in the U.S. Nuclear Regulatory Commission "Safety Evaluation Review of SOG/EPRI Report, "Seismic Hazard Methodology for the Central and Eastern United States" (EPRI NP-4726)).

Thenhaus, Paul C., Perkins, David M., Ziony, Joseph I., and Algermissen, S.T., 1980, Probabilistic Estimates of Maximum Seismic Horizontal Ground Motion in Rock in Coastal California and the Adjacent Outer Continental Shelf, U.S. Geological Survey Open-File Report OF 80-924, 69 p.

Estimation of Volcanic Hazards Related to Tephra Fallout

Charles B. Connor, Brittain E. Hill, Brandi Winfrey, Nathan M. Franklin and Peter C. La Femina

March 22, 2000

Center for Nuclear Waste Regulatory Analyses
Southwest Research Institute
6220 Culebra Road
San Antonio, TX 78235-5166 USA
cconnor@swri.edu

Abstract

The goal of probabilistic volcanic hazard assessment is to translate complex volcanological data and numerical models into practical hazard estimates for communities potentially affected by volcanic eruptions. Probabilistic volcanic hazard assessment quantifies volcanic hazards and illustrates uncertainties about the magnitude and consequences of volcanic activity. Planning based on probabilistic volcanic hazard assessment has the potential of mitigating the effects of volcanic eruptions when they occur. We develop an approach to estimate volcanic hazards related to tephra fallout, and illustrate this approach with a tephra fallout hazard assessment for the city of León, Nicaragua and the surrounding area. Tephra fallout from eruptions of Cerro Negro volcano has caused damage to property, adverse health effects, and has disrupted life in this area. By summarizing the geologic and historical records of past eruptions of Cerro Negro on a probability tree, we show that the inhabitants of León can expect $> 1cm$ tephra accumulation from approximately 30% of eruptions, and $> 4cm$ tephra accumulation from approximately 9% of eruptions of Cerro Negro volcano. This historical record is augmented with simulations of tephra dispersion which estimate the likelihood of tephra accumulation given a range of eruption magnitudes and map the expected distribution of tephra over a broader region. An upper limit value of $0.5m$ is calculated using the tephra dispersion model. Without a fundamental change in the eruptive behavior of Cerro Negro, tephra accumulation in León is not expected to exceed this value.

1 Introduction

Planners and the communities they serve need to know what to expect from volcanic eruptions. Volcanologists can provide the information required to develop strategies for volcanic hazard mitigation and to weigh the relative costs of mitigation efforts. Information about volcanic hazards is best presented in the form of probabilistic hazard assessments, combining elements of the geologic record, numerical simulation of volcanic eruptions, and an accurate picture of the natural uncertainty in estimates of eruption magnitude and timing. Here, we outline steps for a probabilistic hazard assessment of one aspect of volcanic eruptions: tephra fallout. Other phenomena that impact communities near volcanoes, such as pyroclastic flows and lahars, may be treated in a similar fashion, but require separate hazard analyses.

Tephra fallout results from explosive volcanic activity. Tephra, colloquially referred to as volcanic ash, consists of pyroclasts produced during volcanic eruptions and accidental lithic fragments incorporated in the eruption. Tephra fallout results when tephra is carried aloft in a volcanic eruption column and subsequently is

deposited by sedimentation out of the volcanic plume, sometimes at great distances from the volcano (Figure 1) (Fisher and Schmincke, 1984; Sparks et al., 1997).

Magnitudes of volcanic eruptions are classified using the volcano explosivity index (VEI)(Newhall and Self, 1982; Simkin and Siebert, 1994), which primarily reflects eruption volume and the height of the eruption column. Large eruptions (VEI 4-6), such as the May 18, 1980, eruption of Mount St. Helens, and the June 15, 1991, eruption of Mount Pinatubo, produce significant tephra accumulation at great distances from erupting volcanoes, resulting in major disruptions for society (e.g., Sarna-Wojcicki et al., 1981; Punongbayan et al., 1996; Koyaguchi, 1996). Even moderately-sized volcanic eruptions (VEI 3) can significantly affect areas $> 10km$ from the volcano due to tephra fallout (e.g., Hill et al., 1998). Tephra fallout can cause building collapse, disruption of power and water supplies, damage to mechanical systems such as vehicle engines, and widespread damage to agricultural products, including livestock. Although the consequences of tephra fallout may be less severe than other eruption phenomena, vulnerability is often much greater due to the wide dispersal of tephra in volcanic plumes. In some cases, the continual remobilization of tephra deposits by wind or surface-disturbing activities has resulted in respiratory ailments and related deleterious health effects (Baxter, 2000). In recent years, attention has also focused on the harm to aircraft caused by tephra, and the risks associated with such damage for airline passengers and cargo (Miller and Casadevall, 2000).

Our purpose is to describe probabilistic methods for assessment of tephra fallout hazards. We focus on tephra accumulation within tens of kilometers of the volcano, where tephra accumulation is most likely to cause damage. These methods are based on developing a conditional probability of tephra fallout, conditioned on the occurrence of eruptions. Thus, hazard assessments can be prepared in advance of episodes of volcano unrest and mitigation strategies can be developed before volcanic crises occur. In this sense, long-term planning for tephra fallout is a practical goal. To achieve this goal, it is necessary to combine geologic observations of past patterns of tephra accumulation with results of numerical simulations of tephra fallout. These results should account for the natural variability in the physical processes resulting in tephra fallout, and the uncertainty related to our models and observations of these phenomena. Steps in tephra fallout hazard estimation are described in the following, with special reference to Cerro Negro volcano, Nicaragua, a small-volume basaltic cinder cone that experiences comparatively frequent eruptions accompanied by tephra fallout (Figures 2 and 3), and where we have implemented these steps previously (Hill et al., 1998).

Figure 1: Eruptions of Cerro Negro volcano, Nicaragua. The 1968 eruption (left), from USGS files; the 1995 eruption (top right); and damage in León resulting from the 1992 eruption (bottom right).

2 Treating Tephra Fallout as a Conditional Probability

Conditional probabilities enable volcanologists to consider a complicated series of events that typically comprise a volcanic eruption discretely. In the case of tephra fallout, it is simpler to estimate expected tephra accumulation, given eruption conditions, than to combine this disparate information into a single hazard forecast. For example, the geologic record or numerical simula-

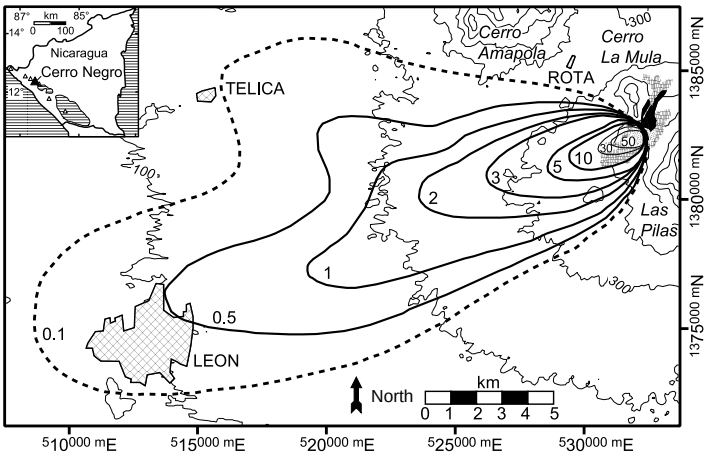

Figure 2: Isopach map for the 1995 eruption of Cerro Negro. Tephra thickness shown in cm.

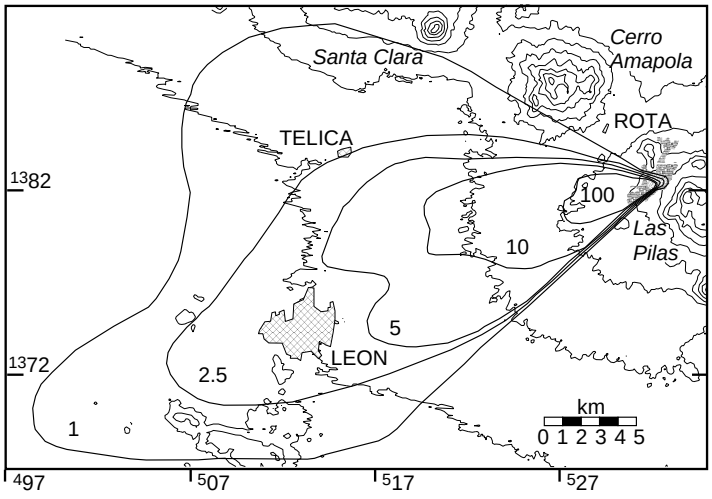

Figure 3: Isopach map for the 1992 erupton of Cerro Negro. Tephra thickness shown in cm.

tions can be used to infer P[tephra accumulation $> 1cm$ at $(x, y) \mid$ VEI 2], the probability that tephra accumulation will exceed $1cm$ at a given location, (x, y), given that a volcanic eruption of intensity VEI 2 has taken place. The probability of an eruption occurring at all, or the probability that an eruption will follow a period of unrest, are treated as separate issues. Once the occurrence probability is determined, the unconditional probability of tephra fallout can be easily estimated.

