Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tectonophysics 413 (2006) 5–12
TECTONOPHYSICS
www.elsevier.com/locate/tecto

The development of statistical seismology: A personal experience

David Vere-Jones

Victoria University and Statistical Research Associates, Ltd, Wellington, New Zealand

Received 3 January 2005; received in revised form 4 April 2005; accepted 4 October 2005
Available online 2 December 2005

Abstract

An account is given of the author's experiences in the development of statistical seismology as a tool for the description, understanding, and forecasting of earthquakes. The discussion is focussed on the formulation and use of stochastic models for earthquake occurrence.

Keywords: Statistical seismology; Earthquake prediction; Probability forecasts; Seismicity models; History of seismology

1. Introduction

In the broader picture, statistics and geophysics have made major contributions to each other. From Halley through Schuster to Jeffreys and beyond, developments in the two subjects have frequently gone hand in hand, impelled by the demands of a science where data gathering and processing is a central task. Seismology in particular has presented statistics with an array of challenging and fruitful problems: epicentre location and velocity modelling; the interpretation of seismograms and a host of related problems in time-series analysis; the analysis of building response and its uses in earthquake-resistant design and the assessment of risk from seismic and other geophysical events. It is almost a quarter of a century since Euan Smith and I attempted to list the major applications of statistics in seismology (Vere-Jones and Smith, 1981), and the range has not exactly diminished since that time.

But for this paper I want to focus on a much narrower issue: the role of statistical (or stochastic) modelling in the analysis of earthquake catalogue data. Can such modelling help in the isolation of features of geophysical interest, in the understanding earthquake mechanism, or in the development of earthquake prediction schemes? My aim is to examine these questions, while at the same time describing how the theme has evolved, from my own particular experience at its statistical end.

2. At the beginning: descriptive models

As far as I am aware, the phrase 'Statistical Seismology' was first used by Keiiti Aki as the title of one of his very early papers (Aki, 1956). My own introduction to the subject came a few years later, at the beginning of 1962, when, as a newly qualified PhD, I started work in the Applied Mathematics Division of the New Zealand Department of Scientific and Industrial Research (DSIR). After a few days of laissez faire introduction, I felt it necessary to seek advice as to what I should do. Several alternatives were suggested, but the one I took up was completing a study started by an earlier recent graduate, Steve Turnovsky, on the statistics of New Zealand earthquakes. This involved ascertaining a period and region over which the catalogue was reasonably complete, attempting to differentiate between main

E-mail address: dvj@mcs.vuw.ac.nz.

0040-1951/$ - see front matter
doi:10.1016/j.tecto.2005.10.005

events and aftershocks, and checking for trends and other features of potential interest. It was a good grounding in work on catalogues, but I hardly anticipated that the topic would still hold me captive 40 years later, or that it would seem more challenging at the end of those 40 years than it did at the beginning.

The New Zealand seismologists with whom I started to work made up a colourful and individualistic group, a fact which undoubtedly contributed to my enthusiasm for the subject. Frank Evison had recently been put in charge of the Geophysics Division of DSIR, Robin Adams headed the Seismological Observatory, and George Eiby was his right hand man. I first collaborated with George, whose booming voice, talents in painting and writing, and enthusiasm for seismology, seemed altogether too large for life in the modest wooden building, located in Wellington's Botanical Gardens, where the Observatory was then housed (it had originally formed part of the Meteorological section, and was still located alongside it). During the war years, when most of the other staff had been seconded for military purposes, George had been forced to remain in New Zealand on health grounds, and had been a key figure in keeping the seismic network running. He knew exactly which stations had been out of action in which periods, where and when the data could be considered reliable, and much of the earlier history of seismology. Besides, his little book (Eiby, 1967) remains an excellent informal introduction to the subject.

Our first studies of the catalogue were unashamedly descriptive, involving histograms, averages and statistics, but this is not necessarily a criticism. If data comes first, as the statistician's creed suggests, then the first job must be to describe it. Finding the right way to display the data, so that the feature of interest is clearly isolated, is often tantamount to solving the problem outright. Moreover plotting the data in various ways is a crucial precaution against drawing flawed conclusions from flawed data. But as the data gets more complex, and the questions asked more subtle, much insight can be obtained by developing a simple statistical model which at least describes, even if it does not explain, the major features of the data.

