attention, some work supports even greater importance for the larger, but temporary, dynamic stress changes caused by the seismic wavefield (Gomberg et al., 2001; Parsons, 2002). Both static and dynamic stress transfer are likely to have an effect on future seismicity and we need to determine which is the most important or if they need to be combined into a single model for accurate applications.

5.1.3 Conclusions

Improved, and especially time-dependent, earthquake probabilities will require new models of earthquake behavior. As we develop a better understanding of the basic processes involved we will need to create physically-based statistical distributions of earthquake recurrence on a regional network of faults. This will be far more complex than the simple point processes currently used. Succeeding at this task will require not only a better understanding of the basic processes but also the ability to produce a variety of numerical models of earthquake behavior on a regional scale and the data needed to validate or reject these models. These goals are just starting to become realistic as computational power increases and new data is collected.

5.2 Probabilistic forecasting of seismicity

[By Y. Y. Kagan, Y. F. Rong, and D. D. Jackson]

Predicting individual earthquakes is not possible now, but long-term probabilistic forecasts can be validated and provide useful information for managing earthquake risk. Short-term probability estimates are important for emergency and scientific response, but they are considerably more difficult to construct and test than long-term forecasts. Here we present three different forecast models, along with some quantitative tests of their effectiveness. The three models are for long-term and short-term forecasts based on seismicity, and for long-term forecasts based on geodetically observed strain rate.

Our long-term seismicity model is described in Kagan and Jackson (1994) and Jackson and Kagan (2000). We made specific forecasts of earthquakes over magnitude 5.8 for two regions of the Pacific Rim at the start of year 1999, and we present here a true prospective test of those forecasts. The short-term model is much harder to test formally, but we show here some reasons to be optimistic. We also present a 'pseudo-prospective' test (that is, using data collected after 1993 to test a model based on pre-1993 data) of the geodetic strain model for southern California.
We assume that there are two processes causing earthquakes: a steady state process for which earthquake occurrence is independent of past events, and a clustering process in which each earthquake depends on previous 'parent' events (Vere-Jones, 1970; Kagan, 1991; Ogata, 1998; Jackson and Kagan, 1999). Seismicity is approximated by a Poisson cluster process, in which clusters or sequences of earthquakes are statistically independent with constant rate although individual earthquakes in the cluster are dependent events. We refer to earthquakes with no obvious parents as 'independent' events, and their progeny as 'dependent' events. Dependent events need not to be smaller than their parents. In practice, we need not to identify specific earthquakes as independent or dependent, but the distinction is useful for modeling the separate functions that describe their rates.

5.2.1 Long-term seismic hazard estimates

We have developed a long-term model for independent events based on smoothed seismicity. The method assumes that the rate density (the probability per unit area, magnitude, and time) at a given location is proportional to the magnitude of nearby past earthquakes and approximately proportional to a negative power of the epicentral distance out to a few hundred kilometers. The method is described in more detail in Kagan and Jackson (1994, 2000), Jackson and Kagan (1999, 2000), and Jackson et al. (2001). We assume that the rate density does not depend on time, although our estimate of it changes as the earthquake catalogue evolves. Forecasts based on the model, and tests of the model, can be seen at http://scsees.ess.ucla.edu/ykagan.html. An example of the rate density for earthquakes larger than moment-magnitude $m_w = 5.8$ for the Northwest Pacific region is shown in figure 5.1.

In all of our calculations, we describe earthquakes, and their probability densities, in terms of seismic moment, which is the directly observed quantity in the Harvard earthquake catalogue (Deichmann et al., 2001) we used. However, many readers will be more familiar with a description of earthquake size in terms of moment-magnitude, so we will often refer to magnitude in the discussion below. We use the notation $M$ for the scalar seismic moment and $m$ for the moment-magnitude, which we calculate using

$$m = \frac{2}{3} \log_{10} M - 6.0, \quad (5.1)$$

where $M$ is measured in Nm. Readers should be aware that not all published definitions of moment-magnitude use the same constant 6.0 in equation (5.1). Some authors use values different by as much as 0.1. Because we base our calculations on seismic moment only, our results are self consistent, but the magnitudes for
5.2 Probabilistic forecasting of seismicity

individual earthquakes, and the earthquake rates we report for a given magnitude, might differ slightly from those that would be obtained using a different constant.