Application of conditional probabilities is important for two reasons. First, volcanoes often experience prolonged periods of unrest, during which time the probability of an eruption changes considerably. Hazard assessments of tephra fallout for a given eruption made before the volcano unrest began should not require significant modifications during episodes of unrest, when attention is focused on monitoring. Second, the geologic record often provides information about the magnitude of eruptions, for instance the thickness of tephra accumulated during past eruptions at a particular location, but comparatively little information about the timing and frequency of eruptions.

Conditional probabilities are easily visualized using event trees, which illustrate the observer's ability to predict outcomes (Schafer, 1996). Often in volcanology, our ability to assign conditional probabilities to all possible outcomes is limited by lack of experience or lack of relevant information. In these circumstances, the trees may be only partially probabilized. We have constructed an event tree for tephra fallout from Cerro Negro volcano (Figure 4) as an example of their application in volcanology.

Cerro Negro has erupted 23 times since the volcano first formed in 1850 (McKnight, 1995; McKnight and Williams, 1997; Hill et al., 1998). Many of the early eruptions of Cerro Negro are poorly documented, but there is a reasonably complete record of tephra fall volumes since 1900. Since 1968, four eruptions have occurred at Cerro Negro (Table 1) that produced tephra fall volumes $> 1 \times 10^6 m^3$, and three smaller eruptions have occurred. The most recent of these small eruptions was in August, 1999 (Hill et al., 1999; La Femina et

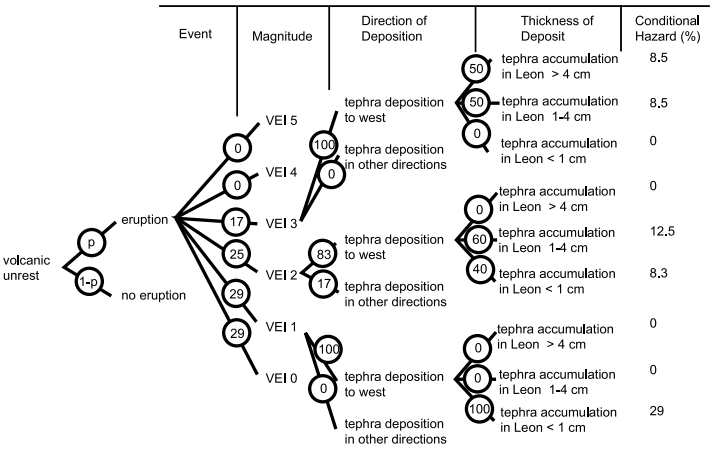

Figure 4: Event tree for tephra accumulation in León. Probabilities (in percent) based on the historical and geologic records of past Cerro Negro eruptions.

al., 1999). Based on a steady-state model of cumulative volume of material erupted from Cerro Negro, a volcanic eruption is expected before 2005 with 95% confidence (Hill et al., 1998; 1999). Many, but not all, of the eruptions of Cerro Negro have resulted in tephra deposition in the city of León, the second largest city in Nicaragua (population $> 200,000$), located $20km$ from Cerro Negro. In addition, an estimated 100,000 people live in the area surrounding Leon, with active agricultural communities located as close as 1 km from Cerro Negro. Tephra deposition in León has varied from trace amounts to $4cm$ in 1992 (Figure 3)(Connor et al., 1993). In the 1992 eruption, at least 2 people were killed and 146 injured through building collapse, over 12,000 people were evacuated, and $19M of crops and infrastructure destroyed (Organización Panamericana de la Salud, 2000). Rates of respiratory and intestinal diseases increased by factors of 4 to 6 in the month following the 1992 eruption (Malilay et al., 1997). The significantly smaller 1995 (Figure 2) eruption still displaced 1,200 people and destroyed $0.7M of crops and infrastructure (Organización Panamericana de la Salud, 2000). Future tephra eruptions from Cerro Negro clearly present a significant hazard to people in this area.

Conditional probabilities are assigned to the branches of the event tree (Figure 4) based on this record of past volcanic activity (Table 1). The event tree is incomplete because there is no probability assigned to the transition from volcanic unrest to volcanic eruption. There is simply inadequate data available to make this assessment at Cerro Negro. Given a volcanic eruption, there is a reasonably equal probability of VEI 0-2 and a slightly lower probability of VEI 3 than VEI 2 eruptions. No eruptions of VEI > 3 have occurred at Cerro Negro. The VEI 0 eruptions do not produce significant tephra fallout, except very close to the volcano, and are not considered further on the tree. Nearly all of the explosive eruptions that have occurred (VEI $1-3$), have resulted in deposition of tephra west of the volcano, toward the city of León. The tree is branched further into tephra accumulation $< 1cm$, $1-4cm$, and $> 4cm$. This bifurcation reflects both the way past tephra accumulation has been reported in León (McKnight, 1995; Hill et al., 1998) and the severity of the consequences of the tephra accumulation.

Based on the geologic and historical records of volcanic eruptions at Cerro Negro, P[tephra accumulation $> 4cm \mid$ volcanic eruption] = 8.5% and P[tephra accumulation $> 1cm \mid$ volcanic eruption] = 29.5% (Figure 4). So, although the record is limited (a ubiquitous circumstance in volcanology), the event tree provides a concise summary of the nature of past volcanic activity and its consequences for the city of León.

3 Modeling Tephra Fallout Numerically

The goal of modeling tephra fallout using numerical simulations is to provide a path from the geological and historical records of tephra accumulation to estimation of parameter distributions, and ultimately to probabilistic estimates of tephra fallout hazard. Why are these additional steps necessary? Numerical simulation of geologic phenomena is a complex undertaking and modeling tephra fallout is certainly no exception. It would be ideal if the geologic and historical records provided an adequate picture of the frequency distribution of tephra

Year (A.D.)	Duration (days)	Fall Volume (m^3)	Fall Volume DRE (km^3)	Lava Volume DRE (km^3)	Column Height (km)	VEI
1850	10	4.3×10^5	0.0002	0.0054		1
1867	16	7.4×10^6	0.0034	N.A.	3	2
1899	7		tr F	N.A.		1
1914	6	2.8×10^6	0.0013	N.A.		2
1919	10		tr F	N.A.		0
1923	49	1.7×10^7	0.0077	0.0100	2	2-3
1929	19		tr F	0.0001		0
1947	13	2.3×10^7	0.0110	0.0038	6	3
1948	1+		tr F	N.A.		0
1949	1+		tr F	N.A.		0
1950	26	2.8×10^6	0.0013	0.0001	1.5	2
1954	1+		tr F	N.A.		0
1957	20	2.8×10^6	0.0013	0.0045	2	2
1960	89	1.1×10^6	0.0005	0.0052	1	1
1961	1+		tr F	N.A.		0
1962	2		tr F	N.A.		1
1963	1+		tr F	N.A.		0
1968	48	9.7×10^6	0.0045	0.0069	2	2
1969	10		tr F	N.A.		0
1971	10.6	3.0×10^7	0.0139	tr L	4 - 6	3
1992	3.6	2.3×10^7	0.0110	N.A.	3 - 7	3
1995	79.0		tr F	tr L		1
1995	13.0	2.8×10^6	0.0013	0.0037	2 - 2.5	2
1999	3				1	1
Total	436		0.0574	0.0397		

Table 1: Cerro Negro Eruption Data

Note: DRE is Dense Rock Equivalent

N.A. = no lava flows reported or very small volume
tr F = little appreciable tephra fallout occurred or was likely
tr L = very little volume lava flows reported

fallout, and if, as a consequence, we could have a great deal of confidence in hazard estimates solely based on this record. Unfortunately, this is not the case. The geologic and historical records are insufficient as the sole basis for hazard assessment for several important reasons.

First, the geologic record is invariably sparse. In applying probabilistic techniques, we assume that there are parameters, such as mean and standard deviation, that describe the frequency distribution of tephra accumulation in a given area. Given a large number of events, we would expect the values of these parameters to converge toward a fixed limit. But with few events (e.g.,< 20 well-documented eruptions of Cerro Negro), there is no reason to believe with confidence that the frequency distribution of outcomes (e.g., thickness of tephra in León) is well represented by the geologic record, or that the limiting values of the parameters that characterize the frequency distribution are known (e.g., von Mises, 1953). Probabilistic hazard assessments based on poorly known frequency distributions are highly uncertain.

Second, the geologic record is biased toward large events because these are more likely to be preserved. Smaller, generally more frequent, eruptions may leave no discernable geologic record. Conversely, the historical record may also be biased. Because of its brevity, the historical record often poorly reflects the full range of activity a given volcano has experienced. At Cerro Negro, given its short history, it is unlikely that the full range of potential eruptions is captured in the range of eruptions that have already occurred.

The disparity between historical and geologic records is particularly evident at Colima volcano, Mexico, where there is nearly complete disjunction between the historical and the stratigraphic records. Dozens of explosive eruptions have been recorded in the last 400 years, but only a single scoria layer has been preserved in the geological record for that period (Navarro and Luhr, 2000). It is clear that at Colima volcano the preserved sequence includes just a fraction of the eruptive events, biasing the geologic record toward large eruptions. Nevertheless, the style of eruptive activity may have also changed over time. Modeling these data and estimating parameter distributions that best fit the data helps resolve this disparity. For example, if complex, multi-modal parameter distributions are needed to model the historical and geologic data sets together, nonstationary behavior in eruption style may be indicated. Alternatively, if comparatively simple parameter distributions model both historical and geologic data sets, then the brevity of the historical record may account for the disparity. This disjunction between geologic and historical data sets is a classic situation and numerical modeling provides insight into its origin and implications for probabilistic volcanic hazard assessment.