Formulating even the simplest statistical model means writing down a set of assumptions which would allow one to simulate (in the Monte Carlo sense) a data set sharing the major features of the data to hand. Simulation is a great stimulus in addressing the problem to which Jeffreys (1973) drew particular attention: a satisfactory model must quantify not only the features directly explained by physical laws, but also the unexplained variation. The attempt to develop such a model helps one roughly to place the data in the statistical zoo, that is, to identify its basic distributional features, to locate its neighbours, and hence to refine one's impression of the type of process by which the data might have been generated. Normal distributions, for example, arise typically from the summation of many small contributions, while the Poisson process frequently arises from combining series of events which originate from independent sources. Such hints immediately suggest that the Poisson process is likely to be a useful model for describing infrequent events over large time or distance scales. Indeed, it is this feature which keeps the Poisson process in play as the archetypal model for series of large events in engineering and insurance applications.

Examining a regional catalogue in detail, however, immediately reveals that the earthquake process is extremely complex. Even for events with $M_L \geq 4.5$, which were all that could be consistently recorded at that time, clustering is a major feature. The events have depths and magnitudes as well as epicentral locations and origin times, not to mention more sophisticated features (hardly available in those early days) such as focal mechanisms. Just plotting the events shows that they are far from homogeneously located in space, and at best marginally stationary in time. Variances increase rapidly and nonlinearly with the scale of the space-time region. Under such conditions, even the simplest statistical task, such as estimating a mean or a trend line, becomes surprisingly difficult. Nor is the situation helped by the many sources of error (systematic and otherwise) in the catalogue entries, including those arising from the natural compulsion practical seismologists feel to improve and hence alter their network and methodology. One is rapidly pushed into the penumbral region where standard assumptions such as independence and stationarity start to break down, the data is unreliable, and conclusions drawn from the analysis have to be interpreted with great caution. Scientists such as Ted Habermann and Yan Kagan have been drawing attention to such difficulties, including the danger of taking man-made artefacts for physical realities, for some time, but their real seriousness has only been felt in recent years, as long-term or even medium-term hazard estimates are increasingly requested for policy decisions, and differences in seismic and geological estimates of quantities such as slip-rates become of increasing concern.

Our own first steps towards formulating a more serious model were made during the next summer, and were initiated by the arrival of two talented young scientists, Robert Davies and Selwyn Gallot, one of whom (Robert) was assigned to work with me on time series properties of the catalogue data. Robert was already something of a time series specialist, and together we developed a simple 'trigger model' (a modification of the classical 'shot noise' model that translates into a one-dimensional (time) cluster model with cluster centres followed by sequences of independently located secondary events) that was able to mimic the main time series properties of the catalogue data, while based on a crude physical picture of the process generating the events.

In trying to fit this model to the data, we first 'binned' the numbers of earthquakes into equi-sized time intervals. We were teased by the sense that the correlation properties of this binned data should be representable in terms of some underlying function that was independent of the bin size, but we could not quite pin down what this was. Then Bartlett's paper on point-process spectra (Bartlett, 1963) reached New Zealand (after the usual six-month delay required for surface mail to reach our shores) and the explanation became clear. The concept we had been missing was the 'covariance density', and its Fourier transform, now called the 'Bartlett spectrum' (see, for example, Chapter 8 of Daley and Vere-Jones, 2003). The seismicity studies had brought us close to discovering these concepts for ourselves; whether we would have done so is a moot point, but at least we were moving in the right direction.

Even the relatively conventional study that Robert and I were able to put together on the basis of this model (Vere-Jones and Davies, 1966) was enough to confirm some obvious but nonetheless important features of the catalogue data. The model allowed us to quantify the high degree of clustering, to note the differences in the extent of clustering with depth, and to pick up the fact that clustering was a long-tailed (power-law decay) rather than a short-tailed feature.

This study was a turning point. It brought our efforts to the attention of seismologists overseas, and encouraged me to start a more systematic attempt to consolidate the mathematical background to such point process models. Gradually, I came into contact with a group of seismologists with a shared interest in the statistical analysis of seismic data. Notable among these were Cinna Lomnitz, then in Chile, V. Gaisky in the Soviet Union, T. Utsu and G. Suzuki in Japan. This group, bolstered by the sympathetic interest of some key figures such as Keiiti Aki, Bruce Bolt, T. Mogi and others, might be considered the veterans of the Statistical Seismology brigade.

