

S&G 3703

S&G3703

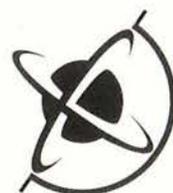
99/388

Client Report 2000/118

**Probability of
Rupture of the
Alpine Fault
Allowing for
Uncertainties**

**DA Rhoades &
RJ Van Dissen**

September 2000



Institute of
**GEOLOGICAL
& NUCLEAR
SCIENCES**
Limited



Institute of
**GEOLOGICAL
& NUCLEAR
SCIENCES**
Limited

Probability of Rupture of the Alpine Fault Allowing for Uncertainties

DA Rhoades & RJ Van Dissen

Prepared for

**EQC Research Foundation
Project No 99/388**

**Institute of Geological & Nuclear Sciences client report 2000/118
Project Number: 41999U**

**The data presented in this Report are
available to GNS for other use from
October 2000**

September 2000

COMMERCIAL - IN - CONFIDENCE

This report has been prepared by the Institute of Geological & Nuclear Sciences Limited exclusively for and under contract to EQC Research Foundation. Unless otherwise agreed in writing, all liability of the Institute to any other party other than EQC Research Foundation in respect of the report is expressly excluded.



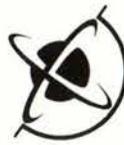
CONTENTS

EXECUTIVE SUMMARY	1
1.0 INTRODUCTION	2
2.0 SUMMARY OF CURRENT INFORMATION ON PAST ALPINE FAULT RUPTURES	2
3.0 STATISTICAL METHODOLOGY.....	4
4.0 DETAILS OF ESTIMATION UNDER THE VARIOUS MODELS	7
4.1 Sampling from the Distribution of the Mean Recurrence Time.....	7
4.2 Sampling from the Distribution of Specific Recurrence Times	7
4.3 Exponential Model	8
4.4 Lognormal Model.....	8
4.5 Weibull Model.....	8
4.6 Inverse Gaussian Model	9
4.7 Numerical Procedure	9
5.0 RESULTS.....	10
5.1 Hazard Estimates for North-east Section of Alpine Fault	12
5.2 Hazard Estimates for South-west Section of Alpine Fault.....	15
5.3 Sensitivity of Results to Standard Error of Date of Second to Last Event.....	17
5.4 Comparison with Previous Estimates	17
6.0 CONCLUSIONS.....	18
7.0 ACKNOWLEDGEMENTS	18
8.0 REFERENCES	18
9.0 APPENDIX.....	21
NON-TECHNICAL ABSTRACT.....	22
TECHNICAL ABSTRACT	23



EXECUTIVE SUMMARY

The time-varying hazard of rupture on the Alpine fault is estimated using a technique that takes account of uncertainties in data and parameter values. The north-east (Karangarua-Haupiri) and south-west (Haast) sections of the fault are considered separately, i.e., it is not assumed that they will necessarily rupture together. Data inputs are based on geological estimates of the long-term slip rate and previous studies of fault traces and forest ages and times of disturbance. The geological average strike-slip rate is taken to be 26 ± 5 mm/yr and the average single-event displacements to be 5.0 ± 1.4 m and 8.0 ± 2.6 m on the north-east and south-west sections, respectively. The last four events on the north-west section are dated at 1717, 1620 ± 10 , 1445 ± 20 and 1150 ± 50 , and the last three events on the south-west section at 1717, 1450 ± 100 and 1150 ± 50 A.D. Using these data and associated uncertainties, the current hazard of rupture on the north-east section of the fault is estimated to be 0.0049, 0.0092, 0.0104 and 0.0074 events per year under the exponential, lognormal, Weibull and inverse Gaussian models respectively. The corresponding probabilities of rupture in the next 20 years are 9%, 17%, 19% and 14%, respectively. The current hazard on the south-west section of the fault is estimated to be 0.0032, 0.0072, 0.0064 and 0.0052 events per year for the four models, and the 20 year probabilities 6%, 14%, 13%, and 10%, respectively. Increased precision in the date of the second to last event on the south-west section of the fault would result in only small changes to these rates and probabilities. The hazard under the lognormal model is about double the long-term average (exponential) rate but less than half of that previously estimated without taking account of uncertainties in the data and parameter values.



1.0 INTRODUCTION

In the vicinity of the Alpine fault, recent investigations of fault traces (Cooper and Norris 1990, 1995; Berryman *et al.*, 1992; Norris and Cooper, in press; Yetton *et al.*, 1998), forest ages and times of disturbance (Wells *et al.*, 1998, 1999), and rock-fall and lichen growth (Bull *et al.* 1994; Bull, 1996; Bull & Brandon, 1998) have greatly improved information on the times of rupture of segments of the fault over the past thousand years. Geological and geodetic observations and associated modelling have helped to improve estimates of the long-term slip rate (e.g. Sutherland and Norris, 1995; Beavan *et al.*, 1999). Much of the data on past rupture on the Alpine fault has been brought together by Yetton *et al.* (1998) who also undertook statistical modelling of recurrence, aimed at estimating the current and future probability of rupture. However, uncertainties in the data, including dates of rupture, long-term slip rate and magnitude of events, and in the parameters of the recurrence time model, have not been integrated into their hazard estimates. Estimates of the current hazard may thus be higher than the data justify. A methodology for incorporating uncertainties developed by Rhoades *et al.* (1994) is elaborated upon and used here to bring together both data and parameter uncertainties into a single estimate of the time varying probability of rupture of sections of the Alpine fault. The choice of model is independent of the data, but its effect on the hazard estimates may be pronounced. The usual lognormal recurrence time distribution, proposed as a generic model by Nishenko and Buland (1987), is compared to the well-known Weibull and exponential models and to the inverse Gaussian, i.e., Brownian passage-time model, recently proposed by Ellsworth *et al.* (1999), to ascertain the influence of model choice on the hazard estimates

2.0 SUMMARY OF CURRENT INFORMATION ON PAST ALPINE FAULT RUPTURES

The following observations, deductions and assumptions are admitted for the present study. Uncertainties are standard errors, unless otherwise stated.

1. The geological average slip rate (5,000-50,000 years) is 25-30 mm/yr. This is the total slip rate, including the vertical component.
2. The geological average strike-slip rate is 26 ± 5 mm/yr. This is the horizontal component of the slip rate only.
3. The geodetic strike-slip rate (5 years) is 24 ± 3 mm/yr.
4. The seismic slip (paleoseismology, 1000 years) is 25 ± 5 mm/yr.
5. The north-east (Karangarua-Haupiri) section of the fault has had four major ruptures in the past 1000 years with a total displacement of 20+ metres.