Third, hazard estimates based on the geologic record are difficult to update, as more information about changing conditions becomes available. Although not the focus of this paper, if the goal is to update probabilistic volcanic hazard forecasts in near-real-time, a model of tephra fallout is required (Carey, 1996).

3.1 General Model Description

The overall benefit of numerical simulation lies in the promise of improved ability to quantify the expected variability in the volcanic system, and consequently improve hazard estimates. Development of tephra fallout models have been recently reviewed by Carey (1996), Sparks et al. (1997), and Rosi (1998). Here we illustrate the integration of tephra fallout models into a probabilistic volcanic hazard assessment using the model developed by Suzuki (1983), and subsequently modified and applied to volcanic eruptions by Armienti et al. (1988), Glaze and Self (1991), Jarzemba (1997), and Hill et al. (1998). Suzuki's model is empirical; the erupting column is treated as a line-source reaching a maximum height governed by the energy and mass flow of the eruption. A linear decrease in the upward velocity of particles is assumed, resulting in segregation of tephra particles in the ascending column by settling velocity, which is a function of grain size, shape,

and density. Tephra particles are removed from the column based on their settling velocity, the decreasing upward velocity of the column as a function of height, and a probability density function that attempts to capture some of the natural variation in the parameters governing particle diffusion out of the column. Dispersion of the tephra that diffuses out of the column is modeled assuming a uniform wind field and is governed by the diffusion-advection equation with vertical settling. Thus, this model relies on estimation of numerous parameters that describe the volcanic eruption and the atmosphere it penetrates. Although not as comprehensive in addressing the physics of the eruption column (c.f., Woods, 1988; 1995), the computational ease of Suzuki's (1983) approach makes it a worthwhile method for hazard assessment, especially in light of the practical difficulties inherent in characterizing the variability in eruption and meteorological parameters (e.g., GVN, 1999).

This model abstracts the thermo-fluid-dynamics of ash dispersion in the atmosphere using the following expression:

$$X(x,y) = \int_{\phi_{min}}^{\phi_{max}} \int_0^H \frac{5 f_Z(z) f_\Phi(\phi) Q}{8\pi C(t+t_s)^{5/2}} exp[-\frac{5((x-ut)^2+y^2)}{8C(t+t_s)^{5/2}}]dz d\phi \quad (1)$$

where: X is the mass of tephra accumulated at geographic location (x, y) relative to the position of the volcanic vent (x increases in the downwind direction and y is orthogonal to this direction); $f_Z(z)$ is a probability density function for diffusion of ash out of the eruption column, treated as a line-source extending vertically (z) from the vent to total column height, H; $f_\Phi(\phi)$ is a probability density function for grain size, ϕ; Q is the total mass of material erupted; u is wind speed in the x-direction; t is the tephra particle fall time, and t_s is the tephra diffusion time; and C is eddy diffusivity in the atmosphere.

In our implementation, we assume that tephra particle diameter is distributed normally in phi units (ϕ) (e.g., a grain diameter of $1\phi = -\log_2(0.5mm)$):

$$f_\Phi(\phi) = \frac{1}{\sqrt{2\pi}\sigma} exp[-\frac{\phi-\mu^2}{2\sigma^2}] \quad (2)$$

where μ and σ are the mean and standard deviation of particle diameter (Fisher and Schmincke, 1984). The tephra particle fall-time in the atmosphere is given by:

$$t = 0.752 \times 10^6 [1 - exp(-\frac{0.0625z}{v_0(\phi)})]^{0.926} \quad (3)$$

(Suzuki, 1983), where z is height and v_0 is the settling velocity, which depends on particle diameter:

$$v_0(\phi) = \frac{\rho_{ash} g \phi^2}{9\eta p_f^{-0.32} + \sqrt{81\eta^2 p_f^{-0.64} + 1.5\rho_{ash}\rho_{air} g \phi^3 \sqrt{1.07 - p_f}}} \quad (4)$$

where ρ_{ash} is the density of tephra particles, here taken to be independent of grain size; g is gravitational acceleration; η and ρ_{air} are air viscosity and density, respectively; and p_f is a particle shape factor:

$$p_f = \frac{p_b + p_c}{2p_a} \quad (5)$$

where subscripts a, b, and c refer to the diameter of the particle along its principle axes and $p_a > p_b > p_c$. The complexity of the expression for settling velocity stems from the complex relationship between drag on the particle and the particle Reynolds number:

$$Re = \frac{v_0(\phi)\phi}{v_{air}} \quad (6)$$

where v_{air} is the kinematic viscosity of air. At low Re, the settling velocity of particles obeys Stokes law. At higher Re, however, there is a transition to inertial flow and finally turbulent boundary flow (Bonadonna et al., 1998). Equation 4 is only one approximation to these

conditions (Suzuki, 1983; Armienti et al., 1988; Sparks et al., 1997; Bonadonna et al., 1998).

In addition to settling, tephra particles diffuse in the atmosphere as they advect downwind. The diffusion time in equation 1 is given by Suzuki (1983) as:

$$t_s = \left[\frac{5z^2}{288C}\right]^{5/2} \tag{7}$$

derived from the assumption that the standard deviation of the tephra plume width is one-third the height (c.f., Woods, 1988).

The probability density function $f_Z(z)$ describes the diffusion of tephra out of the erupting column and into the surrounding atmosphere where it is available for downwind transport:

$$f_Z(z) = \frac{\beta w_0 Y(z) e^{-Y(z)}}{v_0(\phi) H (1 - (1 + Y_0)e^{-Y_0}))} \tag{8}$$

and:

$$Y(z) = \frac{\beta w(z)}{v_0(\phi)} \tag{9}$$

$$Y_0 = \frac{\beta w_0}{v_0(\phi)} \tag{10}$$

$$w(z) = w_0 \left[1 - \frac{z}{H}\right]^{\lambda} \tag{11}$$

where w_0 is the initial eruption velocity at the vent. The parameter β controls the shape of function $f_Z(z)$. Larger values of β result in a greater proportion of tephra reaching high in the eruption column. The upward velocity of particles $w(z)$ is assumed to decrease with height as a function of the parameter λ, where for most applications $\lambda = 1$. Equations 9 - 11 differ slightly from those provided in Suzuki (1983) and are written to conserve mass in the eruption column.

In practice, eruption column height, H, total eruption volume, V, and eruption duration, T, are the best known parameters for a given eruption. Walker et al. (1984) applied:

$$\frac{dV}{dt} = \left[\frac{H}{1.67}\right]^4 \tag{12}$$

to estimate the mass eruption rate, where H is in kilometers and dV/dt is in m^3 (dense rock equivalent). This application assumes that the eruption rate is constant over the duration of the eruption, i.e.,

$$V = \frac{dV}{dt} T \tag{13}$$

These relations provide a check on input parameters used in the Suzuki (1983) model.

3.2 Tephra Fallout Hazard Curves for Cerro Negro Volcano

In practice, the model is calculated numerous times using a range of input parameters (e.g., total mass, column height, etc.) that reflect the historical range of activity at Cerro Negro. Each realization represents a possible combination of eruption style, magnitude, duration, and atmospheric conditions that control the amount and distribution of tephra fallout. Using this stochastic approach, the range of likely eruption styles (VEI 0-3) and their consequences for tephra fallout hazard are evaluated.

Successful application of the model relies on well-chosen input parameters. Our experience is that model results are strongly controlled by eruption volume, column height, and eruption velocity. Eruption volumes were sampled from a log-uniform random distribution, with tephra volumes ranging from 5×10^5 to $1 \times 10^8 m^3$, based on estimates of the volumes of past eruptions. Column height was sampled from a uniform random distribution, U[$2km$, $8km$]. The relationships between column height, mass flow, eruption duration, and total eruption volume (e.g., equations 12 - 13) were used

to check these input parameters. Eruptions with very long durations, > 120 days, were eliminated from the sample set because long duration explosive activity at Cerro Negro has never occurred. Eruption velocity was sampled from a uniform random distribution U[$50m/s$, $100m/s$], independent of mass flow. Wilson and Head (1981) suggested a relationship between eruption velocity and mass flow that, for Cerro Negro, yields eruption velocities much lower than the $75m/s$ to $100m/s$ velocities that have been observed (Connor et al., 1993; Hill et al., 1998) or the eruption velocities predicted for buoyant tephra columns at low mass flow rates, $1 \times 10^6 kg/s$ (Woods and Bursik, 1991).

Wind speed and direction also have a strong influence on the tephra accumulation. At Cerro Negro, tradewinds are quite consistent. Average monthly wind direction is to the west or west southwest, and average wind speed on the ground is 4 to $6m/s$, with maximum wind speeds of $15m/s$ measured during most months (National Climatic Data Center, 2000). Furthermore, with the exception of the 1914 eruption (McKnight, 1995), all significant tephra deposition has occurred west of Cerro Negro, with León lying near the major axis of dispersion for the 1968, 1971, 1992, and 1995 eruptions (Figures 2 and 3). Therefore wind speed was sampled from a U[$5m/s$, $15m/s$] distribution and wind direction from a U[$-5°$, $-35°$] distribution, where wind direction is with respect to due west. In the absence of meteorological observations, we assume that these ranges apply to wind data at higher elevations. This assumption at least works well for the 1995 eruption, during which wind speeds in the tephra plume were $9m/s$ (Hill et al., 1998).