This phase of my work culminated in the review paper, Vere-Jones (1970), the value of which was greatly enhanced by the contributions to the discussion, following the Royal Statistical Society's excellent practice of seeking and publishing comments from informed specialists. Among other things, the review paper was instrumental in helping me to link up with Yan Kagan, one of the most original and active figures of the next few decades, and now a long-time friend and colleague. Yan at this time was working in the Institute of Mines not far from Moscow, and prevented, for complex reasons, from joining the equally unconventional group set up around the same time by Keilis-Borok. Yan had read the review paper, and was seeking contacts beyond the Soviet Union. When I first met him in the early 1970s, I was taken aback to discover that he knew more about stochastic point processes and their analysis than I had been able to pick up in the West. I played a small role in trying to get his work more widely recognized, and in supporting Leon Knopoff in his efforts to find Yan a position outside the Soviet Union.

3. Models for fracture

In descriptive models it is not essential that the model reflects fully the physical processes generating the data. What is important is that the description be close enough that Monte Carlo simulations, based on the model, reproduce at least the main features of the data. Then the model can be used to quantify the strength of the key features (such as clustering in our case) to compare the strengths of such feature in different data sets, and to judge the extent to which the real data set differs from the data produced by the model.

There is, however, a quite different role for statistical models, exemplified by the classical statistical mechanical models for fluids, phase changes and turbulence. Here the model has a genuine explanatory function, the statistical theory being used, typically, to show how macroscopic properties flow from a broad class of models with similar microscopic behaviour. Then the exact behaviour at the microscopic level, which may be extremely complex and beyond direct measurement, is seen to be inessential, and can therefore be crudely approximated without impairing the validity of the resulting macroscopic theory.

It has always been one of my main hopes that a statistical model could achieve something similar for the processes of earthquake fracture, i.e. that fracture can be regarded as a macroscopic feature resulting from the mass interaction of microscopic features such as dislocations, whose behaviour in detail did not need to be specified precisely. During the 1970s I became excited by the possibility that a simple branching process model might serve this role. Over many summer nights I gazed at the cracks in the ceiling, hoping for a mathematical revelation.

Eventually a consistent model did result, spurred into being by a paper of Scholz's which I wanted to refute; it resulted in the papers of Vere-Jones (1976, 1977). Even now, I think it is a model that deserves attention, not because it is a very realistic model of the process of fracture development, but because it is very simple, so that its properties can be developed in detail, and yet it manages to reproduce a number of key features of the earthquake process. Let me briefly outline the model and what can be deduced from it.

We imagine a fault surface, comprising a vast number of irregularities or weaknesses, any one of which may fail, and then become an 'ancestor' in a cascade of dislocations each one of which may trigger additional dislocations ('offspring') in its neighbourhood. Percolation processes form a wider class of models of the same general kind. The distinguishing feature of the branching model, which renders it susceptible to analysis, but also represents the main approximating factor in this context, is that each 'ancestor' produces 'offspring' independently of its neighbours or predecessors. The simplest version of this model assumes further a homogeneous, indefinitely extended environment. These assumptions imply that, in the model, the spread of the rupture (birth of additional generations) will not be affected by boundary effects or depletions in the stored energy, an obvious idealization from a physical viewpoint.

As is well-known from its use as a model for nuclear fission, a branching process can be either subcritical (when ρ, the expected number of offspring per ancestor, satisfies $\rho < 1$), critical ($\rho = 1$), or supercritical ($\rho > 1$). The latter case leads with high probability to an uncontrolled or infinite rupture; the first case leads to small events ('families') which quickly die out, while the critical case leads still to finite events, but ones of very variable size. Indeed, it is the distribution of the size of the total family which is of greatest interest here, for it controls the frequency-magnitude law. Supposing the size to be proportional to the total energy release (and hence to the exponential of the magnitude), a closer analysis, shows that for a subcritical but nearly critical process the size distribution is approximately of the tapered Pareto form (see, e.g. Kagan, 1999)

$$Pr\{\text{Energy} > E\} \propto E^{-1/2} e^{-E/E_0}. \qquad (1)$$

As the process approaches criticality, the parameter E_0 approaches infinity, and the distribution approximates a pure power-law form corresponding to a Gutenberg–Richter law with b-value $b \approx 2/3$.