We use a statistical model to fit the catalogue of earthquake times, locations, and seismic moments, and we base forecasts on this model. While most model components have been validated by extensive geophysical research, some require further investigation. We assume that the occurrence rate \( \phi(x, M) \) of independent events at location \( x \) with magnitude \( M \) may be written purely in terms of the marginal rates, i.e.,

\[
\phi(x, M) \propto \phi_x(x) \times \phi_M(M),
\]

(5.2)

where \( x \) are horizontal spatial coordinates, and \( \phi_x(x) \) and \( \phi_M(M) \) are normalized probability densities of events in area and magnitude, respectively. Equation (5.2)

**Figure 5.1:** Color tones show the probability of earthquake occurrence calculated using the Harvard 1977-2000 catalogue; earthquakes 2001 are shown in white. Northwest Pacific long-term seismicity forecast: latitude limits from 0.25°S to 60.25°N, longitude limits from 109.75°E to 170.25°E.
signifies that the spatial and moment distributions of earthquake clusters are independent. The distribution \( \phi_M(M) \) is defined by equation (5.3) below; a preliminary methodology for its use is described by Kagan and Jackson (1994), Kagan and Jackson (2000) and Jackson and Kagan (1999, 2000). The seismic moments are modeled as following the tapered Gutenberg-Richter distribution (Kagan and Jackson, 2000; Kagan and Schönberg, 2001; Kagan, 2002a):

\[
F_M(M) = \Phi_M(M) = M^{-\beta} \exp(-M/M_c) \quad \text{for} \quad M_t \leq M \leq \infty,
\]

where \( F_M(M) \) is a cumulative distribution function, \( M_t \) is a catalogue completeness threshold (cutoff), taken here to be \( M_t = 10^{17.7} \) Nm \((m_t = 5.8)\); and \( M_c \) is the parameter that controls the distribution in the upper ranges of \( M \) ("corner moment"). We assume that \( \beta \) and \( M_c \) are uniform over the entire area of study, and we estimate them by choosing the values that give the maximum likelihood fit to the earthquake catalogue for the entire region.

This moment distribution is equivalent to a normalized tapered Gutenberg-Richter magnitude distribution, defined by a "\( b \)-value" and a "corner magnitude". The "\( b \)-value" is 1.5 times the value \( \beta \) in the equation above, and the corner magnitude is the value obtained by converting \( M_c \) to magnitude using equation (5.1). The "\( a \)-value" does not appear in equation (5.3), because the moment distribution is normalized, but the function \( \phi(x, M) \) plays the role of a locally variable "\( a \)-value".

The rate density function of equation (5.2) is factored into the product of functions of location and magnitude only, reflecting our assumption that the relative moment distribution is independent of location. Only the "\( a \)-value" depends on location in this model. Thus the map of Figure 5.1 represents the rate density for any earthquake size, after scaling the mapped values by a constant depending on the moment distribution.

The assumption of a uniform \( b \)-value and corner magnitude is somewhat controversial. Many seismologists have proposed that these parameters should depend on tectonic environment, geometry of faults, or other factors. However, our studies to date show that the ratio of small earthquakes to the overall tectonic moment rate is remarkably consistent across different tectonic environments in continental interiors and subduction zones (Kagan, 2002a, b). Because this ratio depends strongly on the \( b \)-value and corner magnitude, we propose that these parameters are relatively uniform. By formalizing this hypothesis quantitatively, as we have done here, we make it possible to test contrasting hypotheses against it using future large earthquakes. The alternative hypotheses for such a test must also be stated in terms of earthquake rate density.