Grain size distribution was integrated from -5 to 5ϕ, with mean grain size = -1ϕ and standard deviation = 1ϕ. Clast density was varied as U[0.9, $1.2gm/cm^3$], equant grains were used ($p_f = 0.5$), and $\beta = 0.5$. This value for β forces a higher proportion of erupted pyroclasts to elevations close to the top of the eruption column, a choice consistent with field observations and thermo-fluid-dynamic models of volcanic plumes (Woods, 1995).

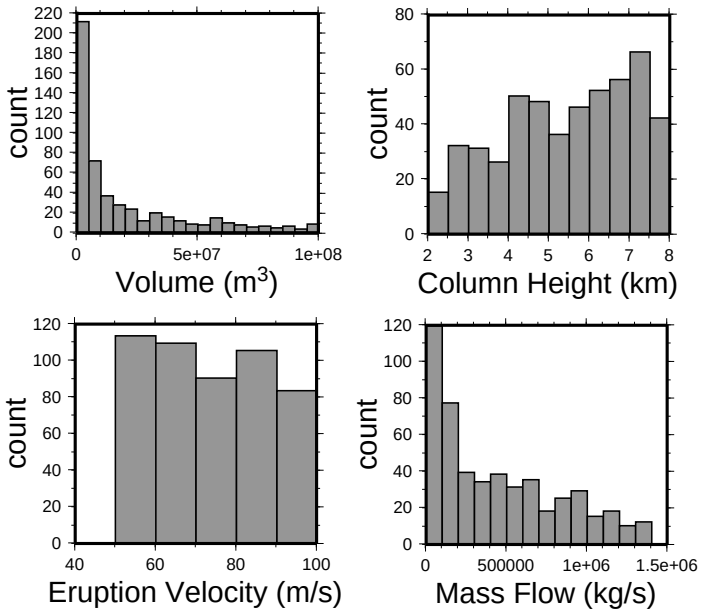

Figure 5: Histograms for several input parameters for the Suzuki model.

Stochastically sampled distributions for volume, column height, and eruption velocity are shown in Figure 5. Estimated mass flow given column height (equation 12) is also shown. For these parameter distributions, simulated tephra accumulation in León varies from $<< 0.1cm$ to $30cm$ in the 500 realizations.

The results of the analysis are best illustrated using an exceedence probability plot, also known as a hazard curve. For this range of input parameters, 50% of eruptions result in tephra accumulation $< 0.2cm$ in León (Figure 6). Approximately 26% of eruptions result in tephra accumulation $> 1cm$, and 11% of eruptions result in tephra accumulation $> 4cm$, in reasonable agreement with the historical record. The numerical simulation suggests that P[tephra $> 10cm$ in León $|$ volcanic eruption] = 5%. Probability of tephra accumulation $> 1cm$ and $> 4cm$ is contoured for the region about Cerro Negro volcano in Figures 7a and 7b. These maps illustrate the expected outcomes of eruptions at Cerro Negro for the population living closer to the volcano than León.

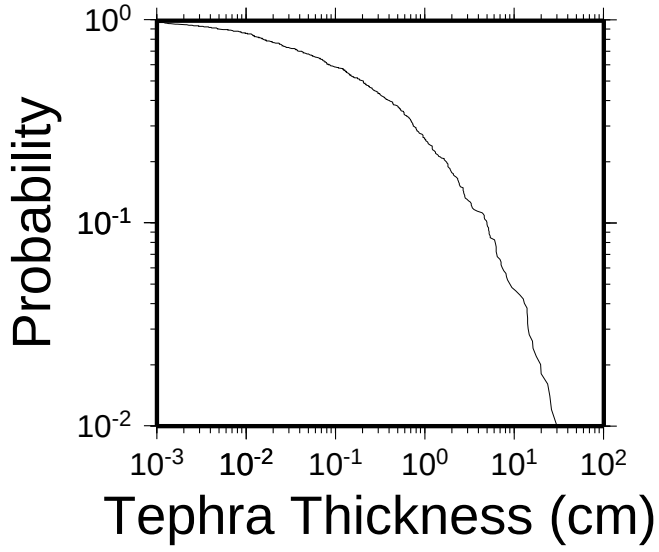

Figure 6: A conditional hazard curve for tephra accumulation in León.

3.3 Extreme Events at Cerro Negro

In addition to the probable outcomes summarized in Figures 6-7, it is useful to produce worst-case scenarios of volcanic activity in order to communicate the potential magnitude of hazards. Worst case scenarios bound the consequences of volcanic activity, with the goal of conveying the limits of potential volcanic devastation based on reasonable or conservative assumptions. Volcanologists often hesitate to convey worst case scenarios because of the fear, often well-founded, that worst-case scenarios will be misinterpreted as "base case" or "expected" scenarios. Conversely, worst-case scenarios can be ridiculed as overly conservative or alarmist. Nevertheless, it is worthwhile for volcanologists and public officials to think freely about large magnitude events and their impact on risks to public health and safety.

Development of worst-case scenarios can be made palatable through the introduction of the concept of upper limit values (ULVs). Essentially, ULVs are deterministic assessments of hazard using conservative assumptions. The concept of ULVs (aka screening distance values) was developed in seismic risk assessment for sensitive facilities, such as nuclear power plants,

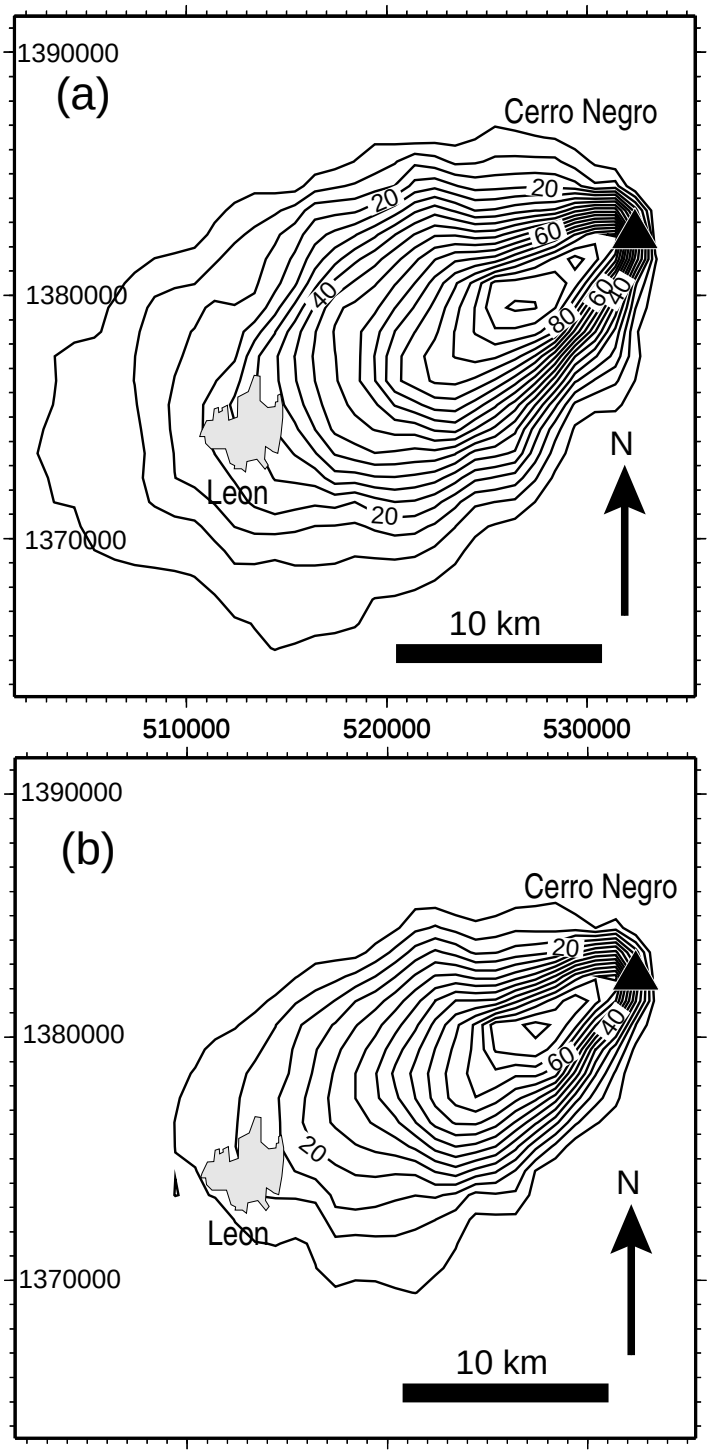

Figure 7: Probability maps for tephra accumulation from an eruption of Cerro Negro. P[tephra > 1 cm | volcanic eruption] (a) and P[tephra > 4 cm | volcanic eruption] (b) are contoured in percent, at 5% contour interval.

that must be located in areas of very low geologic risk. These techniques have now been extended to volcanic hazards for nuclear facilities (International Atomic Energy Agency, 1997) and are readily adapted to general hazard mitigation efforts.