At a qualitative level, at least, it is not difficult to imagine how such a process might operate in a more heterogeneous region. Driven by the larger-scale tectonic forces, pockets in the stressed region, smaller at first but larger in the sequel, accumulate energy until, locally, the parameters of the branching process approach or even exceed the values required for the process to become critical. Before this stage, either events are difficult to initiate, or, once initiated, quickly run their course. Beyond this stage, they start to spread, as in the critical or supercritical state, until terminated either by the inherent variability of the process, or by reaching a physical boundary of the stressed region.

Although this is a 'toy' model in the sense that its physical assumptions are skimpy and approximate, it is enough to indicate those features of the earthquake process which might reasonably be attributed to the randomly terminated spreading of a fracture from point to point across the fault face. The Gutenberg–Richter law is the prime example of such a feature, but it can be derived from such a large range of potential models that it is not a conclusive demonstration of the model's underlying relevance. More convincing is the appearance of Kagan's distribution as the natural modification to be expected when the system falls just short of criticality. Further variations, including the appearance of Gutenberg–Richter laws with higher b-values which can arise when the data under analysis is a mixture from sources with possibly different stress environments or spatial limitations (cf the discussion in Vere-Jones et al., 2001), also lend support to the underlying correctness of the conceptual basis, however over-simplified.

Of course, similar behaviour can be expected from other models possessing subcritical, critical, and supercritical phases, percolation processes being a case in point (although we expressed some hesitations about the use of percolation models in Bebbington et al., 1990). There is also much in common with phase-change models, and with processes showing self-organized criticality. Indeed, the application of statistical mechanical ideas to fracture processes has been taken much further in recent years, for example by Ian Main and colleagues in Edinburgh and by John Rundle and colleagues in Boulder. The problem with such developments, at least from my point of view, is that as the models increase in technical difficulty, they lose the advantage of simple analysis while remaining approximate and hard to fit to observational data.

4. Evolutionary models and probability forecasts

The ability to develop simulated catalogues from a stochastic model is one of the most significant benefits conferred by the great computing power currently available. It allows the model characteristics to be determined numerically in full detail, graphically as well as numerically, and opens up direct paths both to test the model, and to make probability forecasts based on the model. For goodness of fit tests, for example, one simply makes a comparison between the observed value of some feature in the real catalogue, and the histogram of values (for the same feature) obtained from repeated simulations of the artificial catalogue. If the value from the real data lies in the far tail of the histogram, then the real data differs significantly from the model data in this respect. For forecasting one must do a little more. First, one must fit the model using all available data up to the present. Then one must simulate this fitted model into the future, repeating the simulations many times to make a histogram of values for the feature to be forecasted. This histogram then gives an approximate error distribution for the forecast value. From it one can extract the forecast probability that the value will lie within certain bounds. Equally, one can take averages over the simulations to obtain expected values of the forecast quantities. Using a Bayesian approach, the simulations can be modified to take into account initial uncertainties in the parameters, or even uncertainties in the model formulation itself (cf the discussion in Vere-Jones, 1995).

Not all models lend themselves equally to this dual process of successively fitting the model and simulating forward in time. Our original trigger model is relatively difficult to handle from this point of view because it has a complicated likelihood function, requiring averaging over possible cluster centres, which renders fitting difficult. The next decade, however, saw the appearance and rapid development of ‘conditional intensity models’ (‘evolutionary models’) in which the structure of the process is completely summarized by the so-called ‘conditional intensity function’, representing an instantaneous rate (intensity) conditioned by all available past information:

$$\lambda(t|\mathcal{H}_t) = E\left(\frac{\mathrm{d}N(t)}{\mathrm{d}t}|\mathcal{H}_t\right),$$

where $\mathcal{H}_t$ represents the past information available up to time t, including in particular the catalogue of past events. Although $N(t)$ here is an increasing step function, its expectation increases smoothly and for many processes the conditional intensity can be precisely defined. Provided the conditional intensity can be so defined for the particular model in view, and its form is easily computable from the past of the process, then fitting, simulation and forecasting can all be reduced to relatively straightforward, routine procedures.