6. The south-west (Haast) section of the fault has had three major ruptures in the last 1000 years with a total displacement of 24 metres.
7. It is not assumed that the south-west and north-east sections rupture at the same time, although it is probable that they did rupture at the same time in 1717 and in c.1150 AD.
8. The last four ruptures of the north-east section of the fault occurred in 1717, 1620 \pm 10 and 1445 \pm 20 and 1150 \pm 50 AD.
9. The last three ruptures of the south-west section of the fault occurred in 1717, 1450 \pm 100 (uncertainty not well determined) and 1150 \pm 50 AD.

The information in 1-9 above was almost entirely derived from a summary of agreed conclusions on past events and rates on the Alpine fault that emerged from a combined workshop on the Alpine fault held in September 1998. The summary was prepared and made available to us by R.J. Norris.

The central date and standard error of the second to last event on the south-west section did not come from the summary of agreed conclusions referred to above, although a wide uncertainty in this date was acknowledged. The standard error of 100 years on this date is intended to reflect the wide uncertainty, and the central value of 1450 AD is consistent with a rather regular recurrence time on the south-west section of the fault. In the hazard calculations described below, the alternative value of 50 years was also considered for the standard error, to assess the sensitivity of the hazard to the level of uncertainty in the date of this event.

The measurements (5 and 6) of total displacement in recent major ruptures give mean single-event displacements of 5m and 8m, respectively, on the north-east and south-west sections of the fault, but do not afford estimates of the uncertainty of these means. To estimate the uncertainty on these values, it is necessary to look to more detailed studies that have been carried out elsewhere.

Stein *et al.* (1997) have presented data of the amount of right lateral slip occurring in 10 earthquakes during 1939-1992 on the North Anatolian Fault, Turkey, at different points along the fault. Their figure 1 shows that the amount of slip in a single event typically varies between locations along the fault, and at a few locations where more than one event has been measured the amount of slip varies between events. The Appendix Table A.1 tabulates slips for individual events measured from figure 1 of Stein *et al.* (1997) at intervals of 25 km along the fault.

The data from Table A.1 can be used to get an indication of the variability of displacements in earthquakes on a single fault, as follows. The within-earthquake



(between locations) coefficient of variation is 0.59 and the within location (between earthquakes) coefficient of variation is 0.57. Assuming the same coefficients of variation on the Alpine fault, the mean of four observations will have a coefficient of variation of $0.57/\sqrt{4}$, i.e., 0.285, and the mean of three observations will have a coefficient of variation of $0.57/\sqrt{3}$, i.e., 0.329. This gives standard deviations of 0.285×5 and 0.329×8 metres respectively for the mean single-event displacement on the north-east and south-west section of the fault.

A New Zealand example of measurements of repeated movements at a single location, although not on a strike-slip fault, is provided by the sequence of uplifted beaches at Turakirae Head, near Wellington. The most recent uplift occurred in the 1855 Wairarapa earthquake. The sizes of the last four uplifts at Turakirae Head are estimated to be 5.98, 9.13, 5.51 and 3.00 metres (Hull and McSaveney, 1996), giving a coefficient of variation of 0.43, not dissimilar to the value of 0.57 for repeated ruptures on the North Anatolian fault. The value from the North Anatolian fault is adopted, because it is supported by a larger data set and represents single event displacements on a strike slip fault.

We thus assume the following in addition to 1-9 above.

10. The mean single event displacement at the point where measurements were made on the north-east section of the Alpine fault is 5.0 ± 1.4 m.
11. The mean single-event displacement at the point where measurements were made on the south-west section of the Alpine fault is 8.0 ± 2.6 m

The uncertainties in mean single event displacement, like those in average slip rate are assumed to follow a lognormal distribution. This assumption ensures that any sampled values from these distributions will be positive.

3.0 STATISTICAL METHODOLOGY

A method for handling uncertainties when estimating fault hazard in a recurrence-time modelling context was presented by Rhoades *et al.*(1994). Both data values and parameter values enter into their analysis by way of probability distributions. In this way they explicitly considered both data uncertainties and parameter uncertainties, and incorporated them into the hazard estimate by averaging over these distributions.

Following the analysis of Rhoades *et al.* (1994), let f denote the recurrence time probability density function for events that rupture the fault at some point. Then f depends explicitly on the parameter values θ of the assumed recurrence time model α . The parameter values in turn depend on the data x . Let $f(t|\theta, x; \alpha)$ denote the recurrence time distribution for particular values of θ , x and α . Then



$$f(t | \mathbf{x}; \alpha) = \int_{\theta} f(t | \theta, \mathbf{x}; \alpha) g(\theta | \mathbf{x}; \alpha) d\theta \quad (1)$$

where $g(\theta | \mathbf{x}, \alpha)$ is the conditional density of θ given \mathbf{x} under the model α , commonly called the “likelihood function”. The hazard at time t since the last rupture of the fault, conditional on the particular data values \mathbf{x} , is given by

$$h(t | \mathbf{x}; \alpha) = \frac{f(t | \mathbf{x}; \alpha)}{1 - F(t | \mathbf{x}; \alpha)}, \quad (2)$$

where F is the cumulative distribution function associated with the probability density function f . Averaging over the distribution of possible data values, we have

$$h(t | \alpha) = \int_{\mathbf{x}} h(t | \mathbf{x}; \alpha) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \quad (3)$$

Thus, averaging over parameter values is by a “mixture of distributions” (1) and the averaging over data values is by a “mixture of hazards” (3). The mixture of distributions is the means by which information on the repose time since the last rupture is incorporated into the parameter estimates. As time passes without an earthquake occurring, those values of θ that are more consistent with longer repose times automatically receive a higher weighting in the mixed density (1), and hence in the hazard also (2). Thus, the hazard function of the mixed distribution does not necessarily have the same properties as the hazard function of a particular distribution following the same model.

Here we consider four different recurrence time models, giving in each case the density function $f(t | \theta; \alpha)$ and cumulative distribution function $F(t | \theta; \alpha)$.

α_1 . Exponential:

$$f(t | \lambda; \alpha_1) = \lambda \exp(-\lambda t) \quad (t > 0; \lambda > 0). \quad (4)$$

$$F(t | \lambda; \alpha_1) = 1 - \exp(-\lambda t). \quad (5)$$

α_2 . Lognormal:

$$f(t | \mu, \sigma; \alpha_2) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{(\log t - \mu)^2}{\sigma^2}\right] \quad (t > 0; \mu > 0; \sigma > 0). \quad (6)$$



$$F(t | \mu, \sigma; \alpha_2) = \Phi\left(\frac{\log t - \mu}{\sigma}\right). \quad (7)$$

where Φ is the standard normal integral, i.e., $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^x \exp(-u^2/2) du$.