Accurate development of ULVs relies on basic geologic investigations at many volcanoes, not necessarily those most likely to erupt. For example, the recent ($< 5,000$ yr b.p.) history of Crater Lake suggests that eruptions in this magmatic system are highly unlikely in the coming decades. Nevertheless, the history of Crater Lake and ancestral Mt. Mazama provides a benchmark for the potential magnitudes of future silicic volcanic eruptions in the Cascade mountains. In constructing hazard assessments for other Cascade volcanic systems, such as South Sister, geologic insights from studies of the Mt. Mazama eruption can be considered through the use of ULVs. Similarly, the Katmai eruption of 1912 and the Mt. Pinatubo eruption of 1991 should not be considered expected events in the Cascades during the next several decades, but such eruptions are completely within the realm of possibility and can be considered in hazards assessments through the use of ULVs.

As an example, we constructed an ULV for tephra accumulation in León, based on a set of parameters outside the range of past activity, but nonetheless possible, given eruptions from basaltic cinder cones in general. This model assumes an $8km$ high eruption column, $100m/s$ eruption velocity, and total volume of $1 \times 10^8 m^3$. For comparison, the $71day$ eruption of Tolbachik volcano, Kamchatka, sustained $3 - 10km$ high eruption columns and erupted $9 \times 10^8 m^3 (0.42km^3 DRE)$ of tephra (Doubik and Hill, 1999). It is also assumed that León lies on the major axis of dispersion and that the wind speed is $15m/s$ during the eruption. Such an eruption, larger than past eruptions at Cerro Negro, is at the approximate upper bound of VEI 3 activity. Given these parameters and based on the Suzuki model, the ULV for tephra accumulation in León is $47cm$.

4 Discussion and Conclusions

The hazard models presented here are strictly empirical, rather than predictive. The historical and geologic record of events is used to temper and refine the numerical analysis, and to temper the interpretation of the results. Although short, the record of eruptions at Cerro Negro is more complete than many volcanoes. For many volcanoes, analogous volcanic eruptions at other volcanoes may be needed to augment the record and to completely reflect the natural variation in expected eruptions.

The Suzuki model does not capture the physics of volcanic eruption completely. Rather, this model simplifies the physics in several ways. For example, the wind field is considered to be uniform with height above the volcanic vent, and particle motion in the column is treated probabilistically rather than with analytical determinism. Used in a stochastic fashion and calibrated by independent observations, the model works reasonably well despite such limitations. This dynamic exists in much of volcanic hazard assessment. Other examples of the effective use of simplified models in volcanology include Iverson, et al. (1998) and Wadge et al. (1994). When integrated with geologic data, as we have attempted here (Figures 2 and 3), such simplified models become efficient tools for volcanic hazards mitigation. This is different from attempting to forecast the outcome of a specific eruption, during which some variables, such as atmospheric conditions, are possibly well known. Although more dramatic to forecast the trajectory of an ash cloud as an eruption progresses, it is extremely useful to provide communities with a long-term forecast, prior to volcanic activity, so that informed decisions about building location, construction, and response can be formulated.

For Cerro Negro volcano, we conclude from analysis of the geologic data and numerical simulation that the probable tephra accumulation in León, given a volcanic eruption, of $> 1cm$ is approximately 29%, $> 4cm$ is approximately 9%, $> 10cm$ is approximately 5%. For smaller, more frequent eruptions our analysis relies most heavily on observation. For larger, less frequent

eruptions, our analysis relies more on the results of the numerical simulation. The ULV for tephra accumulation is León is approximately $0.5m$. In other words, given our current knowledge about this volcano, we do not envision any circumstances under which an eruption would result in more than approximately $0.5m$ of tephra deposition in León.

The main issues that emerge from application of this style of hazard assessment involve estimation of parameter distributions. For tephra accumulation, a clearer understanding of the links between magma properties and the parameters column height, eruption velocity, and total eruption volume has the potential of improving the hazard assessment. In time, the uniform random distributions used for parameters like eruption velocity might be replaced by more realistic distributions, or calculated directly from magma rheologic properties, as more about the physical basis of these links emerge. The importance of such links can only be appraised by progressing from the basic data and models through the hazard assessment. In this sense, results of probabilistic assessments can provide guidance about the volcanological research most likely to result in volcanic hazard reduction.

5 Acknowledgments

The authors thank Ron Martin and Laura Connor for their assistance in coding and document preparation. Field assistance at Cerro Negro volcano by Wilfried Strauch and Marta Navarro of the Instituto Nicaragüense de la Estudios Territoriales is gratefully acknowledged. Reviews by James Weldy and Budhi Sagar improved the manuscript. This manuscript documents work performed at the Center for Nuclear Waste Regulatory Analyses (CNWRA) for the Nuclear Regulatory Commission (NRC) under Contract No. NRC-02-97-09. The activities reported were performed for the NRC Office of Nuclear Safety and Safeguard, Division of Waste Management. This report is an independent product of the CNWRA and does not necessarily reflect the views or regulatory position of the NRC.

6 References

Armienti, P., G. Macedonio, and M. Pareschi, 1988, A numerical model for simulation of tephra transport and deposition: Applications to May 18, 1980, Mount St. Helens eruption, Journal of Geophysical Research 93: 6463-6476.

Baxter, P.J., 2000, Impacts of eruptions on human health, in: H. Sigurdsson, ed., Encyclopedia of Volcanoes, Academic Press, San Diego, 1035-1056.

Bonadonna, C., G. G.J. Ernst, and R.S.J. Sparks, 1998, Thickness variations and volume estimates of tephra fall deposits: The importance of particle Reynolds number, Journal of Volcanology and Geothermal Research 69: 217-239.

Carey, S.N., 1996, Modeling of tephra fallout from explosive eruptions, in: R. Scarpa and R. Tilling, eds., Monitoring and Mitigation of Volcanic Hazards, Springer-Verlag, Berlin, 429-461.

Connor, C.B., L. Powell, W. Strauch, M. Navarro, O. Urbina, and W.I. Rose, 1993, The 1992 eruption of Cerro Negro, Nicaragua: An example of Plinian-style activity at a small basaltic cinder cone, Eos, Transactions of the American Geophysical Union, 74: 640.

Doubik, O., and B.E. Hill, 1999, Magmatic and hydromagmatic conduit development during the 1975 Tolbachik eruption, Kamchatka, with implications for hazards assessment at Yucca Mountain, NV, Journal of Volcanology and Geothermal Research, 91: 43-64.

Fisher, R.V., and H.-U. Schmincke, 1984, Pyroclastic Rocks, Springer-Verlag, Berlin 471 pp.

Glaze, L. and S. Self, 1991, Ashfall dispersal of the 16 September 1986, eruption of Lascar, Chile, calculated using a turbulent diffusion model, Geophysical Research Letters 18, 1237- 1240.

GVN, 1999, Report on the October eruption of Guagua Pinchincha volcano, Ecuador, Global Volcanism Network Bulletin, 24:10.

Hill, B.E., C.B. Connor, P.C. La Femina, W. Strauch, G. Davoli, G. Guevara, and A. Saballos, 1999, August

1999 eruption of Cerro Negro volcano, Nicaragua, successfully forecast using time- volume relationships, Eos, Transactions of the American Geophysical Union, 80(46): F1111.

Hill, B.E., C.B. Connor, M.S. Jarzemba, P.C. La Femina, M. Navarro, and W. Strauch, 1998, 1995 eruptions of Cerro Negro volcano, Nicaragua and risk assessment for future eruptions, Geological Society of America, Bulletin 110: 1231-1241.

International Atomic Energy Agency, 1997, Volcanoes and Associated Topics in Relation to Nuclear Power Plant Siting, Provisional Safety Standard Series 1, 49 pp., Vienna, Austria.

Iverson, R.M., S.P. Schilling, and J.W. Vallance, 1998, Objective delineation of lahar-inundation hazard zones: Geological Society of America Bulletin, 110: 972-984.

Jarzemba, M.S., 1997, Stochastic radionuclide distributions after a basaltic eruption for performance assessments for Yucca Mountain, Nuclear Technology 118: 132-141.

Koyaguchi, T., 1996, Volume estimation of tephra-fall deposits from the June 15, 1991, eruption of Mount Pinatubo, by theoretical and geological methods, in: C.G. Newhall and R.S. Punongbayan, eds., Fire and Mud, Eruptions and Lahars of Mount Pinatubo, Philippines, University of Washington Press, Seattle, 583-600.

La Femina, P.C., C.B. Connor, B.E. Hill, S. Sandberg, N.T. Rogers, T. Dixon, W. Strauch, A. Saballos, 1999, Shallow dike emplacement during August, 1999, seismic and volcanic activity at Cerro Negro, Nicaragua, Eos, Transactions of the American Geophysical Union, 80(46): F972.

Malilay, J., M.G. Real, A.R. Vanegas, E. Noji, and T. Sinks, 1997, Vigilancia de la salud pública después de una erupción volcánica: lecciones aprendidas en Cerro Negro, Nicaragua, 1992, Revista Panamericana de Salud Pública 1(3): 213219.

McKnight, S.B., 1995, Geology and Petrology of Cerro Negro volcano, Nicaragua. Unpublished Masters Thesis, Arizona State University, Tempe, Az,132 pp.

McKnight, S.B., and S.N. Williams, 1997, Old cinder cone or young composite volcano? The nature of Cerro Negro, Nicaragua, Geology 25: 339-342.

