Statistical Science
2003, Vol. 18, No. 4, 412–421

Three Environmental Probabilistic Risk Problems

David R. Brillinger

Abstract. Risk analysis may be defined as the problem of estimating the probabilities of rare events and the magnitudes of associated damages. The topic unifies the environmental sciences. This paper considers risk analyses for earthquakes, wildfires and floods. The computation of insurance premiums is used to motivate and unite the work.

Key words and phrases: Catastrophe modelling, disasters, earthquakes, floods, insurance, natural disasters, natural hazards, risk assessment, risk theory, wildfires.

1. INTRODUCTION

1.1 The Problem

Tremendous financial losses occur from environmental disasters such as floods, tropical storms, droughts, tornadoes, forest fires and earthquakes. To add to societal concern, the number and costs of these seem to be increasing; see Figure 1, adapted from [29]. One can speculate on the cause. Global warming and population movement to hazardous areas have both been suggested.

Risk management and insurance purchase seek to ameliorate the impacts of disasters. To determine insurance premiums actuaries need estimates both of probabilities of occurrence and of distributions of damages. Further, government regulators are concerned with solvency of insurance companies, and civil engineers need estimates of probabilities for design purposes and building codes. The field of statistics becomes involved for a variety of reasons, including large data sets, small data sets and uncertain inferences. Novel analytic problems arise.

1.2 Risk

Risk may be defined as the probability of some hazardous event or catastrophe, the chance that something bad will occur. The occurrence becomes a catastrophe when the losses sustained are severe. In many cases huge amounts of money are involved [30]. A principal concern therefore is low probability–high consequence events, occurrences that lead to damage, loss, injury, death and/or substantial environmental impairment. Often the work is done as an aid to decision making. In consequence risk models and risk management pervade modern technical life.

A common tool of workers in the field is a *catastrophe model*. These models have been defined as sets of databases and computer programs designed to analyze the impact of different scenarios on hazard-prone areas [30]. In practice these models combine scientific risk assessments of hazard with historical records in order to estimate the probabilities of disasters of different magnitudes and the resulting damages. The information may be presented in the form of expected annual losses and/or the probability that in a given year the claims will exceed a certain amount.

Risk analyses may be required by government agencies. To mention a specific example, in nuclear accident prevention a "core damage frequency" rate maximum of 10^{-4} events per reactor year is required by the U.S. Nuclear Regulatory Commission [33].

A formal risk analysis may include (i) estimation of probabilities, (ii) determination of the distribution of damage and (iii) preparation of products such as formulas, graphics and hazard risk maps. Typically there is extensive use of computer science, systems analysis, substantive subject matter and statistical methods. Important analytic tools include box-and-arrow diagrams, software packages, simulation, decision tools, GIS, visualization and database management.

David R. Brillinger is Professor, Department of Statistics, University of California–Berkeley, 367 Evans Hall, Berkeley, California 94720-3860 (e-mail: brill@stat.berkeley.edu).

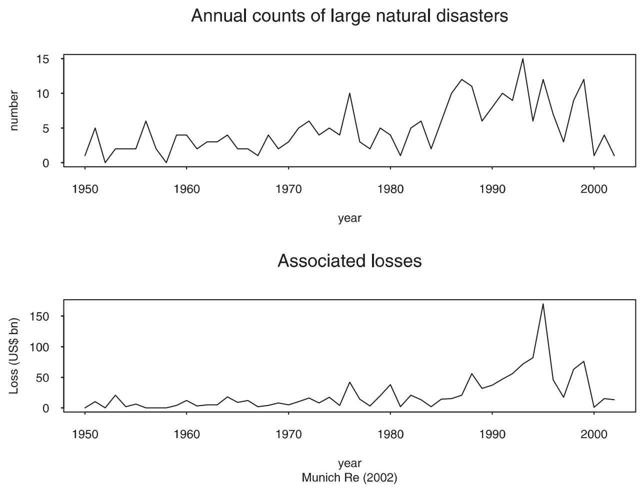

FIG. 1. *Counts and losses of worldwide great catastrophes. The figures are adapted from* [29].

A pair of nice examples of statistical work in risk analysis is provided by Fairley [17] and Miller and Leslie [27]. The Fairley work is concerned with the probability of a spill of liquified natural gas during its importation at U.S. ports. In particular the paper makes the case for describing risks by probabilities, not by their reciprocals, the return periods. The Miller–Leslie paper concerns the probability of a ship hitting the Tasman Bridge at Hobart, Australia. In both these papers there is careful evaluation of the probabilities of component events.

1.3 Insurance and Some Historical Background

Insurance measures are often taken to manage catastrophic phenomena, for the events of concern typically have substantial financial implications. Insurance against loss has been around for many years. One can mention the code of Hammurabi around (1780 BC), which involved so-called bottomry, a form of marine insurance [30]. The catastrophes such as the Great Fire of London 1666 and the U.S. floods of the late 1800s led to the development of insurance as a major business [11, 37].

An insurance premium paid is meant to reflect the risk potential [15]. A tentative premium may be computed making use of probability and damage estimates for possible events. In practice it will be "loaded" to cover the costs of doing business and adding a profit. The pure risk or net premium for damage L in a single instance is given by

$$P = E\{L\}$$

and is often usefully written as

$$\text{Prob}\{L \neq 0\} E\{L | L \neq 0\}.$$

This second form separates the probability and the consequence.

Insurers are not generally concerned with single events or short time periods. They maintain stability by pooling many risks over time and type to deal with random fluctuations. Formulae for loaded premiums include

$$P = (1+\alpha)\mu_L, \quad \mu_L + \beta\sigma_L, \quad \mu_L + \gamma\sigma_L^2, \quad \alpha\mu_L + \beta\sigma_L + \gamma\sigma_L^2,$$

where

$$\mu_L = E\{L\}, \quad \sigma_L^2 = \text{var}\{L\}$$

and α, β, γ are parameters to be given numerical values [3].