α_3 . Weibull:

$$f(t | \beta, c; \alpha_3) = \frac{c}{\beta} \left(\frac{t}{\beta}\right)^{c-1} \exp\left[-\left(\frac{t}{\beta}\right)^c\right] \quad (t > 0; \beta > 0; c > 0). \quad (8)$$

$$F(t | \beta, c; \alpha_3) = 1 - \exp\left[-\left(\frac{t}{\beta}\right)^c\right]. \quad (9)$$

α_4 . Inverse Gaussian:

$$f(t | \mu, \eta; \alpha_4) = \sqrt{\frac{\mu}{2\pi\eta^2 t^3}} \exp\left[-\frac{(t-\mu)^2}{2\eta^2 \mu t}\right] \quad (t > 0; \mu > 0; \eta > 0). \quad (10)$$

$$F(t | \mu, \eta; \alpha_4) = \Phi\left(\frac{t-\mu}{\eta\sqrt{\mu t}}\right) + \exp\left(\frac{2}{\eta^2}\right) \Phi\left(\frac{t+\mu}{\eta\sqrt{\mu t}}\right). \quad (11)$$

The exponential distribution (4) is the model of least information. It is the model associated with rupture times occurring according to a Poisson process, most commonly assumed in probabilistic seismic hazard analysis. The parameter λ represents the hazard rate and the mean recurrence time is $1/\lambda$. If uncertainties in parameters and data are ignored, the hazard under this model is constant.

The lognormal distribution (6) has been widely used in the context of fault rupture recurrence modelling and has been proposed as a generic model (Nishenko and Buland, 1987). The parameters μ and σ are the mean and standard deviation of the logarithm of the recurrence time. The mean recurrence time is $\exp(\mu + \frac{1}{2}\sigma^2)$ and the coefficient of variation is $\sqrt{\exp(\sigma^2) - 1}$. If uncertainties in parameters and data are ignored, the hazard under this model increases from a low value immediately after the last rupture to some maximum, after which it begins to decline and approaches zero as $t \rightarrow \infty$.



The Weibull distribution (8) is widely used in failure time modelling for manufactured items, where “new is better than used”, and has been proposed as a model of fault rupture recurrence (Hagiwara, 1974). The mean recurrence time is $\beta\Gamma(c^{-1} + 1)$ and c is a shape parameter. If uncertainties in parameters and data are ignored and if $c > 1$, the hazard under this model increases monotonically starting from zero at the time of the last rupture.

The inverse Gaussian distribution (10), which is the distribution of the first passage time of Brownian motion with positive drift (Johnson & Kotz, 1970), has recently been proposed by Ellsworth *et al.* (1999) as a physically realistic model of earthquake recurrence. The mean recurrence time is μ and η is a dispersion parameter, called the aperiodicity. If uncertainties in parameters and data are ignored, the hazard of rupture increases from zero at the time of the last rupture to some maximum, after which it tails off to a rate of $[1/(2\mu\eta^2)]$ as $t \rightarrow \infty$. Ellsworth *et al.* (1999) suggested $\eta = 1/2$ as a generic value of the aperiodicity of recurrence-time distributions. At this value, the asymptotic hazard rate is $2/\mu$, i.e., twice the average hazard rate.

4.0 DETAILS OF ESTIMATION UNDER THE VARIOUS MODELS

4.1 Sampling from the Distribution of the Mean Recurrence Time

The data include distributions for the long-term average slip rate, and the mean single event displacement at specific points on the fault. If single values, r and d respectively, are sampled independently from these distributions, then the mean recurrence time T associated with this sample is

$$T = \frac{d}{r}. \quad (12)$$

A large sample of T -values generated in this way will reflect our knowledge of the distribution of the average recurrence time. This distribution can be fed in directly to the various statistical models as prior information on (one of) the parameter values.

4.2 Sampling from the Distribution of Specific Recurrence Times

The only other data to be incorporated is the information on specific recurrence times at a given location. Each rupture date has a standard error, which is assumed to represent the standard deviation of a normal distribution. To construct a single sample of the last two and three recurrence times from the north-western and south-eastern sections of the fault, respectively, we generate a random sample independently from the distribution of each rupture date and take differences.



4.3 Exponential Model

Each value of $1/T$, where T is a sample from the distribution of the mean recurrence time generated as described above, represents a sample from the prior distribution of the parameter λ of the exponential recurrence-time distribution. Now let T_1, \dots, T_k be samples of the last k specific recurrence times. Then the likelihood function is

$$f(\lambda | T_1, \dots, T_k) \propto \lambda^k \exp(-\lambda \sum_{i=1}^k T_i). \quad (13)$$

4.4 Lognormal Model

Each value of T , where T is a sample from the distribution of the mean recurrence time generated as described above, represents a sample from the prior distribution of the mean γ of the lognormal recurrence-time distribution. The prior distribution of the coefficient of variation is taken to be uniform on the interval $(0,1)$. A sample (μ_s, σ_s) from the prior distribution of (μ, σ) is given by

$$\mu_s = \log T - \frac{1}{2} \log(\delta_s^2 + 1), \quad (14)$$

$$\sigma_s = \sqrt{\log(\delta_s^2 + 1)}, \quad (15)$$

where δ_s is a sample from the prior distribution of δ . If T_1, \dots, T_k are samples from the distribution of the last k specific recurrence times, then the likelihood function is

$$f(\mu_s, \sigma_s | T_1, \dots, T_k) \propto \frac{1}{\sigma_s^k (2\pi)^{k/2}} \prod_{i=1}^k \frac{1}{T_i} \exp\left[-\frac{1}{2} \frac{(\log T_i - \mu_s)^2}{\sigma_s^2}\right]. \quad (16)$$

4.5 Weibull Model

Each value of T , where T is a sample from the distribution of the mean recurrence time generated as described above, represents a sample from the prior distribution of the mean $\beta\Gamma(c^{-1} + 1)$ of the Weibull recurrence-time model. The prior distribution of the parameter c is taken to be such that $1/c$ is uniformly distributed on the interval $(0,1)$. Thus we are only entertaining the possibility of values of c greater than 1. A sample β_s from the prior distribution of β is given by

$$\beta_s = T / \Gamma(c_s^{-1} + 1) \quad (17)$$

where c_s is a sample from the prior distribution of c . If T_1, \dots, T_k are samples from the distribution of the last k specific recurrence times, then the likelihood function is



$$f(\beta_s, c_s | T_1, \dots, T_k) \propto \left(\frac{c_s}{\beta_s}\right)^k \prod_{i=1}^k \left(\frac{T_i}{\beta_s}\right)^{c_s-1} \exp\left[-\left(\frac{T_i}{\beta_s}\right)^{c_s}\right]. \quad (18)$$

4.6 Inverse Gaussian Model

Each value of T , where T is a sample from the distribution of the mean recurrence time generated as described above, represents a sample from the prior distribution of the mean μ of the inverse Gaussian recurrence-time distribution. The prior distribution of the aperiodicity η is taken to be uniform on (0,1). Thus we are not entertaining the possibility of the recurrence-time distribution being highly aperiodic. If T_1, \dots, T_k are samples from the distribution of the last k specific recurrence times, then the likelihood function is

$$f(\mu_s, \eta_s | T_1, \dots, T_k) \propto \left(\frac{\mu_s}{2\pi\eta_s^2}\right)^k \prod_{i=1}^k \left(\frac{1}{T_i^3}\right) \exp\left[-\frac{(T_i - \mu_s)^2}{2\eta_s^2 \mu_s T_i}\right] \quad (19)$$

where $\mu_s = T$ and η_s is a sample from the prior distribution of η .