Miller, T.P., and T.J. Casadevall, 2000, Volcanic ash hazards to aviation, in: H. Sigurdsson, ed., Encyclopedia of Volcanoes, Academic Press, San Diego, 915-930.

von Mises, R. 1957, Probability, Statistics, and Truth, Second Revised English Edition, Dover Publications, New York, 244 pp.

National Climatic Data Center, 2000, International Station Meteorological Climate Summary, STA 787410, Managua - Telecom Summary, http://www.ncdc.noaa.gov.

Navarro Ochoa, C., and J. F. Luhr, 2000, Late-Holocene tephrachronology at the Colima volcanic complex, Mexico, Colima Volcano Seventh International Meeting, Colima, Mexico, March, 2000, pp. 44-45.

Newhall, C.G., and S. Self, 1982, The volcanic explosivity index (VEI): an estimate of explosion magnitude for historical volcanism, Journal of Geophysical Research 87: 1231-1238.

Organización Panamericana de la Salud, 2000, Datos históricos de desastres naturales que han afectado Nicaragua en los ultimos años. http://www.ops.org.ni/desastre/d-civil/i-history.htm

Punongbayan, R.S., C.G. Newhall, L.P. Bautista, D. Garcia, D.H. Harlow, R.P. Hobblitt, J.P. Sabit, and R.U. Solidum, Eruption hazard assessments and warnings, in: C.G. Newhall and R.S. Punongbayan, eds., Fire and Mud, Eruptions and Lahars of Mount Pinatubo, Philippines, University of Washington Press, Seattle, 67-86.

Rosi, M., 1998, Plinian eruption columns: Particle transport and fallout, in: A. Freundt and M. Rosi (eds.), From Magma to Tephra, Modelling Physical Processes of Explosive Volcanic Eruptions, Elsevier, Amsterdam, 139-172.

Sarna-Wojcicki, A.M., S. Shipley, R.B. Waite, D. Dzurisin, and S.H. Wood, 1981, Areal distribution, thickness, mass, volume, and grain size of air-fall ash from the six major eruptions of 1980, in: P.W. Lipman and D.R.

Mullineaux, eds., The 1980 Eruptions of Mount St. Helens, Washington, U.S. Geological Survey Professional Paper 1250: 577-600.

Shafer, G., 1996, The Art of Causal Conjecture, The MIT Press, Cambridge, Massachusetts, 511 pp.

Simkin, T., and L. Siebert, 1994, Volcanoes of the World, Second Edition, Smithsonian Institution, washington D.C., 349 pp.

Sparks, R.S.J., M.I. Bursik, C.N. Carey, J.S. Gilbert, L.S. Glaze, H. Sigurdsson, and A.W. Woods, 1997, Volcanic Plumes, John Wiley and Sons, West Sussex, 574 pp.

Suzuki, T., 1983, A theoretical model for the dispersion of tephra, in: D. Shimozuru and I. Yokoyama, eds., Arc Volcanism: Physics and Tectonics, Terra Scientific Publishing, Tokyo, 95- 113.

Wadge, G., P.A.V. Young, and I.J. McKendrick, 1994, Mapping lava flow hazards using computer simulation, Journal of Geophysical Research 99: 489-504.

Walker, G.P.L., S. Self, and L. Wilson, 1984, Tarawera 1886, New Zealand - A basaltic plinian fissure eruption. Journal of Volcanology and Geothermal Research 21: 61-78.

Wilson, L., and J. Head, 1981, The ascent and eruption of magma on the Earth and Moon, Journal of Geophysical Research, 86: 2971-3001.

Woods, A.W., 1988, The fluid dynamics and thermodynamics of eruption columns, Bulletin of Volcanology, 50: 169-193.

Woods, A.W., 1995, The dynamics of explosive volcanic eruptions, Reviews in Geophysics, 33: 495-530.

Woods, A.W., and M.I. Bursik, 1991, Particle fallout, thermal disequilibrium, and volcanic plumes, Bulletin of Volcanology, 53: 559-570.

Probabilities of Volcanic Eruptions and Application to the Recent History of Medicine Lake Volcano

Manuel Nathenson
U.S. Geological Survey
Menlo Park, California 94025

An underlying assumption of USGS hazards assessments for the Cascades is that the probability of volcanic eruptions may be treated as a Poisson process. Time histories for some volcanoes match this assumption well (Klein, 1982). In order to calculate an annual or 30-year probability for an eruption, the relation for the occurrence rate is used. For a Poisson process, this relation is obtained from the exponential distribution for the probability that an eruption will occur in a time T less than or equal to the interval time t:

$$P\{ T \le t\} = 1 - e^{-\mu t} \approx \mu t , \quad \text{for } \mu t \text{ small,}$$

where μ is the mean occurrence rate of events per year. Since occurrence rates are small in the Cascades, the approximate relation shown is normally used (e.g. Scott and others, 1995). For the case of lava flows from isolated vents covering an area a in a volcanic field of total area A, a factor $p = a/A$ can be factored in as μtp to account for the probability of areal coverage (e.g. Crowe and others, 1982). This analysis assumes that the occurrence of vents are homogeneous in space within the defined area of the volcanic field, but areally varying probability models have been developed (Connor and Hill, 1995).

The properties of a Poisson process include the characteristic that the conditional probability of waiting a time until an eruption occurs does not depend on the time that we have already waited, but only on the time that is in the future. For some volcanoes, the time history contains disparate time intervals between eruptions, with some being short and others being much longer. Some examples of time histories with disparate eruption time intervals are: Mullineaux's (1974) data for eruption times of tephra layers at Mount Rainier have three long intervals (>2000 years) and seven short intervals (<600 years). Mullineaux's (1996) data for Mount St. Helens has one interval of 8600 years, one of 1500 years, and 34 less than 640 years. In these cases, other probability distributions are a more accurate representation of the data, and the conditional probability for these distributions will depend on the time since the last eruption. Use of the Weibull distribution introduced by Bebbington and Lai (1996) has shown mixed results in dealing with these disparate intervals. An alternate distribution is the mixed exponential (Cox and Lewis, 1966, p. 140):

$$P\{ T \le t\} = 1 - p_1 e^{-\mu_1 t} - p_2 e^{-\mu_2 t}$$

where p_1 is the fraction of short intervals, μ_1 is the average occurrence rate for the short intervals, and p_2 and μ_2 are the same parameters for the long intervals. The basic notion embodied in this relation is that there are two states, one involving short intervals and a second involving long intervals. The probability of an eruption occurring in either of these states is each governed by an exponential distribution. The mixed-exponential distribution appears to match the available data reasonably well and resolves the conceptual problem of eruption probability for volcanoes with disparate eruption time intervals.

Donnelly-Nolan and others (1990) provide a chronology for Medicine Lake volcano for 17 numbered units (with 17 being the most recent) erupted in the last 11,000 carbon-14 years. Their units 1 and 4 are contemporaneous and are treated as single eruption. Using the data set in Stuiver and others (1998), the carbon-14 ages can be converted to calendar years and single dates chosen with some interpretation. However, unit 12 (Medicine dacite flow) is only dated by using paleomagnetic data to correlate its direction of magnetization with directions from dated lava flows in other locations. The direction correlates with lava flows having calendar-year ages of 2000, 3590, and 4880 years (calendar-year ages are given as before A.D. 2000 in this report). We analyze eruption time intervals for three cases depending on the various age possibilities of unit 12. Using the 2000-year calendar age of unit 12, Medicine Lake has one repose-time interval of 7700 years, one of 1640 years, and 13 that are less than 780 years.

The probability distribution for eruption time intervals is shown in Figure 1 assuming a 2000-year age for unit 12. The exponential distribution is a poor match to the data, whereas the Weibull and mixed exponential are better matches. The Weibull, in this case, fits the data better than the mixed exponential. For a calendar-year age of unit 12 of 3590 years, Figure 2 shows the resulting probability distributions. In this case, the Weibull and mixed exponential are equally good matches to the data. In calculating the conditional probability that an eruption will occur in the next 30 years given a time since the last eruption, the exponential distribution has no dependence on the time since the last eruption (Figure 3), whereas the Weibull and mixed exponential distributions start with higher probability and decrease as the time since the last eruption increases (Figure 3). The principle being demonstrated is that after a long time since the last eruption, the interval to the next eruption can only be long, and the probability of an eruption in a certain future time decreases. The Weibull distribution has the distinctive property of having a spike in probability values at early time since the last eruption (Figure 3). This very high probability at short time seems to be a characteristic of the Weibull distribution rather than required by the data and contrasts with the smoother behavior of the mixed exponential curve. Surprisingly, the Weibull and mixed exponential have very similar probabilities of an eruption in the next 30 years (790 years after the last eruption) of around 2.5 %, and both are significantly less than the estimate from the exponential distribution of 3.7 %.

In addition to the probability of an eruption occurring in a given time period, an important consideration is the likely size of an eruption. Figure 4 shows cumulative events per thousand years versus volume erupted. The data follow a logarithmic relationship with volume. Although the eruption probability characteristics developed in the previous analysis are not involved in the analysis, the cumulative time/volume relationship allows estimates that once per every thousand years, there should be an eruption of about 0.02 km^3, whereas an eruption of as much as 0.9 km^3 would occur only once every 10,000 years at Medicine Lake volcano.