We have begun a "prospective" test of our model and earthquakes since the beginning of 1999 (Kagan and Jackson, 2000) are quite compatible with
5.2 Probabilistic forecasting of seismicity

![Graph showing cumulative histograms of the likelihood scores of synthetic catalogues relative to the observed ones for calendar year 2001. Figures for calendar years 1999 and 2000 were very similar.](http://scesess.ucla.edu/~ykagan/tests_index.html)

Figure 5.2: We use earthquakes from 1977 to 2000 as a control set. The solid line is the best Gaussian curve, having the same standard deviation as simulations. The dashed curve corresponds to simulation distributions for the NW-Pacific; the dotted curve to the SW-Pacific. Curves on the right from the Gaussian curve correspond to simulations worse than a real earthquake distribution; curves on the left correspond to simulations better than a real earthquake distribution.

...To test the model we generated and scored a million synthetic earthquake catalogues each having the same number of events as the observed earthquake catalogue, and a spatial distribution consistent with the assumed $\phi_X(x)$. The 'score' for each catalogue is the log-likelihood, computed by summing the log of the theoretical rate density $\phi_X(x)$ at the location of each quake. We then computed the score for the actual catalogue, and compared it with the simulated values. Figure 5.2 shows cumulative histograms of the likelihood scores of synthetic catalogues relative to the observed ones for calendar year 2001. Figures for calendar years 1999 and 2000 were very similar (http://scesess.ucla.edu/~ykagan/tests_index.html). In most areas the observed likelihood score was within the central 95% confidence limits based on the simulations. The exceptions were the SW Pacific in 1999, and the NW Pacific in 2001 (the latter is shown in figure 5.2). In both of the exceptional cases, the actual catalogue had a higher likelihood score than 97.5% per cent of the simulations.

When we defined spatial smoothing kernels $\phi_X(x)$, we optimized them to forecast the latter half of a catalogue using the first catalogue half as a training set (Kagan and Jackson, 1994, 2000). Clearly, the kernels set to forecast 10–12 years...
in advance may not perform as well in yearly forecasts. More research is needed to develop appropriate spatial kernels.

We judge the model to represent the catalogue well when the likelihood of the observed catalogue falls in the middle range of synthetic values. When most of the synthetic catalogues plot to the right of the observed one, then the observed earthquakes are occurring in lower probability regions than expected, and the model fails. When most of the synthetic catalogues plot to left of the observed, then the observed earthquakes are falling in the high probability regions even more than expected, and the model fails by over-smoothing the earthquake probabilities.

The relationship between the forecast and observed earthquake locations can be more easily seen in displays of probability and earthquake location concentration, shown in figure 5.3a, b. To make these diagrams we divide the region into small cells; estimate the theoretical rate of earthquakes above the magnitude threshold for each cell, using equation (5.2); count the events that actually occurred in each cell; sort the cells in decreasing order of theoretical rate; and compute cumulative values of forecast and observed earthquake rates as plotted in the figures. We use two criteria to evaluate the forecasts using these figures. First, the most useful forecasts will have most of the probability concentrated in a small area; for such a forecast, the cumulative forecast rate will rise very steeply on the left side of the diagram and flatten out to the right. Second, for a successful forecast the observed earthquake distribution should mimic the theoretical forecast distribution. In fact the likelihood score for the actual catalogue can be computed directly from the table used to produce the concentrations diagrams like figure 5.3a, b.

In figure 5.3a (NW Pacific) all earthquakes are concentrated in the 'hottest' 30 per cent of the area, whereas the forecast predicts only about 93 per cent of epicenters to be in this part of the Pacific. The forecast for the NW Pacific over-smoothed the seismicity, and a more concentrated kernel function \( \phi(x) \) would have provided a better fit. This explains why the model fails for having too high a likelihood, as shown in figure 5.2. Earthquakes in the SW Pacific were concentrated more nearly as forecasted, as shown in figure 5.3b. This explains why the forecast is within the 95 per cent confidence limits as shown in figure 5.2.

### 5.2.2 Short-term seismic hazard estimates

We are exploiting short-term trends in earthquake catalogues to provide daily estimates of the probabilities of strong earthquakes throughout the globe, and especially on the Pacific Rim. Short-term hazards may be viewed as temporary perturbations to the long-term earthquake potential. By short-term we mean earthquake hazard estimates on the order of a few days, weeks or even months. Our preliminary investigations indicate that following most mod-
5.2 Probabilistic forecasting of seismicity

![Smoothed Seismicity: NW Pacific]

![Smoothed Seismicity: SW Pacific]

Figure 5.3: Concentration diagrams for forecast probability and earthquakes occurring in 2001. Solid line – probability distribution, thin line – earthquakes. (top) NW Pacific; (bottom) SW Pacific.