4.7 Numerical Procedure

Taking account of the above considerations, the following numerical procedure was carried out to compute the hazard estimates for each section of the fault, under each of the models.

1. Generate a sample of size n from the distribution of the geological average strike-slip rate in mm/yr, i.e., from the lognormal distribution $L(26,5)$ with mean 26 and standard deviation 5.
2. Generate a corresponding sample of size n from the distribution of the mean single event displacement in metres, i.e. from the lognormal distributions $L(5,1.4)$ and $L(8,2.6)$ for the north-east and south-west sections of the fault, respectively.
3. Hence calculate n sample values from the distribution of the mean recurrence-time T , as described above.
4. Generate a corresponding sample of size n from the distribution of the times of occurrence, assumed normal, of individual ruptures during the last thousand years on each section of the fault. By subtraction obtain a sample of the last k recurrence-times T_1, \dots, T_k , where $k=3$ for the north-east section and $k=2$ for the south-west section of the fault.



5. Each sample value of (T, T_1, \dots, T_k) comprises a sample data vector x .
6. For each of the n sample values of x , generate a further sample of size m from the conditional distribution of the parameters θ given x , using the likelihood functions described in equations (13-19). A method for generating random samples from a distribution was described by Rhoades *et al.* (1994). The procedure used here is the same, except that the sample values of θ are selected initially from a prior distribution as described in the section above. They are then accepted or rejected by a random procedure in which the probability of acceptance is proportional to the value of the likelihood function.
7. For each sample value of x , calculate the density function for the time from the last rupture to the next as a mixture of distributions (c.f. equation (1)):

$$f(t | x, \alpha) = \frac{1}{m} \sum_{j=1}^m f(t | \theta_j, x; \alpha). \quad (20)$$

8. Calculate the hazard function $h(t | x; \alpha)$ by equation (2).
9. Perform the hazard mixing over the data sample, i.e., calculate

$$h(t | \alpha) = \frac{1}{n} \sum_{i=1}^n h(t | x_i; \alpha). \quad (21)$$

10. Finally, the probability of an earthquake occurring in any time interval (t_1, t_2) of interest is calculated by

$$P[E_{(t_1, t_2)}] = 1 - \exp\left[-\int_{t_1}^{t_2} h(t | \alpha) dt\right]. \quad (22)$$

5.0 RESULTS

The method outlined above was applied with the number of data samples set at $n=200$, and the number of parameter samples for each data sample set at $m=30$. These numbers, although somewhat arbitrary, are sufficiently large to make the results repeatable to the level of precision reported here.

Figure 1 illustrates the differing hazard functions for samples from the distribution of parameters under the four different models. Each of the hazard functions illustrated conforms exactly to the standard form of the model concerned. Thus, for example, all of

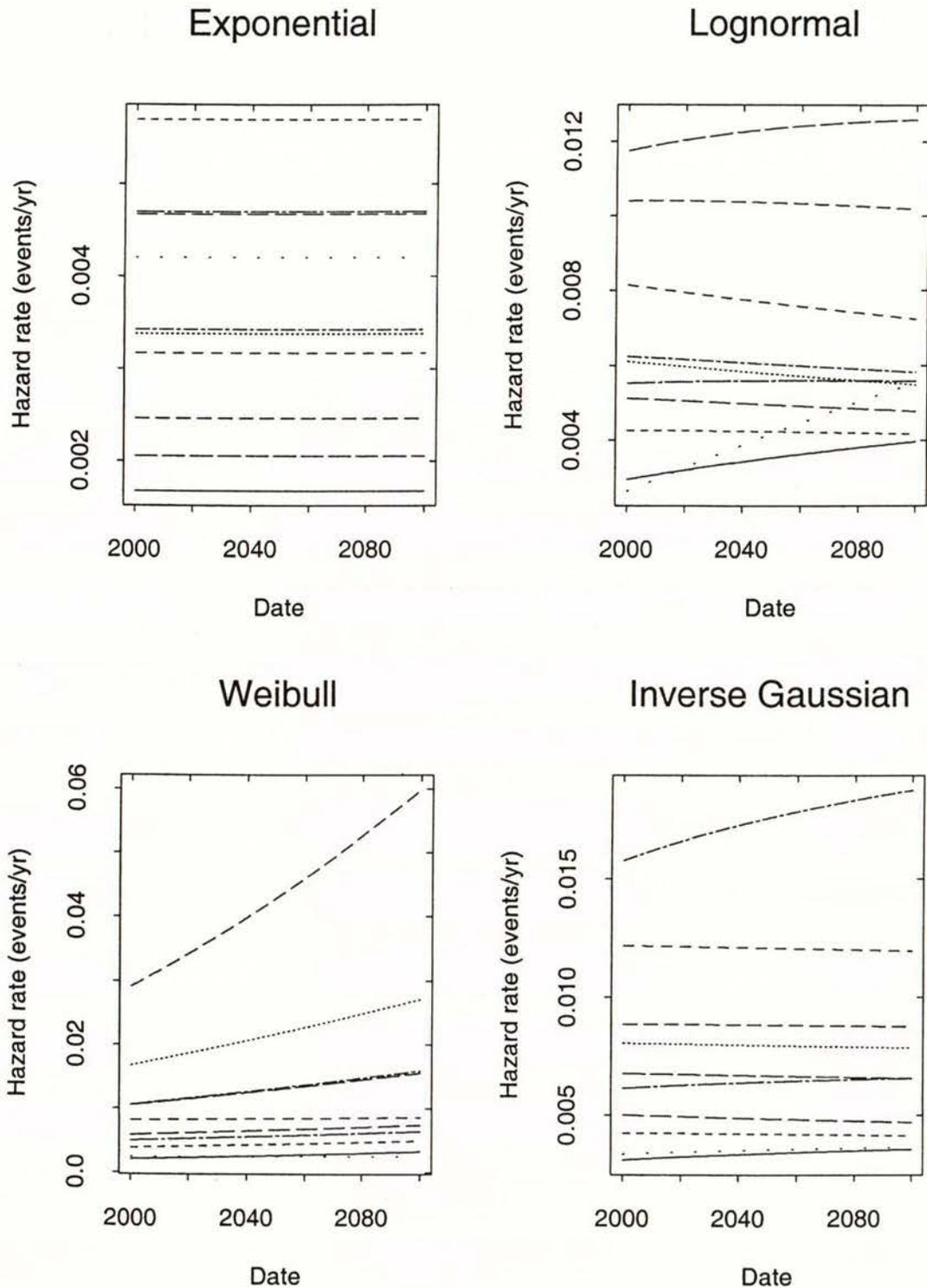


Figure 1. Hazard functions for ten samples from the distribution of parameter values under the exponential, lognormal, Weibull, and inverse Gaussian recurrence-time models applied to the south-west section of the Alpine fault. Note that the vertical scale differs between plots.