References

Bebbington, M. S., and Lai, C. D., 1996, On nonhomogeneous models for volcanic eruptions: Mathematical Geology, v. 28, p. 585-600.

Connor, C. B., and Hill, B. E., 1995, Three nonhomogeneous Poisson models for the probability of basaltic volcanism: Application to the Yucca Mountain region, Nevada: Journal of Geophysical Research, v. 100, p. 10,107-10,125.

Cox, D. R., and Lewis, P. A. W., 1966, The Statistical Analysis of Series of Events: Matheuen, London, 285 p.

Crowe, B. M., Johnson, M. E., and Beckman, R. J., 1982, Calculation of probability of volcanic disruption of a high-level radioactive waste repository within southern Nevada, USA: Radioactive Waste Management and the Nuclear Fuel Cycle, v. 3, p. 167-190.

Donnelly-Nolan, J. M., Champion, D. E., Miller, C. D., Grove, T. L., and Trimble, D. A., 1990, Post-11,000-year volcanism at Medicine Lake Volcano, Cascade Range, Northern California: Journal of Geophysical Research, v. 95, p. 19,693-19,704.

Klein, F. W., 1982, Patterns of historical eruptions at Hawaiian volcanoes: Journal of Volcanology and Geothermal Research, v. 12, p. 1-35.

Mullineaux, D. R., 1974, Pumice and other pyroclastic deposits in Mount Rainier National Park, Washington: U. S. Geological Survey Bulletin 1326, 83 p.

Mullineaux, D. R., 1996, Pre-1980 tephra-fall deposits erupted from Mount St. Helens, Washington: U.S. Geological Survey Professional Paper 1563, 99 p.

Scott, W. E., Iverson, R. M., Vallance, J. W., and Hildreth, Wes, 1995, Volcano hazards in the Mount Adams region, Washington: U.S. Geological Survey Open-File Report 95-492, 11 p., 2 plates, scale 1:500,000, 1:100,000.

Stuiver, Minze, Reimer, P. J., Bard, Edouard, Beck, J. W., Burr, G. S., Hughen, K. A., Kromer, Bernd, McCormac, Gerry, van der Plicht, Johannes, and Spurk, Marco, 1998, INTCAL98 radiocarbon age calibration, 24,000–0 cal BP: Radiocarbon, v. 40, p. 1041-1083.

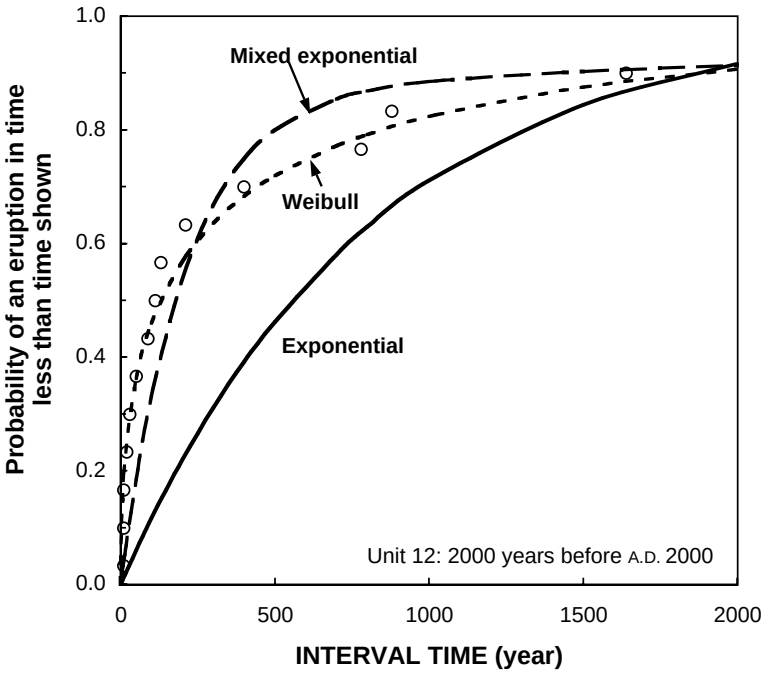

Figure 1. Probability of an eruption in a time less than some interval time between eruptions. Circles are data for Medicine Lake Volcano assuming unit 12 erupted 2000 years before A.D. 2000. Curves are for model distributions.

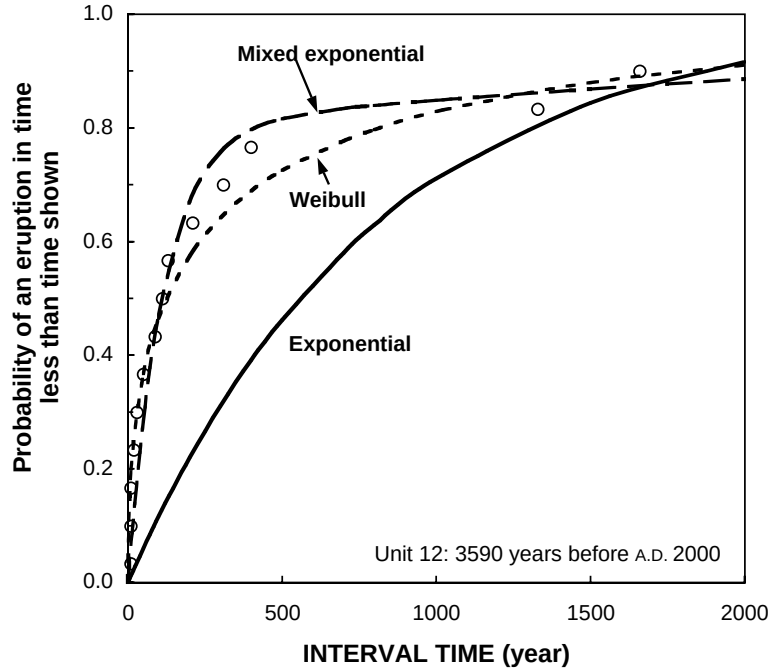

Figure 2. Probability of an eruption in a time less than some interval time between eruptions. Circles are data for Medicine Lake Volcano assuming unit 12 erupted 3590 years before A.D. 2000. Curves are for model distributions.

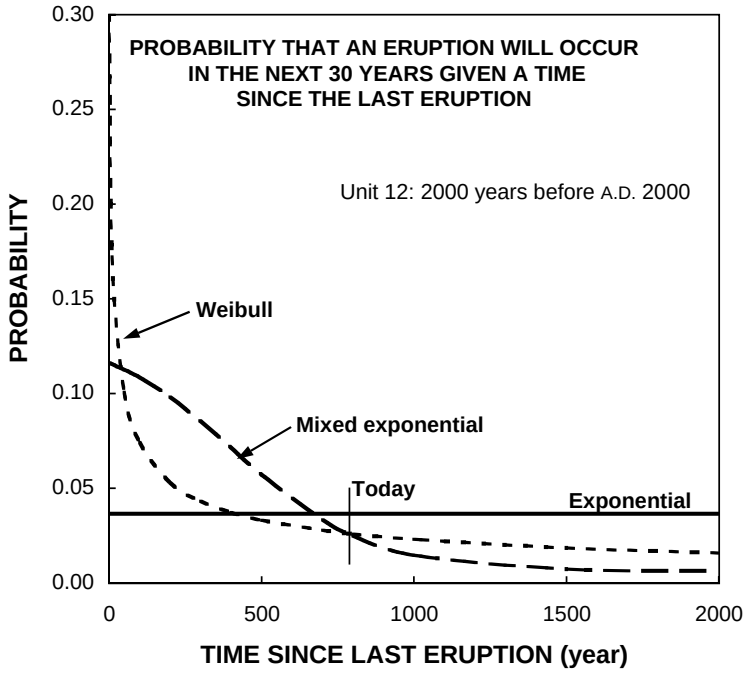

Figure 3. Conditional probability that an eruption will occur in the next 30 years after waiting some time since the last eruption. Calculated assuming unit 12 erupted 2000 years before A.D. 2000. Today notes time since last eruption at Medicine Lake Volcano.

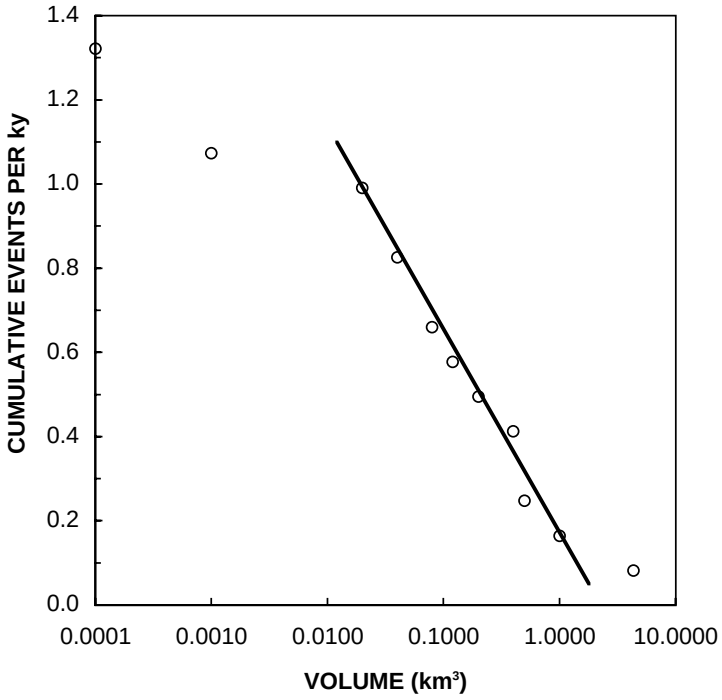

Figure 4. Cumulative number of events per thousand years (ky) versus volume erupted for the eruptions in the last 13,000 years at Medicine Lake Volcano.