Catastrophic earthquakes (magnitude 6–7), seismic activity is dramatically increased for a few weeks, although the probability is measurably higher for a few years. Our forecasts are based on the most recent seismic record. They
are and will be updated continuously and published on the World-Wide Web (http://ssec.ess.ucla.edu/~ykagan/predictions_index.html).

Our short-term forecasts are largely based on earthquake clustering, the best-known manifestations of which are foreshock-mainshock-aftershock sequences. Kagan and Knopoff (1987), Kagan (1991), Reasenberg and Jones (1989), Utsu and Ogata (1977), Ogata (1988, 1998), Console (1998), Michael and Jones (1998), Wiemer (2000), and Ebel et al. (2000) described quantitative models of clustering. Although short-term clustering has been recognized for some time, its application to real time forward forecasting was not feasible until recently. The present availability of earthquake CMT solutions within a few hours after an event, the ability of fast computers to generate earthquake hazard maps in a few minutes, and the capability of the World-Wide Web for instant distribution of results to a wide audience make short-term forecasting worthwhile.

Short-term forecasts may have important practical benefits. Early alerts to probable strong earthquakes will allow emergency response personnel to anticipate and begin planning for some possibly disastrous earthquakes. Our preliminary analysis indicates that on average about 9 per cent of earthquakes are followed by a stronger dependent event within a relatively short time (a few days or weeks). This fact explains why statistical short-term prediction does not meet the popular definition of ‘prediction’ (Kagan, 1997) as only a few alarms would be followed closely by earthquakes large enough to cause damage. However, investigations (Kagan and Knopoff, 1987; Reasenberg and Jones, 1989; Michael and Jones, 1998; Reasenberg, 1999) show that from 20 per cent to 40 per cent of mainshocks are preceded by identifiable foreshocks. This means that some inexpensive mitigation measures (Molchan and Kagan, 1992) can be used. Examples of such measures include putting emergency services on a higher alert status, stepping up InSAR and other geophysical measurements in the dangerous regions, and deploying temporary seismic stations.

5.2.3 Experimental short-term forecasts for Western Pacific

As an exploratory step we started to calculate short-term hazard estimates for the western Pacific region. The forecasts are computed from the preliminary Harvard CMT catalogue (Dziewonski et al., 2001). We use email messages sent by the Harvard team to update our catalogue for all earthquakes $M \geq 10^{17.7}$ Nm ($m \geq 5.8$); the time delay between earthquake occurrence and update is on the order of a few hours, or sometimes one day. This catalogue is then used to estimate both long- and short-term daily probabilities of future earthquake occurrence. The short-term values are for one-day periods from the current midnight to the subsequent midnight, Los Angeles time. These forecasts are stored in a table and displayed in two figures, which are accessible via the World Wide Web.
5.2 Probabilistic forecasting of seismicity

Figure 5.4: Short-term seismicity forecast for southwest Pacific. Color tones show the probability of earthquake occurrence calculated using the Harvard catalogue starting with January 1, 1977. This forecast as well as the forecast for the northwest Pacific is updated every day. The updated plots are available from http://scec.ess.ucla.edu/~ykagan/predictions/index.html.

(http://scec.ess.ucla.edu/~ykagan.html, see "FORECASTS FOR 1999-2002: TABLES AND FIGURES"). An example of such a forecast for the southwest Pacific region is shown in figure 5.4. The red glow in the New Ireland region shows the influence, as of February 13, 2002, caused by the sequence of strong earthquakes starting on November 16, 2000. After 15 months the earthquake occurrence rate is still much higher than the background rate, the recent sequence of four moderate events near Solomon islands (April 19, 2001, latitude 7.5° S, longitude 136° E) is also highlighted by its red color.