the exponential hazard functions are constant, and all of the Weibull hazard functions are monotone increasing, while the lognormal and inverse Gaussian hazard functions can be increasing or decreasing in the range of times plotted, depending on the particular parameter values sampled.

The hazard curves in Figure 1 can be contrasted with those in Figure 2, which represent the hazard functions for samples from the distribution of data. Here, the distributions have been averaged over all sampled parameter values. The hazard functions of the mixed distributions do not conform to the standard form of the model concerned. For example, the exponential hazard functions are now monotone decreasing because, as the time since the last movement increases, the estimate of hazard is increasingly re-weighted in favour of parameters consistent with longer mean recurrence-times. Similarly, for the other three models a wider range of hazard function shapes is now possible, e.g., the Weibull hazard functions are not necessarily monotone.

5.1 Hazard Estimates for North-east Section of Alpine Fault

Figure 3 shows the mean hazard function under each of the models for the north-east section of the fault. This curve is the average of the hazard functions over all samples from the data distribution computed by equation (21). It shows the rate of hazard that will be applicable at any time between the years 2000 and 2100, if no fault rupture occurs before that time, taking into account the uncertainties in data and parameter values.

The differences between the models are quite clear. The exponential model gives the lowest hazard, at a slightly declining rate of about 0.005 events per year. This represents an average recurrence of about 200 years. The decline in the hazard rate over time for all models is the adjustment for the increasing elapsed time since the last rupture. This adjustment is achieved by the mixing of distributions for different parameter values, as described above. The Weibull model gives the highest hazard, at about 0.01 events per year or twice the exponential model rate. This was to be expected because this was the only model for which the underlying hazard functions are all monotone increasing. The lognormal and inverse Gaussian models give intermediate levels of hazard at about 0.009 and 0.007 events per year respectively.

Table 1 shows estimates of the probabilities of rupture of the fault on the north-east section for the next year, 20 years, 50 years and 100 years. These values were computed from the hazard curves of Figure 3 using equation (22). These values are in general much lower than corresponding estimates by Yetton et al (1998), which were based on the lognormal distribution and did not have regard to uncertainties in the data or parameter values.

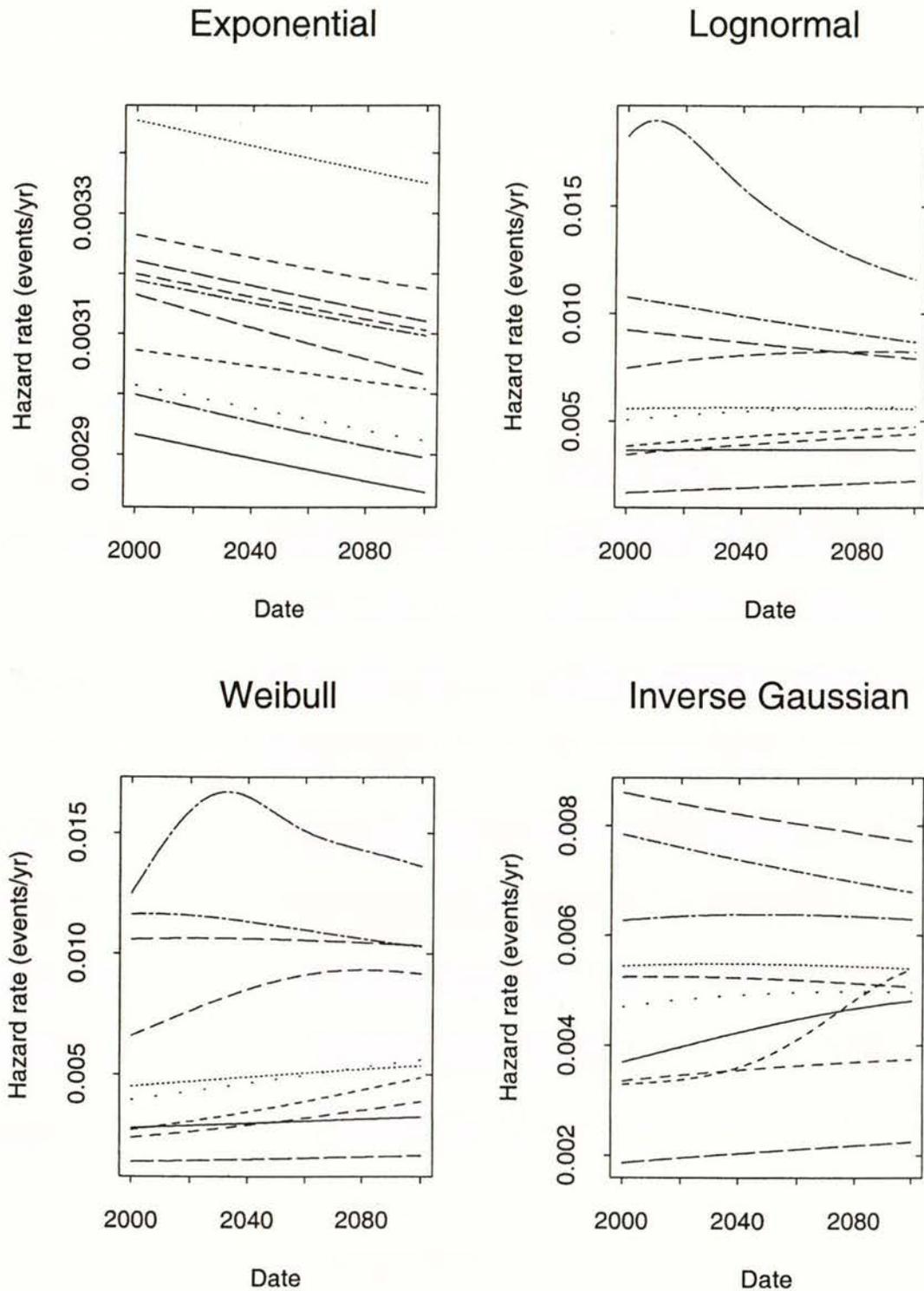


Figure 2. Hazard functions of the mixed distributions, i.e., averaged over all sampled parameter values, for ten samples from the distribution of data under the exponential, lognormal, Weibull, and inverse Gaussian recurrence-time models applied to the south-west section of the Alpine fault. Note that the vertical scale differs between plots.