The premiums may also be based on considerations of the following type. Suppose that the reserve available to the insurer in the time period t is R_t. Then the insurer becomes insolvent in period t if $L_t > R_t + P_t$ with L_t denoting the total of claims and P_t the total of the premiums in period t. It now makes sense to consider the probability of ruin,

$$\text{Prob}\{L_t > R_t + P_t \text{ for some } t \text{ in a specified time period}\}. \tag{1}$$

This is a crossing probability for $L_t - R_t - P_t$. Various models and approximations have been proposed for (1). For example, there is the Cramér–Lundberg model [14, 18], based on a Poisson process for the times of the claims included in L_t.

In determining premiums insurers have to deal with the following difficulties: (i) ambiguity of risk; that is, the event or its probability is not well defined; (ii) adverse selection; that is, those most at risk are more likely to purchase insurance; (iii) moral hazard; that is, an increase in probability of loss caused by some type of behavior on the part of the policy holder; and (iv) correlated risk; that is, the simultaneous occurrence of losses from a single event [23].

There are various practical details to be dealt with, including taxes, reinsurance, exposure, inflation, investment return, lags, interest rates and the tax authority's treatment of surpluses. There are other approaches. For example, an extreme value approach is taken in [16] and there is a market-driven approach [4]. The latter is adaptive and evolutionary and uses time series data on income and expenses to compute a premium from predicted future expenses. It is interesting to read in [11] of the empirical efforts of companies in the 18th century to find effective premium rates on the basis of experienced gains and loses.

In summary, probabilities of occurrence need to be estimated as well as distributions of losses. One finds oneself dealing with probabilities and distributions associated with phenomena in space and time. Modern actuarial science deals with some other aspects as well, for example, the relationship between a particular risk and claims in the whole market, the attitude toward risk in the market and the total assets of all insurers [1, 2].

In the sections that follow we restrict ourselves to the details of examples from the fields of seismology, forest science and hydrology.

2. EXAMPLES OF RISK ANALYSIS

EXAMPLE 1 (Earthquake damage). (a) *Background*. Cornell [13] is the seminal paper on seismic risk assessment. His definition of the subject is a variant of the following:

- *Seismic risk assessment*—the process of estimating the probability that certain performance variates at a site of interest exceed relevant critical levels within a specified time period as a result of nearby seismic events.

The approach presented here is one of breaking the problem down conceptually into manageable parts, including (i) damage, (ii) site, (iii) attenuation and (iv) event locations, times and sizes. These parts are illustrated in Figure 2. The figure contains both series and parallel structures and shows two possible events. The "?" between "Site" and "Structure" in the figure refers to the possibility of feedback between the two being present.

The presentation flows backward from a structure at the site of interest to the locations, times and sizes of earthquakes. This has the advantage of better anticipating the requirements at each stage of the analysis.

(b) *Damage*. There are a variety of ways to describe and estimate earthquake damage. An old and elementary one, yet an important one, uses the modified Mercalli intensity (MMI). One reason for this scale's importance is that sometimes values may be derived from historic accounts. A second is that it refers to damage directly.

MMI values are given by roman numerals I to XII (and sometimes 0 referring to nothing felt or noticed). The scale is ordinal increasing with growing severity of damage. For example, the definition of MMI VIII includes "Damage slight in specially designed structures; considerable in ordinary substantial buildings; ..." while that of MMI IV includes "Dishes, windows, doors disturbed; walls make creaking sound; ..." [10].

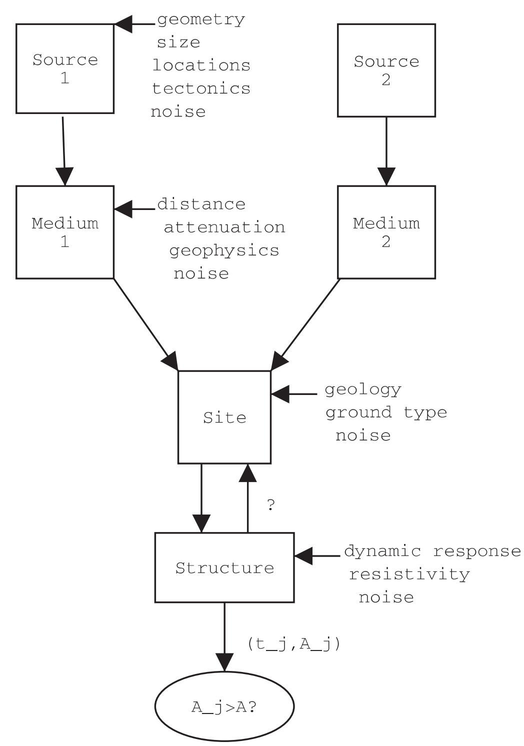

FIG. 2. *Box-and-arrow diagram highlighting components of seismic risk analysis. The symbol A_j refers to the level of some performance variate associated with the jth event*; *A is a threshold level.*

There are functions that have been proposed to convert MMI values into damage percentages for different types of structures. Table 1 is an example, a so-called damageability matrix, taken from [28]. The entries are loss ratios per risk category in present.