Each cluster is closely approximated by a stochastic space-time critical branching process. The space-time distribution of interrelated earthquake sources within a sequence is controlled by simple relations justified by analyzing the available statistical data on seismicity. Usually the first event in a sequence is the largest one and it is called a mainshock. Other dependent events are called aftershocks. If the first event in a sequence is smaller than subsequent shocks, it is called a foreshock. Retrospectively, it is relatively easy to subdivide an earth-
quake catalogue into foreshocks, mainshocks, and aftershocks. However, in real time any event might be a foreshock. Although it is likely that dependent events will be smaller and called aftershocks, there is a significant chance that such earthquakes may be bigger than the events that preceded them (Kagan, 1991; Michael and Jones, 1998; Reasenberg, 1999).

We assume that the distribution of the dependent events within a cluster may also be broken down into the product of its marginal distributions. That is, the conditional probability density of the j-th shock dependent on the i-th mainshock \((j > i)\) with seismic moment \(M_i\) is modeled as

\[
\psi(\tau, \rho, M_j|M_i) = \psi_{\tau}(\tau) \times \psi_{\rho}(\rho) \times \psi_{M_i}(M_i) \times \phi_M(M_j),
\]

where \(\tau = t_j - t_i\) and \(\rho\) is the horizontal distance between the i-th and j-th centroids. The functions \(\psi_{\tau}, \psi_{\rho},\) and \(\psi_{M_i}\) are the marginal temporal, spatial, and moment densities, and are detailed in our publications (Jackson and Kagan, 1999; Kagan and Jackson, 2000; Bird et al., 2000; Kagan, 2002a, b).

Quantitatively, the marginal densities behave as follows (Kagan, 1991; Kagan and Jackson, 2000). The function \(\psi_{\tau}(\tau)\) describes a power-law decrease in earthquake rate vs. time, which is like the Omori law. We make an adjustment near \(\tau = 0\) to prevent singularity of \(\psi_{\tau}\). The exponent in \(\psi_{\tau}(\tau)\), estimated by the maximum likelihood (Kagan, 1991), differs substantially from the Omori exponent (section 2.3.2) because the earthquake rate density behaves on all previous events, not just on one previous mainshock. The spatial function \(\psi_{\rho}(\rho)\) depends approximately as a Gaussian function of epicentral distance \(\rho\) (Kagan, 1991; Kagan and Jackson, 2000). The moment distribution \(\psi_{M_i}(M_i)\) increases as a power law with the moment of previous events; \(\phi_M(M_j)\) decreases as a negative power of the moment of later events. Thus large events increase the probability more than smaller events, and at any time large events are less probable than small ones. The density \(\phi_M\) is proportional to the derivative of \(\Phi_M\) in equation (5.3).

The earthquake rate density at a location is the sum of the values calculated from equation (5.4) for all past earthquakes. The individual terms from equation (5.4) do not depend on whether the past earthquakes were foreshocks, mainshocks, or aftershocks, so this distinction becomes irrelevant for estimating earthquake potential. However, some other methods use a formula like equation (5.4), but the enhanced earthquake rate is presumed to stem from just one main event, so there is only one term in the sum. In such a case the choice of which event is called the mainshock could affect very strongly the estimated earthquake rate.

Although the focal mechanisms of future earthquakes might depend on the focal mechanism of previous events in the cluster (Kagan and Jackson, 1994), we presently assume that the distribution of focal mechanisms within a cluster is the same as the distribution governing long-term focal mechanism. Within an
5.2 Probabilistic forecasting of seismicity

Figure 5.5: Time-magnitude plot for the New Ireland region.

earthquake sequence the focal mechanisms appear to be very tightly clustered, with the Cauchy rotational distribution describing the scatter of the 3D angle of rotation (Kagan, 2000).