The application of such models to earthquake data is linked in my own history to two major visits to Asia. The first occurred in 1976, when my family and I made a memorable visit to Japan as guests of Akaike and colleagues at the Institute of Statistical Mathematics in Tokyo. Akaike quickly seized on the potential usefulness of such models, in particular the ‘self-exciting’ models introduced by Hawkes at the beginning of the decade (Hawkes, 1971; Hawkes and Oakes, 1974). Two of Akaike’s students, Yosi Ogata and Tohru Ozaki, were assigned to help me during this visit, and with their help we were able to develop and fit prototype versions both of the Hawkes’ model (Ozaki, 1979; Ogata and Akaike, 1982) and a stochastic version of Reid’s elastic rebound model which we called the ‘stress-release model’ (SRM) (Vere-Jones, 1978; Ogata and Vere-Jones, 1984). Both ‘students’ have remained staunch friends, greatly adding to my enjoyment of Japanese life, as well carving out notable careers for themselves.

The work on the Hawkes’ model was subsequently developed by Ogata into the ‘epidemic-type-aftershock-sequence’ (ETAS) model (e.g. Ogata, 1988, 1998). In his hands, it has expanded to include dependence on the location as well as the size of the initiating event, and to accommodate fine variations of the model parameters in space and time. It is currently at the centre of detailed statistical studies of stress shadowing, where its ability to capture local departures from the standard parameters are crucial to show up areas of modified seismicity (Ogata et al., 2003; Ogata, 2004). His student Zhuang has developed an interesting stochastic declustering procedure (Zhuang et al., 2002, 2004), based on the ETAS model, and allowing explicit (but stochastic) allocation of events to specified ‘ancestor’ events, following up an initial idea of Kagan’s (e.g. Kagan, 1991).

If the major use of the ETAS model is in capturing variations in short-term clustering effects, the stress-release model, by contrast, seeks to model slower, longer-term variations in activity, associated with the build-up and release of stress within a region. It preceded the ‘time-predictable’ and ‘stress-predictable’ models, which appear as limiting cases, and has been quantitatively fitted to historical records of large events (e.g. Zheng and Vere-Jones, 1994). More recently, it has been extended by Shi Yaolin and colleagues to a model for several interacting regions (Liu et al., 1998; Shi et al., 1999), although it is still limited in its ability to handle spatial features (cf Bebbington and Harte, 2003). To better capture the spatial effects, a more careful analysis of stress transfer is needed, leading perhaps to statistical models for the evolution of the stress field on a family of interrelated faults, as has long been advocated by Kagan (see, e.g. Kagan, 1982, or the discussion in Kagan and Vere-Jones, 1996 and further references therein), but this still appears a task of daunting complexity.

The two models describe almost completely disjoint aspects of the catalogue. They can be combined (see Schoenberg and Bolt, 2000), but not a great deal is gained thereby: on the longer time-scale, inclusion of aftershocks, as in the ETAS model, merely shifts the proportion of stress released by aftershocks and by smaller independent events, without altering the major large-scale features; on the shorter time-scale, the processes of stress build-up and release have little effect on the aftershock patterns described by the ETAS model. Both models are incorporated into David Harte's 'Statistical Seismology Library' (Harte, 1998), a collection of statistical and graphical routines for modelling earthquake catalogues, written in the statistical language 'R'.

5. Earthquake prediction

I became involved in prediction issues through the influence of Frank Evison. Like him (see, e.g. Rhoades and Evison, 1979), I strongly hold to the view that earthquake prediction is properly set in a probability context, since the uncertainties in making such predictions seem unlikely to disappear, and, while present, they need to be quantified. I have also made liberal use of many other of his ideas, such as the use of 'synoptic maps' to display the spatial variation of probability forecasts. Nevertheless, until a decade or so ago, earthquake prediction in a practical sense appeared to lie outside the modelling work on which we had concentrated.

My attitudes changed substantially with a second Asian visit, to Beijing in 1996. Here I was a guest of Ma Li at the Centre for Analysis and Prediction of the State Seismological Bureau, more recently China Seismological Bureau, most recently (as I understand) China Earthquake Administration. If prediction was a controversial topic for Western seismologists, it was the bread and butter of my Chinese colleagues, indeed the very reason for their existence, since earthquake prediction was the primary task of the government-funded Bureau. Ma Li was in charge of a section concentrating on statistical methods, and had invited me to help them introduce stochastic modelling ideas, and associated probability forecasting techniques, through a series of lectures and demonstrations to her colleagues and the students of Shi Yaolin at the neighbouring University.