Figure 3 shows some of the MMI values observed following the Northridge event of 17 January 1994. The event occured 30 km NW of Los Angeles, California. Its size, as measured by magnitude, was 6.7. There were 57 deaths, 1,500 serious injuries and 12,500 structures moderately to severely damaged. The damage was estimated as US $12.5 billion. There are 554 observations to be used in the analyses. The highest

TABLE 1
Loss ratios per risk category in percentage damaged

MMI	**VI**	**VII**	**VIII**	**IX**	**X**
Residential	0.4	1.7	6	17	42
Commercial	0.8	3	11	27	60
Industrial	0.1	0.7	3	11	30

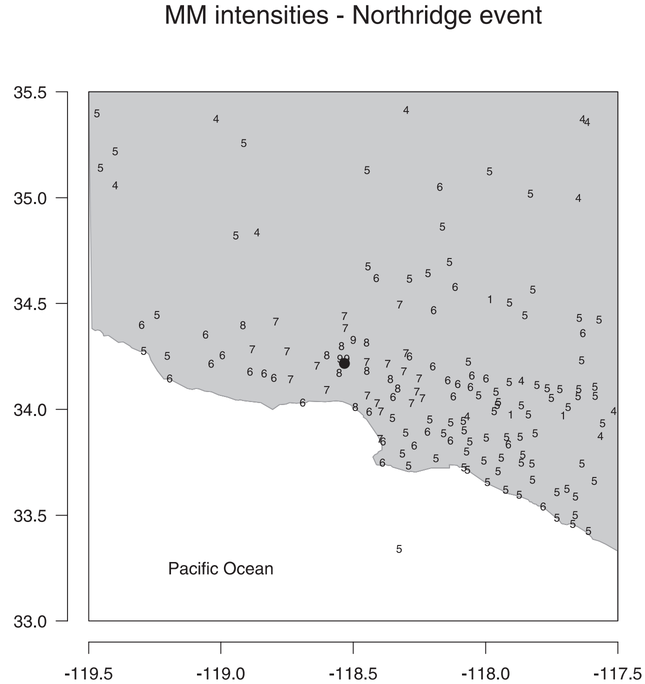

FIG. 3. *A sampling of the MMI values observed for the Northridge earthquake of 1994. The black circle represents the epicenter of the event.*

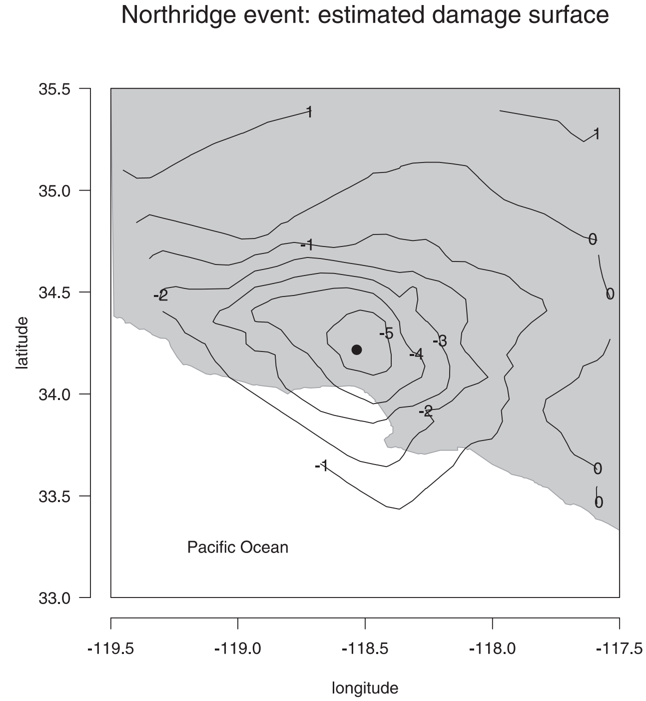

FIG. 4. *The estimate of the function $f_{x,y}$ of* (2). *The smaller the contour level*, *the greater the damage.*

intensity recorded was IX. The 5 = MMI V of the figure, apparently in the ocean, is actually an observation made on Catalina Island. More details of the event may be found in the February 1996 issue of the *Bulletin of the Seismological Society of America*.

As indicated, MMI values are ordinal. Such data are conveniently handled by postulating the existence of a latent variable ζ and cut points a_i such that the MMI value at location (x, y) is given by

$$I_{x,y} = i \quad \text{if } a_i < \zeta_{x,y} \le a_{i+1}.$$

Suppose further that

$$\zeta_{x,y} = f_{x,y} + \varepsilon_{x,y} \tag{2}$$

with $f_{x,y}$ deterministic and smooth and further suppose that the $\varepsilon_{x,y}$ have independent extreme value distributions. The use of the extreme value distribution is plausible given that the concern is damage. It and the corresponding complimentary log–log link mean that the R/S-PLUS function glm$(\cdot)$ may be used for the computations (see [26], Chapter 5). In the related paper [7], $f_{x,y}$ is estimated using the R/S-PLUS smoother loess$(\cdot)$ of Cleveland [12] and the R/S-PLUS function gam$(\cdot)$ of Hastie and Tibshirani [19], with data from the 1989 Loma Prieta event employed. Other smoothers are investigated in [8].

Figure 4 provides the estimate of f obtained for the Northridge event. One can see a general dying off of the estimated values as one moves away from the epicenter of the event. Referring to the contour levels indicated, because of the way things have been set up, the more negative f the greater the damage.

(c) *Site*. In practice further details concerning the observations may be available such as the geology of the site. Such information, and other covariates, could be included directly with $f_{x,y}$ in the linear predictor.

(d) *Attenuation*. Next a relationship describing the falloff in energy with distance from the earthquake origin is needed. This falloff is apparent in Figure 4. Following Joyner and Boore [21, 7], one can use the attenuation form

$$\begin{aligned} &\log\bigl(-\log(1 - \text{Prob}\{I = i\})\bigr) \\ &\quad = \alpha_i + \beta d + \gamma \log(d) + \delta M, \end{aligned} \tag{3}$$

$i = 0, \text{I}, \text{II}, \ldots, \text{XII}$, where d is the distance of the observation point to the hypercenter of the event and M is the event's magnitude. This relationship was fitted to the Loma Prieta data in [5]. For the Northridge data results are presented in Figure 5. It gives the fitted value of Prob$\{I = i\}$, $i = 0, \text{V}, \text{IX}$, as functions of d. One sees a very rapid falloff with distance in the case of MMI IX. The curve for MMI V peaks at a distance