Figures 5.5–5.7 show how the method would apply to the New Ireland sequence of large earthquakes which started on November 16, 2000. In figure 5.5 the time-magnitude sequence of events is shown. It is clear that the first three large earthquakes, which occurred in rapid succession, have very similar moment-magnitudes (8.06, 7.87, 7.83 in the final Harvard catalogue). The magnitude differences are comparable with the magnitude error in the catalogue (Kagan, 2002a). Their magnitudes in a preliminary catalogue are 8.03, 7.61, and 7.57, respectively. The mainshock would be the same in either the preliminary or final catalogue, but it is clear that in some cases the identity of the mainshock (i.e. the largest shock in a sequence) might change between the preliminary and final catalogue. Thus our method of summing over all events to estimate rate density has a strong advantage over methods based on using only a single event.

The closeness of magnitude values and their change from a preliminary to the final catalogue would present a serious difficulty for any forecasting procedure based on the Omori law and empirical rules (Wiemer, 2001). Such procedures depend on a real-time identification of foreshocks, mainshocks, and aftershocks. Probabilities are functions of time after the mainshock, so results are very sensitive to the choice of which event is called the mainshock. No such problem arises
for our method. Since the model is based on a stochastic process formulation, no earthquake identification is necessary and minor adjustments in earthquake magnitude would only cause minor changes in the forecasted earthquake occurrence rate.

We constructed a hybrid earthquake forecast by taking weighted average of the long- and short-term forecasts (Jackson and Kagan, 1999; Kagan and Jackson, 2000). The total rate-density in a hybrid model is the short-term component plus 80 per cent of the long-term rate-density. In figure 5.6 we display the time history of the hybrid model, with the long-term model for comparison, at one point near the center of the November 16, 2000 sequence. The reference point is 5.0° S, 152.5° E, i.e., near the centroid of the sequence (see figure 5.7). Figure 5.6 shows that the estimated short-term rates increase by a few orders of magnitude for a very short time following each nearby event. Each of the three large earthquakes on November 16 and 17 spiked the theoretical earthquake rates, although the first two events were close enough in time that their effects merge in figure 5.6. The November 16 and 17 sequence and its aftershocks kept the short-term rate more than five times the long-term rate for a period of about a month and a half.

Figure 5.7 displays the spatial distribution of earthquakes in the New

![Figure 5.6: Time history of long-term and hybrid (short-term plus 0.8 × long-term) forecast for a point at 5.0° N, 152.5° E. Dark line is the long-term forecast, lighter line is the hybrid forecast.](image)
Ireland region during 2000-2001. Earthquake positions are overlayed on the long-term forecast of seismic activity, which is based on the 1977–1999 earthquake history (Jackson and Kagan, 1999; Kagan and Jackson, 2000; Jackson et al., 2001). Obviously, most of the earthquakes occurred in zones of a high forecasted activity level. A more quantitative measure of the forecasting efficiency (see FORECAST TEST FOR 2001: in http://sec.ess.ucla.edu/~ykagan/predictions_index.html) shows that our long-term forecast is quite satisfactory – the real catalogue falls within the 95 per cent confidence intervals of model prediction.

5.2.4 Experimental forecasts in Southern California

We also applied the technique described above to make a long-term seismicity forecast for southern California, although we did not put it on the web. The seismicity rate in California is too low for estimation of the smoothing kernel.

![New Ireland Forecast 1977-99, Eqs 2000/1/1-2001/05/02 (r = 2.5 km, λ = 1.9, δ = 25)](image)

**Figure 5.7:** New Ireland region long-term seismicity forecast: color shows the rate-density of earthquake occurrence calculated using the Harvard 1977-1999 catalogue. Latitude limits are 3° to 8° South; longitude limits 150° to 157° West. Earthquakes in 2000 and 2001 (after the time of the forecast) are shown as white circles, with radius proportional to magnitude.
parameters by using one part of the catalogue to forecast a later part. Thus, we
adopted the values of the kernel parameters from our analysis of global seismicity.
However, in California we can extend our description of the past seismicity by
representing earthquakes with magnitude 6.5 and greater as extended sources: i.e.,
as a sum of rectangular dislocation sources, and treating the separate dislocations
as separate earthquakes. An example of such a representation of seismic sources
is shown in figure 2 of Kagan (1994).