Hazard on NE section of Alpine fault

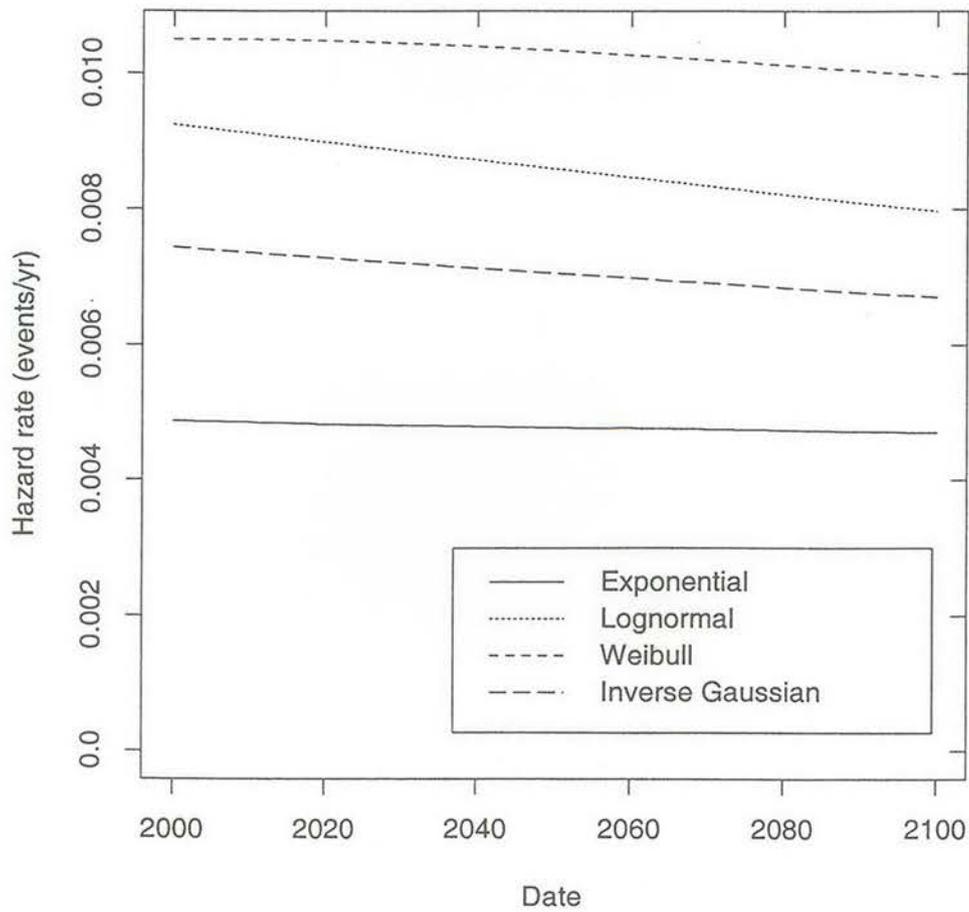


Figure 3. Hazard function averaged over all sampled data values under the exponential, lognormal, Weibull, and inverse Gaussian recurrence-time models applied to the north-east section of the Alpine fault.



east section of the Alpine fault.

Table 1. Estimated probability of rupture of the Alpine fault at points on its north-east section during time intervals starting in the year 2000 A.D., based on the exponential, lognormal, Weibull and inverse Gaussian recurrence-time models, taking into account uncertainties in data and parameter values.

Model	Time Interval			
	1 year	20 years	50 years	100 years
Exponential	0.0049	0.09	0.21	0.38
Lognormal	0.0092	0.17	0.36	0.58
Weibull	0.0104	0.19	0.41	0.64
Inverse Gaussian	0.0074	0.14	0.30	0.51

5.2 Hazard Estimates for South-west Section of Alpine Fault

The results for the south-west section of the fault are summarised in Figure 4 and Table 2. Figure 4 shows the mean hazard function under each of the models. Again the hazard is lowest under the exponential model which indicates that the long-run average rate is about 0.003 events per year, i.e., the mean recurrence time is about 300 years. In this case the present hazard is highest under the lognormal model at about 0.007 events per year, although taken overall there is not a large difference between the mean hazard functions under the lognormal and Weibull models.

Table 2. Estimated probability of rupture of the Alpine fault at points on its south-west section during time intervals starting in the year 2000 A.D., based on the exponential, lognormal, Weibull and inverse Gaussian recurrence-time models, taking into account uncertainties in data and parameter values.

Model	Time Interval			
	1 year	20 years	50 years	100 years
Exponential	0.0032	0.06	0.15	0.27
Lognormal	0.0072	0.14	0.31	0.52
Weibull	0.0064	0.13	0.30	0.52
Inverse Gaussian	0.0052	0.10	0.23	0.40

Table 2 shows estimates of the probabilities of rupture of the fault on the south-west section of the fault for the next year, 20 years, 50 years and 100 years, computed from the hazard curves of Figure 4 using equation (22). These probabilities are lower than the