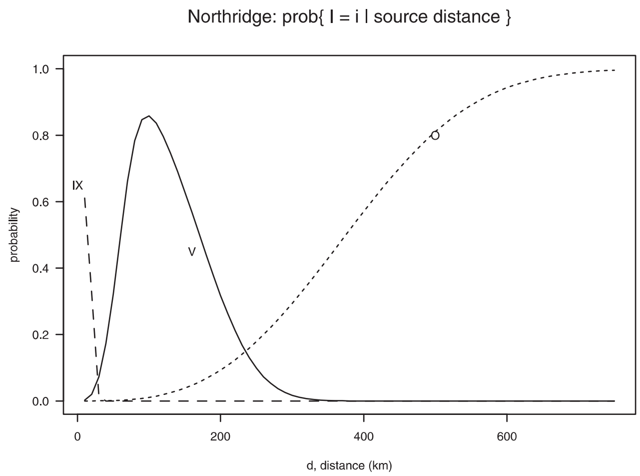

FIG. 5. *Fitted probabilities of the indicated MMIs for a site located at a distance d from the epicenter.*

of about 125 km, while the curve for MMI 0 rises steadily to 1 as d increases.

(e) *Event locations and times*. To complete the analysis one needs probabilities for occurrences in time and space and associated magnitudes. One can define a marked spatial–temporal point process of earthquake locations, times and sizes. In California many faults have been located and found to be sources of earthquakes. These can be used. One might take d to be the distance to the nearest point on the fault from the site. The faults have been modelled as line segments and patches of planes. An event's magnitude can be related to its fault size. In Figure 2 just two sources have been hypothesized, but there could be many. Commonly renewal processes have been employed to model the sequence of times. The intervals between events may be assumed exponential, Weibull or lognormal. Basic references to earthquake statistics and seismic risk analysis include [34, 40].

(f) *An example*. As an example of a fair premium computation, consider a commercial building 25 km from the epicenter of an event like Northridge. For this case the estimated expected loss is

$$\begin{aligned} &0.8 \times 0.178 + 3 \times 0.266 \\ &\quad + 11 \times 0.416 + 27 \times 0.097 = 8.15\% \end{aligned}$$

assuming the highest MMI possible is IX, using the values of Table 1 and employing the results of fitting the model (3) to the Northridge data. The standard error of the estimate is 1.09%. It was obtained via the jackknife, splitting the data randomly into 10 groups.

Papers [5] and [8] contain details for a related example, namely the Loma Prieta event of 1989.

EXAMPLE 2 (Forest fires). (a) *Background*. There are tens of thousands of wildfires in North America each year destroying millions of acres of forests. For the years 1981 through 2002 a National Interagency Fire Control (NIFC) report [31] lists the annual counts of U.S. wildfires as ranging from 81,043 to 249,370. The acres burned ranged from 2,237,714 to 8,422,237. Another NIFC report [32] indicates that for the years 1994–2002 the yearly suppression costs for federal agencies ranged from 256 million to 1.661 billion U.S. dollars. That report further lists an average annual count of 13,879 fires caused by lightning and 102,694 caused by humans. Wildfires are clearly a serious problem.

Consider the task of predicting the average number of fires each day as a function of place and day for a region of interest to foresters. Specifically suppose that one is interested in base values. Dynamic predictions will be considered in later work.

(b) *A model*. Let occurrences be denoted by (x_j, y_j, t_j), $j = 1, 2, 3, \ldots$, with (x, y) location and t occurrence time. These values can be viewed as a realization of a spatial–temporal point process. To illustrate the idea consider Figure 6. The right panel shows the locations of forest fires in a large rectangular region containing the state of Oregon. Fires are indicated that occurred in federal lands during the period 1989–1996. These lands are indicated in the left panel of the figure and are seen to make up much of the state. To proceed to an analysis let space–time be broken up into voxels labelled by (x, y, t) and let

$$N_{x,y,t} = \begin{cases} 1, & \text{if a fire in the } (x, y, t) \text{ voxel,} \\ 0, & \text{otherwise.} \end{cases}$$

(In the computations the voxels have sizes of 1 km by 1 km by 1 day.) Next write

$$\mathrm{Prob}\{N_{x,y,t} = 1\} = p_{x,y,t}$$

and consider the model

$$\operatorname{logit} p_{x,y,t} = g_1(x, y) + g_2(d) + \zeta$$

with d the day of the year, and ζ a year effect. The g functions are assumed to be smooth and in the computations are represented by spline functions. The spatial term g_1 involved is a thin-plate spline,

$$g_1(x, y) = \sum_{j=1}^{J} \delta_j r_j^2 \log r_j,$$

where for nodes (x_j, y_j) the variable $r_j^2 = (x - x_j)^2 + (y - y_j)^2$ [35]. The day term g_2 is a spline with period 1 year.