We calculate probabilities for earthquakes with \( m > 5 \) per unit area and time
on a 5 km grid. For larger earthquakes the probabilities are lower according to
the magnitude distribution, which we assume to be a tapered Gutenberg-Richter
distribution with uniform \( b \)-value of 0.95 and a uniform corner magnitude of 8.5.
The forecast model has just a few adjustable parameters: a normalizing constant,
a smoothing distance, and an anisotropy factor. For California earthquakes we
can almost always determine, using fault geology data, which of two planes in a
focal mechanism solution is the fault plane. In a global catalogue such a decision
is more difficult and is based on a statistical guess (Kagan and Jackson, 1994;
Kagan and Jackson, 2000) which is correct only in about 75 per cent cases.

We have used the model in a 'pseudo-prospective' forecast in southern Cali-
ifornia (figure 5.8). We defined the probabilities using earthquakes before the
beginning of 1993 and tested against later earthquakes, obtaining diagrams similar
to figure 5.3a, b. The smoothed seismicity model predicts that 90 per cent of
the earthquakes should lie in the 'hottest' 58 per cent of the area covered, while
in fact all events after 1993 lay in the hottest 41 per cent. Thus, as in figure 5.3a,
we smoothed too much. Optimizing the parameters with southern California data
will probably result in a 'sharper' forecast. We can also use the model to construct
random synthetic earthquake catalogues, including focal mechanisms. Since the
likelihood score, a measure of compatibility of data and theory, was about the
same for the real catalogue as it was for the simulated ones, the forecast was quite
consistent with observed earthquakes.

We constructed an Earthquake Potential Model (figure 5.9) based on maxi-
mum horizontal shear strain rate, evaluated in 1993, based on the Southern Cali-
ifornia Earthquake Center (SCEC) Crustal Motion Model 2.0 (Shen et al., 1996;
Jackson et al., 1997). We chose 1993 as the 'pseudo-prospective' test date be-
cause it is the earliest date for which we could accurately estimate the strain rate.
Figure 5.9 also shows earthquakes after 1993. This model has a normalization
constant and a smoothing parameter.

We set the normalization constant to fit the long-term earthquake rate, and we
chose the smoothing to fit the observed GPS velocities. We assumed the same
magnitude distribution as for the smoothed seismicity model. From a diagram
similar to figure 5.3a, b, we infer that the geodetic model forecasted 90 per cent of
the earthquakes in the fastest deforming 67 per cent of the area. In fact, all events
5.2 Probabilistic forecasting of seismicity

Figure 5.8: Long-term seismicity forecast for southern California: Latitude limits 32.0–37.0°N, longitude limits 114.0–122.0°W. Earthquakes after December 31, 1992 are shown in white. The size of circles is proportional to earthquake magnitude. Color scale shows the long-term probability of earthquake occurrence calculated with the historical and Harvard 1977–1992 catalogue.

After 1993 occurred in this area. In our calculations we excluded from consideration the parts of the southern California where strain data are not available (NE and SW parts of figure 5.9), thus probability values in our forecast are more uniform. The simulated earthquake catalogues looked very much like the observed one.

5.2.5 Conclusions

The best indicator of earthquake probability appears to be the occurrence of recent earthquakes. We regard the seismicity models as ‘null hypotheses’. In other words, they are necessary first steps required to test more sophisticated models based on earthquake physics, active faults, or tectonic stresses. The long-term seismicity model assumes that earthquake probability does not vary with time,
while the short-term model includes an explicit representation of earthquake clustering.

In general the long-term seismicity model performs well, although in two cases the model failed our two-sided likelihood test because the likelihood was too high. This occurred because the observed earthquakes were concentrated more in the higher probability regions than the model had predicted. Short-term probabilities seldom reach the high levels that would justify the label ‘earthquake prediction’, but they do vary enough to justify preliminary emergency activities in some circumstances. Aftershocks exemplify earthquakes made more probable by previous events, but earthquake clustering goes well beyond simple aftershock sequences. Many large earthquakes are preceded by smaller events, and a quantitative clustering model can forecast those larger events reasonably well.

![S. California strain (Earthquakes 1963-2001, C/I T m>5.0)](image)

**Figure 5.9:** Same as figure 5.8 except that color scale tones show the long-term probability of earthquake occurrence calculated using strain map.