Hazard on SW section of Alpine fault

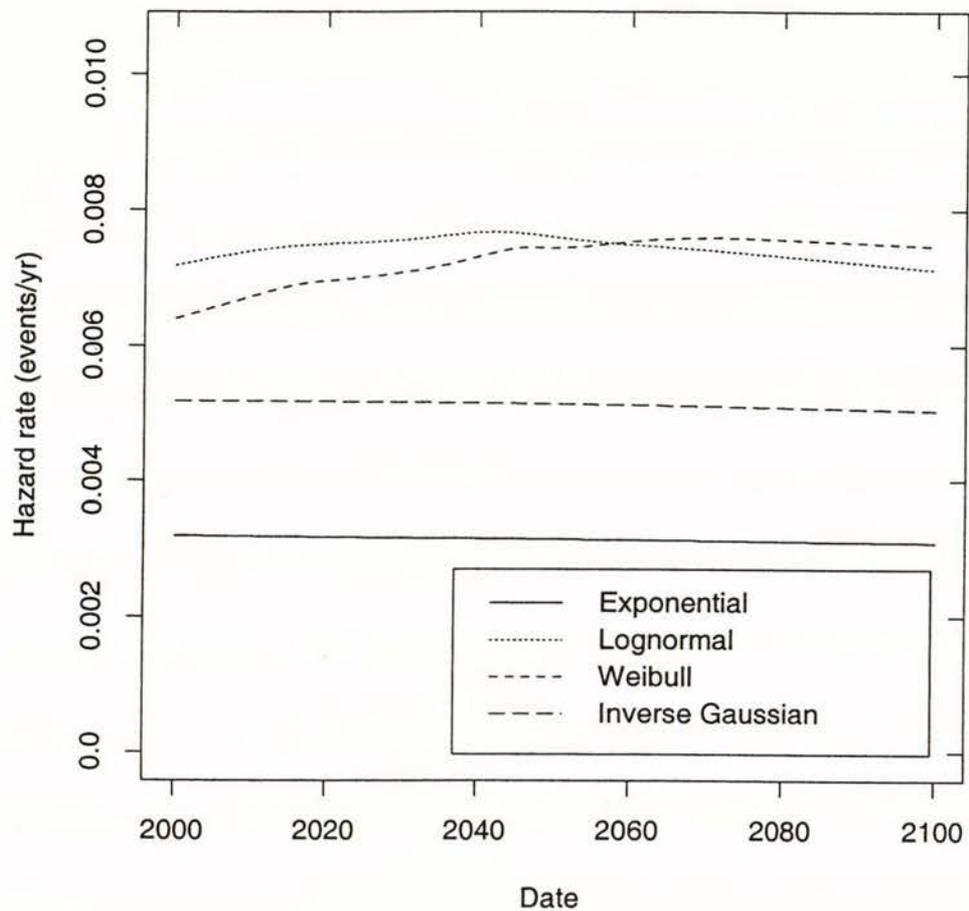


Figure 4. Hazard function averaged over all sampled data values under the exponential, lognormal, Weibull, and inverse Gaussian recurrence-time models applied to the south-west section of the Alpine fault.



corresponding values for the north-east section (Table 1) because of the larger single event displacement, i.e., longer mean recurrence time, estimated for the south-west section.

5.3 Sensitivity of Results to Standard Error of Date of Second to Last Event

The standard deviation of 100 years assumed for date of the second to last event to occur on the south-west section of the fault was not well determined. It is instructive to consider how the results would have been affected if a standard deviation of 50 years had been applied to this date. Accordingly the calculations have been repeated with this data item changed. The results were almost identical to Table 2 for the exponential and inverse Gaussian models but the probabilities were slightly higher in the case of the other two models (see Table 3).

Table 3. Estimated probability of rupture of the Alpine fault at points on its south-west section during time intervals starting in the year 2000 A.D., taking into account uncertainties in data and parameter values, and using 1450 ± 50 yr for the date of the second to last rupture and all other data the same as for Table 2.

Model	Time interval			
	1 year	20 years	50 years	100 years
Exponential	0.0032	0.06	0.15	0.27
Lognormal	0.0092	0.17	0.36	0.58
Weibull	0.0081	0.16	0.36	0.59
Inverse Gaussian	0.0055	0.10	0.24	0.42

It is to be expected that, in general, if the sequence of dated events is short, the effect on the hazard of refining the distribution for the time of occurrence of individual events will be small. Improving the precision of the mean recurrence interval, i.e., refining the distributions of the average slip rate and the average single event displacement is likely to have a greater impact on the hazard.

5.4 Comparison with Previous Estimates

Yetton *et al.* (1988) applied the lognormal recurrence-time model to estimate the hazard on the Alpine fault, and calculated a probability of 0.65 ± 0.15 for a fault rupture occurring during the next 50 years. A 50 year probability of 0.65 corresponds to a mean hazard rate of 0.021 events per year. This is more than double the rate of about 0.009 events per year calculated here for the north-east section of the fault when uncertainties in data and



parameter values are allowed for, and three times the rate of about 0.007 events per for the south-west section of the fault.

6.0 CONCLUSIONS

Of the four models treated here, it is the results for the lognormal model and the inverse Gaussian model that deserve to be taken most seriously. The lognormal model has become widely accepted as a model of fault recurrence. And the inverse Gaussian model, although proposed only recently, has a strong physical basis that lends it some credence. In any case the range of probabilities covered by the four different models is not too wide, despite the very different assumptions underlying them. This range may be a sufficient indicator of the level of hazard for many practical purposes.

The hazard estimated here of future rupture of sections of the Alpine fault is markedly lower than previous estimates under the lognormal model, which did not take account of uncertainty in data estimates and parameter values.

The type of analysis presented in this report is useful for investigating the likely effect on the hazard of improved precision in one or more of the data estimates. It has been shown that the large uncertainty in the date of the second to last rupture on the south-west section of the fault has only a modest effect on the hazard, and then only for two of the recurrence-time models.

A useful outcome of this study is that a computer program is now available to rapidly rework the estimates of hazard in the event that new information on the long-term slip rate, average single event displacements or dates of individual rupture of the Alpine fault should become available. It can also be applied to any other faults for which similar information is available.

7.0 ACKNOWLEDGEMENTS

The authors are grateful to Richard Norris for providing a summary of agreed conclusions on past events and rates on the Alpine fault, and to Mauri McSaveney and Warwick Smith who made internal reviews of the report.

8.0 REFERENCES

Beavan, J., Moore, M., Pearson, C., Henderson, M., Parsons, B., Bourne, S., England, P., Walcott, D., Blick, G., Darby, D., & Hodgkinson, K., 1999. Crustal deformation during 1994-98 due to oblique continental collision in the central Southern Alps, New Zealand,



and implications for seismic potential of the Alpine fault. *Journal of Geophysical Research*, **104**, 25233-25255.

Berryman, K.R., Beanland, S., Cooper, A.F., Cutton, H.N., Norris, R.J., & Wood, P.R., 1992. The Alpine Fault, New Zealand: variation in Quaternary structural style and geomorphic expression. *Annales tectonicae* **VI**, 126-163.

Bull, W.B., 1996. Prehistorical earthquakes on the Alpine Fault, New Zealand. *Journal of Geophysical Research, Solid earth* **101(B3)**, 6037-6050.

Bull, W.B., & Brandon, M.T., 1998. Lichen dating of earthquake-generated regional rockfall events, Southern Alps, New Zealand. *Geological Society of America Bulletin*, **110**, 60-84.

Bull, W.B., King, J., Kong, F., Moutoux, T., & Phillips, W.M., 1994. Lichen dating of coseismic landslide hazards in alpine mountains. *Geomorphology* **10**, 253-264.

Cooper, A.F., & Norris, R.J., 1990. Estimates for the timing of the last coseismic displacement on the Alpine Fault, northern Fiordland, New Zealand. *New Zealand Journal of Geology and Geophysics* **33**, 303-307

Cooper, A.F., & Norris, R.J., 1995. Displacement on the alpine fault at Haast River, south Westland. *New Zealand Journal of Geology and Geophysics* **33**, 303-307

Ellsworth, W.L., Matthews, M.V., Nadeau, R.M., Nishenko, S.P., & Reasenber, P.A., 1999. A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities. Workshop on Earthquake Recurrence: State of the Art and Directions for the Future, Istituto Nazionale de Geofisica, Rome, Italy, 22-25 February, 1999, preprint.