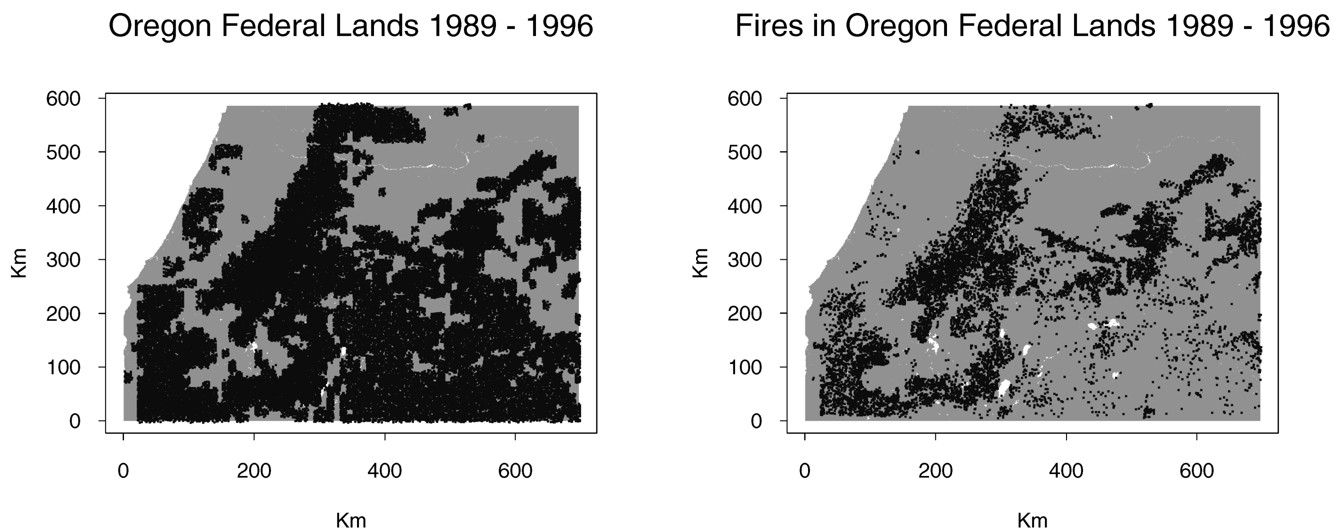

FIG. 6. *The left panel shows the Federal lands in a region containing Oregon. The right panel provides locations of fires in the region during the years 1989–1996.*

The data set for Oregon was very large, 578,192,400 voxels and 15,786 fires. To be able to carry out exploratory data analyses, a sample of the locations where no fires occurred was selected, while all the voxels with fires were employed. The chosen sampling fraction of the voxels with no fires was $\pi = 0.00012$. This lead to a total of 58,094 cases.

With the logit link, conditional on the sample, one had a generalized linear model (glm) with an offset of $\log 1/\pi$. It was a suprise that the new logit was simply $\operatorname{logit} p' = \operatorname{logit} p + \log(1/\pi)$, that is, an offset. This meant that standard generalized linear model computer programs could be used for the analysis. Logit models have been used previously in estimating fire risk; see, for example, [25].

(c) *Results*. The basic results are provided in Figure 7. One has estimates of the functions g_1, g_2 and the effects ζ. (The ζ are assumed fixed here but in work in progress they are random.) Examining the top panels one sees fewer fires in SE Oregon, as could have been anticipated from the right panel of Figure 6. From the bottom left panel of Figure 7 one notes a definite day effect—more fires in the summer. From the bottom right panel there appears to be a definite year effect. The year effect values are relative to 1996, which is taken to be level 0. The horizontal line is at 0. Also included in the bottom panels of the figure are ± 2 s.e. bounds. In the thin-plate computations 60 nodes were employed and they were taken to be 10 km apart throughout the region.

One can estimate the mean fire count for a nominated region and given time period by adding estimated probabilities for individual voxels. This is done for each day of the year for the Umatilla Forest. The results are presented in Figure 8 assuming that the year effect is 0. (The Umatilla Forest is approximately the rectangle with $450 < x < 500$ and $360 < y < 410$.) One obtains a Gaussian shaped curve peaking around day 220. The dashed lines provide approximate marginal 95% confidence limits.

If one desires statements concerning the count of fires, not just their average number, one can use the Hodges–Le Cam result [20] and approximate the distribution of the count by a Poisson.

Other results are presented in [9, 36]. Currently the work involves various explanatories such as fire danger indices, random year effects and extensions to other regions and states.

Turning to damage and insurance issues, standard homeowner policies cover wildfire damage. Kovacs [22], referring to an Insurance Services Office report, indicates insurance payments of \$3 billion in the 1990s following several large wildfires in California. For the Oregon data set there is information concerning the sizes of the fires. For example, one can study the probability of a fire becoming a large fire once it has started.

EXAMPLE 3 (Amazon floods). (a) *Background*. One can segue from studies of earthquake risk into studies of flooding risk by noting that earthquakes can cause dams to fail and thereby lead to flooding. The next example concerns the risk of floods on the