HAZARDS MAPPING: A NEED FOR GUIDELINES

Dr. Ute J. Dymon
Department of Geography
Kent State University
Kent, OH 44242-0001
Email: Udymon@kent.edu

Mapping Applications Center
Reston, VA 20192
Email: Udymon@usgs.gov

Abstract

As development in the world keeps expanding over the countryside, the costs of natural and technological disasters keep escalating. For less costly and more effective response to these extreme events, our use of information before, during and after disasters must be enhanced. Hazard maps provide a unique organization of vital information for hazard identification, risk estimation and allocation of resources. They are a major support for emergency managers at all stages of a disaster. However, currently, there are no unified guidelines for making these hazard maps. Each person charged with producing a new hazard map must develop content and design methods. Content and design not only differs between countries, but often within a government entity there is little map coordination. The lack of agreed upon symbolization, meanings of color ranges and other design characteristics does not provide a common ground for understanding maps in planning and in emergencies. Standardizing content and approach to design of hazard maps produced throughout government agencies will do much to enable managers to absorb vital information as efficiently as possible. The development of guidelines for hazard mapping will illuminate communication between planners and will strengthen the ability of emergency managers to better understand information at a glance during crucial decision making moments. This paper presents an outline for the creation of a common language on hazard maps.

Why Do We Need Hazards Maps?

As humans occupy more space on earth, the burden of mortality and property losses from natural and technological hazards have also increased. Expanding population growth contributes to the daily risk of natural and technological emergencies impacting humans in their everyday lives. Over the past decade, technological accidents and natural disasters in which many people lost their lives were reported. For less costly and more effective response to these extreme events, our use of information before, during and after disasters must be enhanced. Efficient coordination of the efforts of all emergency respondents is vital. Hazard maps provide a unique organization of vital information for hazard identification, risk estimation and allocation of resources. They are a major support for emergency managers at all stages of a disaster. Well-integrated hazards management maps can be a major key device in promoting interdepartmental, inter-agency and cross-jurisdictional coordination of emergency response efforts before, during and after a disaster (Dymon and Winter, 1993). In spite of the critical role maps play in the planning for disasters, hazards management maps are a neglected area of study. In the current literature there is little information about hazards maps as a whole (Dymon and Winter, 1993).

It is commonly assumed that a map is a concrete object. However, people construct mental maps out of an aggregation of the input from their senses over time as they experience and interact with the environment or space (Liben et al., 1981). Dent (1998) defined mental maps as "mental images that have spatial attributes." According to Gould and White (1986) and Downs and Stea (1973), these images vary markedly between individuals and are influenced by each person's social class and location. MacEachren (1989) described the gradual developmental process whereby adults gather information from the environment. Proceeding from an initial spatial framework, details about an environment are subsequently incorporated into an individual's mental framework. MacEachren suggests that the location of landmarks and routes is a major feature of the cognitive learning process. It follows that newcomers will have less recall about an area compared to long-term residents. The diversity of mental images employed by decision makers during any disaster may be a critical ingredient in the way the disaster is managed (Dymon and Winter, 1991).

By providing concrete objects such as hazards management maps, information on these maps can enhance the mental image of the critical situation in the emergency manager's spatial framework. This information may help the manager become more in tune with others who also manage the event from a different perspective or from a particular needs situation. An extended taxonomy of Hazard Management Maps which need to be considered for the formulation of unified guidelines has been developed by Winter (1997) and Dymon (2000).

Why Do We Need Guidelines?

Hazard management maps are being produced and published by various federal, state and local agencies, institutions and the private sector. The Federal Emergency Management Agency (FEMA), responsible for response to disasters, has developed comprehensive event-specific Information Collection Plans which were developed by the Rapid Assessment Work Group (FEMA, 1999). These plans provide detailed information on an emergency event itself, for example, on hurricanes. Listings of contacts and specific information required to deal with the event are described in each plan. Similarly, scientists at other federal agencies such as the United States Geological Survey (USGS) provide guidance to others for data collection methods and analysis, as in the case for seismic hazards whereby Hanks and Cornell (1994). Randall et al (1998) outlined methods for Landslides Hazards Maps. Each of these sets of guidelines addresses one type of hazard. Rather than addressing a single hazard, a different perspective can be gained by addressing the entire field of hazard management and disaster management mapping.

Hazard management mapping encompasses the entire domain of mapping found at the federal government and often in the private sector. A wide variety of maps are employed to identify the hazard itself, the degree of hazardousness, and results from screening and monitoring thereby providing the opportunity for a diagnostic overview. Other maps show the probalistic risk in a given situation. Based on calculations, they estimate the probability of risk from a given hazard in a given area and attempt to explain the degree of consequences a given risk produces. Vulnerability maps have become popular in recent years. They often center on issues of health, economics or environment and integrate data from many different sources to portray how vulnerable a given area is to one or several hazards. Most of these maps are researched and prepared by federal agencies and have been formed with some standards and guidance in place. In contrast, disaster management maps are prepared in most cases by local, state and regional agencies. These maps have little in common, and the emergency manager or planner preparing them often is not familiar with the wider universe of emergency mapping. Guidelines are needed more at this level than at any other.

Problems Associated With Not Having Guidelines

Currently, there are no unified guidelines throughout government agencies for making these hazard maps. Content and design not only differ between federal, state and local agencies, but often within a government entity there is little map coordination. The lack of agreed upon design methods such as symbolization, meanings of color ranges and other design characteristics does not provide a common ground for understanding maps in planning and in emergencies.

The Advantages of Having Guidelines

In 1996, a Natural Disaster Reduction Plan was prepared by the National Science and Technology Council Committee on the Environment and the Natural Resources Subcommittee on Natural Disaster Reduction. This strategic plan clearly states that risk assessment is the necessary starting point with which to identify the resilience needed in society and the sort of overarching mitigation strategies which must be implemented to strengthen societal resilience to hazards. Such an effort is essentially spatial and requires maps as research and planning tools.

Standardizing the content and the approach to design of hazard maps produced throughout government agencies will do much to enable managers to absorb vital information as efficiently as possible. The development of guidelines for hazard mapping will illuminate communication between planners and will strengthen the ability of emergency managers to better understand information at a glance during crucial decision making moments. More generally, Federal mapping guidelines will aid Local Emergency Planning Committees in their production of local hazard maps and will provide clearer communication in the use of hazard maps in planning efforts in the private sector.

Americans today face a plethora of natural hazards from tsunamis to volcanoes and the dangers of technological surprises from explosions to groundwater contamination. The creation of a common language on hazard maps would be a major step in arming planners and emergency managers to handle these disasters. The Federal Government has an opportunity to play a major role in the establishment of hazard mapping guidelines.

References

Dent, B.D., 1990. Cartography: Thematic Map Design. William C. Brown, Dubuque, IA.

Downs R. and D. Stea,1973. Images and Environment. Aldine, Chicago, Il.

Dymon, U. J. , 2000. *A Common Ground for Understanding Hazards Maps*. NMD Research Symposium 2000, Rolla, MO. Pp 90-93.

Dymon, U. J. and N.L. Winter, 1993. Evacuation Mapping: The Utility of Guidelines. Disasters, 17 (1),12-24.

Dymon, U. J. and N.L. Winter, 1991. Emergency Mapping in Grass Roots America: A Derailment Evacuation Case Study. G eoforum, 22 (4) 377- 389.

Federal Emergency Management Agency, 1999. Information Collection Plan. Rapid Assessment Working Group, Washington, D.C.

_____.1999. Geographic Information System: Its Mapping and Analysis Center, Production Catalogue. Washington D.C.

Gould, P. R. and R. L. White, 1986. Mental Maps. Allen and Unwin, Winchester, MA.

Hanks, T. C. and C. A. Cornell, 1994. Probabilistic Seismic Hazards Analysis: A Beginner's Guide. Proceedings of the fifth Symposium on Current Issues Related to Nuclear Power Plant Structures, Equipment and Piping, I/1-1 to I/1-17. North Carolina State University, Raleigh, N.C.

MacEachren, A. M., 1995. How Maps Work: Representation, Visualization, and Design. New York: Guilford Press.

Office of Science and Technology, 1996. The Natural Disaster Reduction Plan. National Science and Technology Council Committee on the Environment and Natural Resources Subcommittee on Natural Disaster Reduction. Washington, D.C.

Jibson, R. W., E. L. Harp and J. A. Michael. 1998. A Method for Producing Digital Probalistic Seismic Landslide Hazard Maps: An Example from the Los Angeles, California, Area. Open-File Report USGS 98-113.

Winter, N. L. 1997. Managing a Mega-Disaster: GIS Applications, Decision Making and Spatial Data Flow Between Local, State and Federal Levels in Hurricane Andrew Disaster Management. Clark University, Goddard Library (unpublished dissertation).