Hagiwara, Y., 1974. Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain. *Tectonophysics* **23**, 321-318.

Hull, A.G., McSaveney, M.J. 1996. A 7000-year record of great earthquakes at Turakirae Head, Wellington, New Zealand. Institute of Geological Sciences client report 33493B.10 to the Earthquake Commission. 14 p.

Johnson, N.L. and Kotz, S., 1970. *Distributions in Statistics: Continuous Univariate Distributions*, 1, 300pp., Houghton Mifflin, Boston Mass.



Nishenko, S., & Buland, R., 1987. A generic recurrence interval distribution for earthquake forecasting. *Bulletin of the Seismological Society of America* **77**, 1382-1399.

Norris, R.J., & Cooper, A.F., in press. Late Quaternary slip rates and slip partitioning on the Alpine Fault, New Zealand. *Structural Geology*.

Stein, R.S., Barka, A.A., & Dieterich, J.H., 1997. Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International* **128**, 594-604.

Rhoades, D.A., Van Dissen, R.J., & Dowrick, D.J., 1994. On the handling of uncertainties in estimating the hazard of rupture on a fault, *Journal of Geophysical Research* **99(B7)**, 13701-13712, 1994.

Sutherland, R., & Norris, R.J., 1995. Late quaternary displacement rate, paleoseismicity, and geomorphic evolution of the Alpine Fault: evidence from Hokuri Creek, South Westland, New Zealand. *New Zealand Journal of Geology and Geophysics* **38**, 419-430.

Wells, A., Stewart, G.H., & Duncan, R.P., 1998. Evidence of widespread synchronous, disturbance-initiated forest establishment in Westland, New Zealand. *Journal of the Royal Society of New Zealand* **28**, 333-345.

Wells, A., Yetton, M.D., Duncan, R.P., and Stewart, G.H., 1999. Prehistoric dates of the most recent Alpine fault earthquakes. *Geology* **27**, 995-998.

Yetton, M.D., 1998. Progress in understanding the paleoseismicity of the central and northern Alpine fault, Westland, New Zealand. *New Zealand Journal of Geology and Geophysics*, **41**, 475-483.

Yetton, M.D., Wells, A., & Traylen, N.J., 1998. The Probability and Consequences of the Next Alpine Fault Earthquake. *EQC Research Report 95/193*.



9.0 APPENDIX

Table A.1. Single event right lateral slip in earthquakes on the North Anatolian fault, over the period 1939-1992. Positions are measured as distances east of 35°E. Data were obtained from figure 1 of Stein *et al.* (1997).

Event (year)	Position (km)	Slip (m)
1967	-400	0.4
1967	-375	1.0
1967	-350	1.5
1957	-325	1.5
1944	-300	2.0
1944	-275	3.0
1944	-250	3.0
1944	-225	2.5
1944	-200	1.9
1944	-175	1.4
1944	-150	1.2
1943	-125	2.5
1943	-100	4.5
1943	-75	3.8
1943	-50	3.0
1943	-25	2.8
1943	0	2.6
1943	25	2.4
1943	50	2.2
1943	75	1.5
1943	100	2.6
1943	125	2.1
1939	150	2.5
1939	175	2.8
1939	200	3.5
1939	225	3.8
1939	250	4.5
1939	275	4.0
1939	300	6.5
1939	325	5.8
1939	350	5.7
1939	375	6.0
1939	400	6.3
1939	425	0.5
1949	525	1.5
1949	550	1.5
1966	600	0.5
1951	-225	0.7
1951	-250	0.3
1939	100	1.3
1939	125	1.8
1942	175	1.8
1992	425	1.0



NON-TECHNICAL ABSTRACT

The probability of a major rupture on the Alpine fault is considered to be increasing gradually since the last event in about 1717 AD. This increase has been estimated by using statistical models of time-varying hazard, and a procedure which takes account of uncertainties in estimates of the long-term slip rate, the average size of individual ruptures and the dates of the last few rupture events. The probability of rupture over the next 20 years is estimated to be up to 20% on the north-east section of the fault and up to 15% on the south-west section, depending on the particular model used. These figures represent about twice the long-term average probability.



TECHNICAL ABSTRACT

The time-varying hazard of rupture on the Alpine fault is estimated using a technique that takes account of uncertainties in data and parameter values. The north-east (Karangarua-Haupiri) and south-west (Haast) sections of the fault are considered separately, i.e., it is not assumed that they will necessarily rupture together. Data inputs are based on geological estimates of the long-term slip rate and previous studies of fault traces and forest ages and times of disturbance. The geological average strike-slip rate is taken to be 26 ± 5 mm/yr and the average single-event displacements to be 5.0 ± 1.4 m and 8.0 ± 2.6 m on the north-east and south-west sections, respectively. The last four events on the north-west section are dated at 1717, 1620 ± 10 , 1445 ± 20 and 1150 ± 50 , and the last three events on the south-west section at 1717, 1450 ± 100 and 1150 ± 50 A.D. Using these data and associated uncertainties, the current hazard of rupture on the north-east section of the fault is estimated to be 0.0049, 0.0092, 0.0104 and 0.0074 events per year under the exponential, lognormal, Weibull and inverse Gaussian recurrence-time models, respectively. The corresponding probabilities of rupture in the next 20 years are 9%, 17%, 19% and 14%, respectively. The current hazard on the south-west section of the fault is estimated to be 0.0032, 0.0072, 0.0064 and 0.0052 events per year for the four models, and the 20 year probabilities 6%, 14%, 13%, and 10%, respectively. Increased precision in the date of the second to last event on the south-west section of the fault would result in only small changes to these rates and probabilities. The hazard under the lognormal model is about double the long-term average rate but less than half of that previously estimated without taking account of uncertainties in the data and parameter values.

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LIMITED

Gracefield Research Centre
69 Gracefield Road
PO Box 30368
Lower Hutt
New Zealand
Phone +64-4-570 1444
Fax +64-4-570 4600

Wairakei Research Centre
State Highway 1, Wairakei
Private Bag 2000
Taupo
New Zealand
Phone +64-7-374 8211
Fax +64-7-374 8199

Rafter Laboratory
30 Gracefield Road
PO Box 31312
Lower Hutt
New Zealand
Phone +64-4-570 4637
Fax +64-4-570 4657

Dunedin Research Centre
764 Cumberland Street
Private Bag 1930
Dunedin
New Zealand
Phone +64-3-477 4050
Fax +64-3-477 5232