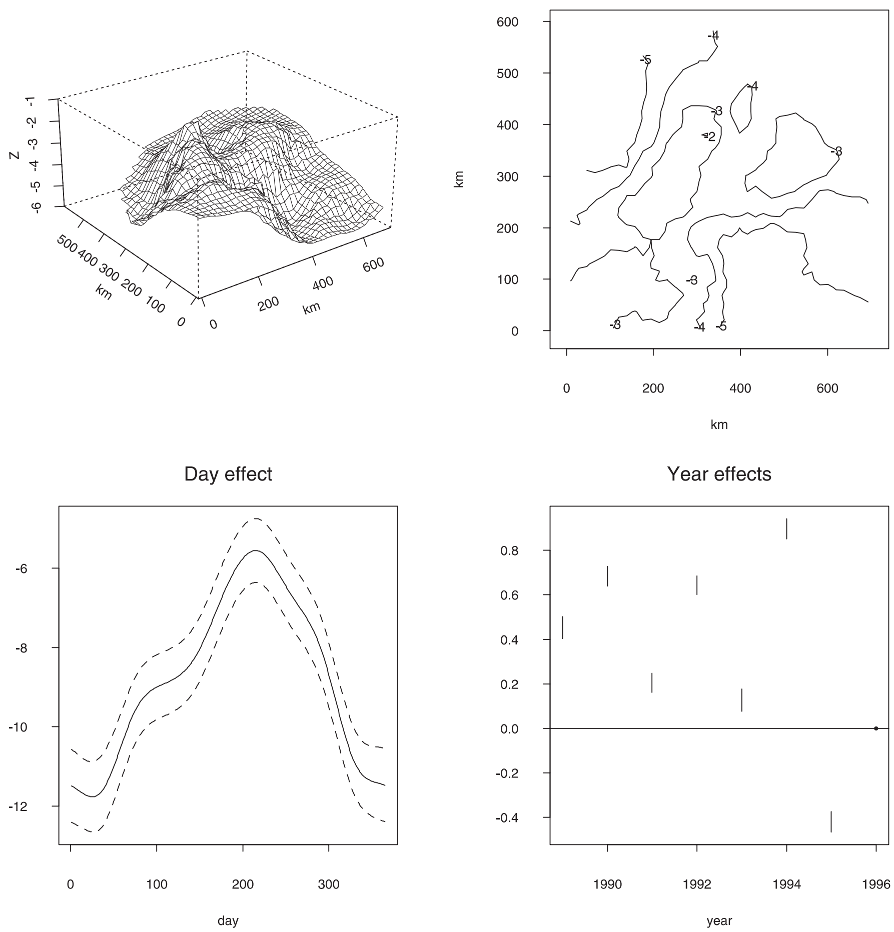

FIG. 7. *The top panels provide the estimated location effect $\hat{g}_1$ in perspective and contour form. The bottom left panel provides the estimated day effect $\hat{g}_2$. The bottom right panel shows the estimated year effect $\hat{\zeta}$. Approximate* 95% *error bounds are indicated in the bottom panels.*

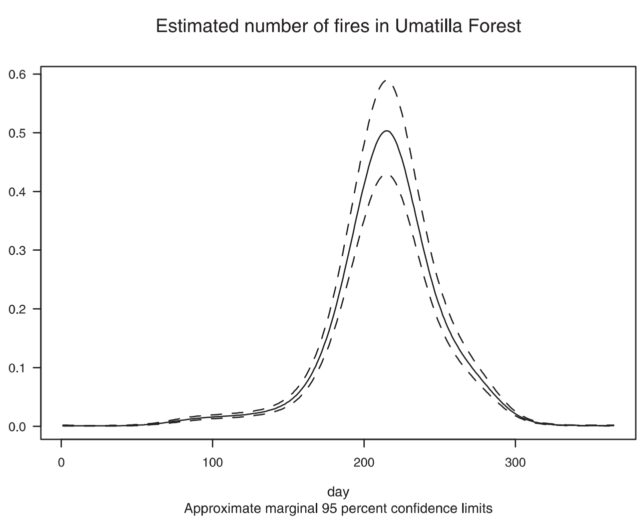

FIG. 8. *Estimated average number of fires for each day of the year for the Umatilla Forest region. The dashed lines provide approximate marginal* 95% *confidence limits.*

Amazon River at the city of Manaus. Data for the years 1892–2002 are employed.

Manaus is a city well up the Amazon River in Central Brazil. It is actually on the Rio Negro, but that river becomes the Amazon not far downstream from Manaus. At a dock in Manaus the Rio Negro's height has been recorded daily since 1903. Also there are newspaper records and journals that may be consulted to determine the dates of some earlier floods [38]. Of real concern is the question of whether the risk of flooding is increasing. Increased flooding will eventually occur because of the deforestation taking place. (See [38, 6].)

The top panel of Figure 9 provides the dates of serious floods in the period 1892–2002. The definition of a serious flood is that the water level exceeds 28.5 m [39]. For the data set studied there were 29 such floods. The lower panel of the figure displays the maximum level achieved each of the years 1903 through 2002. (These particular values were not available for the

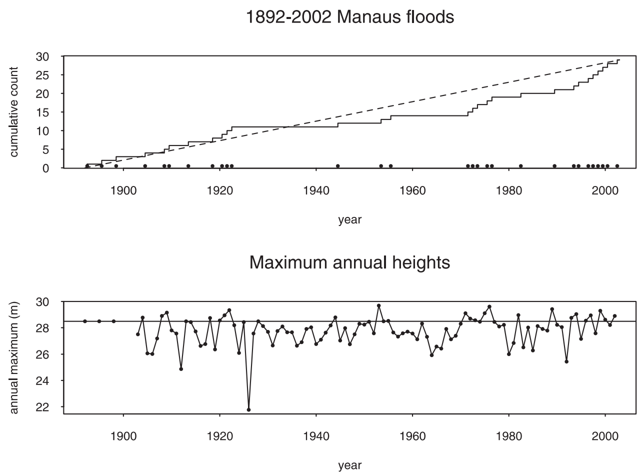

FIG. 9. *The top panel is a step function counting the number of floods since January 1892. The dashed line's slope provides the rate of events. The bottom provides the maximum height reached each year. The horizontal line is at* 28.5 m.

years before 1903.) The year 1924 is an outlier, there having been many fires that year [38].

Of concern is a suggestion of an increase in the number of floods and the maximum yearly height. To assess the reasonableness of this possibility we first examine the degree of serial correlation of the data. We do this, using the periodogram, for the binary series that is 1 the year of a flood and 0 otherwise and also for the results of annual maximum heights. The series are indicated in Figures 9 and 10. In the periodogram computation the outlier in the maximum height series was replaced by the mean of the other values. Also included in the figures are approximate 95% confidence limits. The results suggest independence is a reasonable working assumption. Not many points are outside the bounds.

Figure 11 top panel provides an estimate of the annual rate of serious floods and the bottom panel provides a trend function. Also provided are error bounds assuming independence. As with the earthquake data the computations employed the R/S-PLUS functions gam(·) and lo(·) and the complementary log–log link.

A form of damageability matrix is given in Table 2 [24]. It provides estimates of the numbers of persons affected as a function of the maximum height flood waters reach.

Both long-term and short-term prediction are important in this case. Long term refers to the question of whether there is a continuing trend. Short term can refer to whether there is a change in recent years.

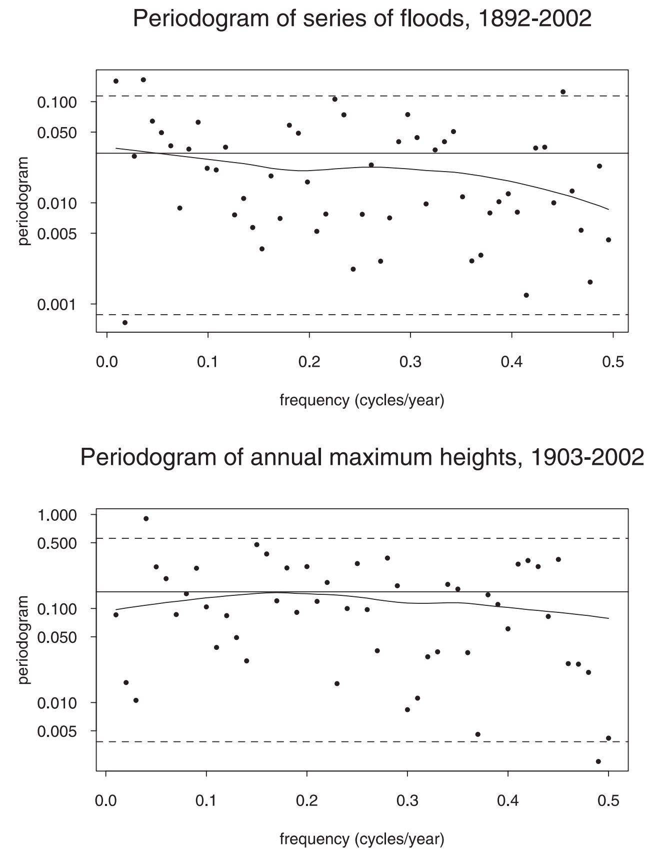

FIG. 10. *The top panel is the periodogram of the* 0–1 *series of Manaus floods. The bottom panel provides the periodogram of the series of annual maxima given in Figure* 9. *The dashed lines give approximate marginal* 95% *confidence limits. The solid curves are the result of smoothing using loess.*

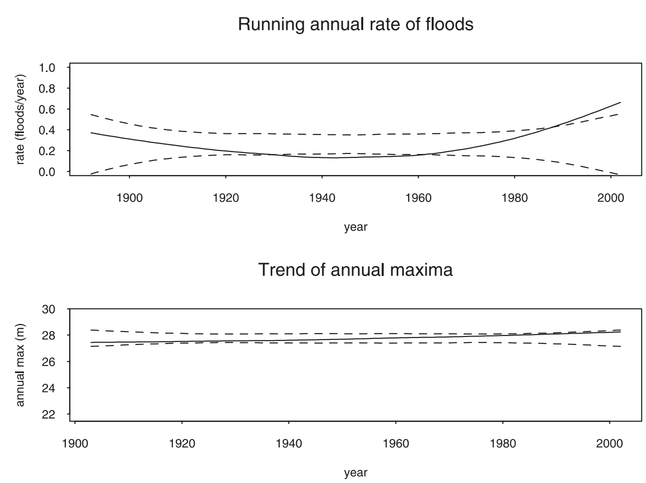

FIG. 11. *The top panel is a smooth estimate of the running rate determined from the* 0–1 *series of Manaus floods. The bottom panel provides a smooth estimate of the mean level. The dashed lines provide approximate marginal* 95% *confidence limits.*

TABLE 2
The estimated number of persons affected when the river level reaches the indicated height

Maximum height (m)	Number affected when height exceeded
27.00	1,000
27.50	2,000
28.00	6,000
28.50	15,000
29.00	30,000
29.50	50,000

Interestingly the authorities in Manaus have the following short-term risk assessment procedure:

> If at the end of March the height is greater than or equal to 26 m, then the chance of reaching 28.5 m is 92%. If, further, at the end of April it reaches 27.55 m the chance becomes 98%.

3. DISCUSSION AND CONCLUSIONS

The demand for risk analyses and insurance for environmental catastrophes is growing, in part because the costs of replacing destroyed structures are growing and in part because of the steady increase in the number of people living in hazardous areas. Statistical methods are basic to risk assessments and the computation of insurance premiums. In part this is obvious because probabilities and data are involved. It is also the case because statistics adds important things to what actuaries, engineers and scientists tend to know and do on their own. Statisticians offer things like efficiency results, extensions to different data types and uncertainty analyses.

Three examples have been presented. What do they have in common? Each is seeking probabilities and distributions. What do the solutions have in common? Data and subject matter are basic and the solutions each made use of the generalized linear model in some form. What has been learned from the examples presented? There are difficulties and opportunities. There are solutions and there are lots of open problems. The stochastic approach is highly effective. Considering the insurance problem has helped to focus the work.

ACKNOWLEDGMENTS

Supported by NSF Grants DMS-97-04739, DMS-99-71309 and DMS-02-03921 and U.S. Forest Service Contract 02-JV-11272165-020.

Collaborators are basic to research on problems of the type considered in this paper. I worked with John Benoit and Haiganoush Preisler of the U.S. Forest Service on the wildfire problem and with my colleague Hilgard O'Reilly Sternberg on the Rio Negro flooding question. Knut Aase commented on the insurance section. I thank them.

Jim Dewey of the USGS provided the Northridge data. Daniel Ortega of Companhia de Pesquisas de Recursos Minerais, Manaus, provided recent Rio Negro flood data.

People who provided help concerning the computing included Chong Gu, Trevor Hastie and Phil Spector. The Guest Editor and referees made comments improving the presentation of the paper considerably.

Part of the material was presented in the keynote talk at the 2002 Interface Meeting in Montreal.

REFERENCES

[1] AASE, K. K. (2001). A Markov model for the pricing of catastrophe insurance futures and spreads. *J. Risk and Insurance* **68** 25–49.

[2] AASE, K. K. (2002). Perspectives of risk sharing. *Skand. Actuar. J.* **2002**(2) 73–128.

[3] BEARD, R. E., PENTIKÄINEN, T. and PESONEN, E. (1969). *Risk Theory*. Chapman and Hall, London.

[4] BOHMAN, H. (1979). Insurance protection for large risks with a low claim frequency. *Bull. Int. Statist. Inst.* **48** 125–135.

[5] BRILLINGER, D. R. (1993). Earthquake risk and insurance. *Environmetrics* **4** 1–21.

[6] BRILLINGER, D. R. (1994). Trend analysis: Time series and point process problems. *Environmetrics* **5** 1–19.

[7] BRILLINGER, D. R. (1997). Random process methods and environmental data. *Environmetrics* **8** 269–281.

[8] BRILLINGER, D. R., CHIANN, C., IRIZARRY, R. A. and MORETTIN, P. A. (2001). Automatic methods for generating seismic intensity maps. *J. Appl. Probab.* **38A** 189–202. (*Probability, Statistics and Seismology,* Special volume of *J. Appl. Probab.*)

[9] BRILLINGER, D. R., PREISLER, H. K. and BENOIT, J. W. (2003). Risk assessment: A forest fire example. In *Statistics and Science* (D. R. Goldstein, ed.) 177–196. IMS, Beachwood, OH.

[10] BULLEN, K. E. and BOLT, B. A. (1985). *An Introduction to the Theory of Seismology*, 4th ed. Cambridge Univ. Press.

[11] CLARK, G. (1999). *Betting on Lives*. Manchester Univ. Press.

[12] CLEVELAND, W. S., GROSSE, E. and SHYU, W. M. (1992). Local regression models. In *Statistical Models in S* (J. M. Chambers and T. J. Hastie, eds.) 309–376. Wadsworth, Pacific Grove, CA.

[13] CORNELL, C. A. (1968). Engineering seismic risk analysis. *Bull. Seismol. Soc. Amer.* **58** 1583–1606.

[14] CRAMÉR, H. (1955). *Collective Risk Theory*. Nordiska, Stockholm. (Scandia Jubilee Volume.)

[15] DAYKIN, C. D., PENTIKÄINEN, T. and PESONEN, M. (1994). *Practical Risk Theory for Actuaries.* Chapman and Hall, London.

[16] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[17] FAIRLEY, W. B. (1977). Evaluating the "small" probability of a catastrophic accident from the marine transportation of liquefied natural gas. In *Statistics and Public Policy* (W. B. Fairley and F. Mosteller, eds.) 331–354. Addison-Wesley, Reading, MA.

[18] GRANDELL, J. (1991). *Aspects of Risk Theory*. Springer, New York.

[19] HASTIE, T. J. (1992). Generalized additive models. In *Statistical Models in S.* (J. M. Chambers and T. J. Hastie, eds.) 249–308. Wadsworth, Pacific Grove, CA.

[20] HODGES, J. L., JR. and LE CAM, L. M. (1960). The Poisson approximation to the Poisson binomial distribution. *Ann. Math. Statist.* **37** 737–740.

[21] JOYNER, W. B. and BOORE, D. M. (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bull. Seismol. Soc. Amer.* **71** 2011–2038.

[22] KOVACS, P. (2001). Wildfires and insurance. Research Paper 11, Inst. Catastrophic Loss Reduction, Toronto.

[23] KUNREUTHER, H. (1998). Insurability conditions and the supply of coverage. In *Paying the Price* (H. Kunreuther and R. J. Roth, Sr., eds.) 17–50. Henry Press, Washington.

[24] LA ROVERE, A. L. N. and CRESPO, S. (2002). *Relatório Urbano Ambiental Integrado—Informe GEO—Manaus*. IBAM/ISER/REDEH, Brazil.

[25] MARTELL, D. L., OTUKOL, S. and STOCKS, B. J. (1987). A logistic model for predicting daily people-caused forest fire occurrence in Ontario. *Canad. J. Forest Res.* **17** 394–401.

[26] MCCULLAGH, P. and NELDER, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman and Hall, London.

[27] MILLER, A. J. and LESLIE, J. S. (1979). The risk of collisions between ships and bridges. *Bull. Int. Statist. Inst.* **48** 137–148.

[28] MUNICH RE (1991). *Insurance and Reinsurance of Earthquake Risk*. Munich Re, Munich.

[29] MUNICH RE (2002). *Annual Review*: *Natural Catastrophes 2002*. Munich Re, Munich.

[30] NATIONAL ACADEMY OF SCIENCES (1998). *Paying the Price* (H. Kunreuther and R. J. Roth, Sr., eds.) Henry Press, Washington, DC.

[31] NATIONAL INTERAGENCY FIRE CENTER (1999). 10-year comparison statistics. Available at www.nifc.gov/fireinfo/1999/hilites.html.

[32] NATIONAL INTERAGENCY FIRE CENTER (2002). Wildland fire statistics. Available at www.nifc.gov/stats/wildlandfirestats.html.

[33] NUCLEAR REGULATORY COMMISSION (1997). *Federal Register* **62**(122) (June 25, 1997) 34,321–34,326.

[34] OGATA, Y. (1983). Likelihood analysis of point processes and its application to seismological data. *Bull. Int. Statist. Inst.* **50** 943–961.

[35] POWELL, M. J. D. (1992). The theory of radial basis function approximation in 1990. In *Advances in Numerical Analysis* **2** (W. Light, ed.) 105–210. Oxford Univ. Press.

[36] PREISLER, H. K., BRILLINGER, D. R., BURGAN, R. E. and BENOIT, J. W. (2004). Probability based models for the estimation of wildfire risk. *Int. J. Wildland Fire*. To appear.

[37] RAYNES, H. E. (1964). *A History of British Insurance*. Pitman, London.

[38] STERNBERG, H. O'R. (1987). Aggrevation of floods in the Amazon River as a consequence of deforestation? *Geografiska Annaler A* **69** 201–219.

[39] STERNBERG, H. O'R. (1988). Personal communication.

[40] VERE-JONES, D. (1973). The statistical estimation of earthquake risk. *New Zealand Statistician* **8** 7–16.