The major problem I discovered was the lack of statistical expertise in the Seismological Bureau. As far as I could ascertain, the Bureau did not employ a single statistician among its many thousands of scientific staff. In fact, it is only very recently that the Chinese universities have started to produce statistics graduates with the applied modelling and analysis background that is needed in such an environment. The other major revelation to me was the wide variety of phenomena being systematically studied and recorded, and the use of this data, at least informally, in preparing prediction statements which from time to time were passed on to senior government officials and turned into official warnings. This paradoxical conjunction, of the collection of a large amount of varied data, its real-life use in making predictions, and the lack of any statistical expertise to properly model or assess the data, or even to exercise quality control over its selection and archiving, was both tantalizing and frustrating — rich in possibilities, hard to alter.

While I see much to criticize in some of the work of the Chinese scientists, their inclusive approach to the phenomena which should be examined as part of a comprehensive programme to predict earthquakes is not one of them. I only wish that Western seismology had been similarly embrasive. For in that case we would have thirty years' worth of first class data with which to address the question of whether such phenomena are or are not relevant. As it is, their relevance still remains uncertain — probably not large, but perhaps sufficient to reinforce estimates of risk derived from catalogue and deformation studies. Certainly, our recent study of the Chinese data on electrical signals (Zhuang et al., 2005) provides a prima facie case for reconsidering this phenomenon on a wider scale, just as it illustrates the possible nature and uses of stochastic models for this wider class of phenomena.

Statistical seismology has no reason to disparage the work of the Chinese seismologists, for it was Ma Li's energy and enthusiasm that initiated the current programme of international workshops in statistical seismology, the first two of which were held in China, the third in New Zealand, and the fourth and most recent in Mexico. The next is likely to be in Japan, and we look forward to future appearances in the United States and Europe.

6. Concluding remarks

The last decade has seen an influx of new concepts, new data, and new procedures, which combine to make the present time as exciting as any for statistical seismology. New concepts include new mathematical structures, such as self-similarity, fractal growth and dimension and self-organizing criticality, for which existing statistical techniques, based as most of them are on assumptions of stationarity and ergodicity, are inappropriate. In this area at least, seismology is once more challenging the statisticians to enlarge and update their tool box. New data include in the first instance data from the precise strain-measurements made by GPS technology. Already these are having a major impact in elucidating the occurrence of 'slow earthquakes' which may require a major reevaluation of existing models and model classes. In the past, the complexity of stress transfer, and the difficulty of direct stress measurements, impeded the incorporation of stress changes into stochastic earthquake modelling, but new understanding, new algorithms and more powerful computers may reverse this situation. Finally, there remains that whole suite of postulated, non-catalogue-based precursory phenomena, which, again in my personal opinion, have yet to be adequately tested, having first been embraced incautiously, without sufficient awareness of the statistical pitfalls, in the heady days of the 1970s, and more recently relegated to a low priority, still without adequate data, in the sceptical atmosphere of the 1990s.

A long time ago, after the first seminar I gave to the Geophysics series at Victoria University, Trevor Hatherton remarked a little caustically that "I won't really believe all this statistics stuff until it gives me something useful." (Or words to that effect, since he expressed himself forcefully even at seminars.) I recognized that he had laid down a challenge I could hardly ignore. I do indeed believe that statistical models can contribute significantly both to understanding seismic processes and to increasing the utility of seismological studies. To demonstrate this usefulness is a task on which I have been engaged, with considerable enjoyment but not a huge amount of success, for the last 40 years. Perhaps even now Trevor would not be convinced; but at least I feel that he would have to consider the possibility more seriously.

References

Aki, K., 1956. Some problems in statistical seismology. Zisin 8, 205–228 ((in Japanese) English translation by A.S.Furimoto, Hawaii, 1963).

Bartlett, M.S., 1963. The spectral analysis of point processes. Journal of the Royal Statistical Society. Series B 29, 264–296.

Bebbington, M., Harte, D., 2003. The linked stress release model for spatio-temporal seismicity: formulations, procedures and applications. Geophysical Journal International 154, 925–946.

Bebbington, M., Vere-Jones, D., Zheng, X., 1990. Percolation theory: a model for earthquake faulting? Geophysical Journal International 100, 215–220.

Daley, D.J., Vere-Jones, D., 2003. An introduction to the theory of point processes (second edition). Vol I: Elementary Theory and Methods. Springer, New York.

Eiby, G.A., 1967. Earthquakes: A Complete Account of Their Causes, Effects and Known History (revised edition). Frederick Muller, London.

Harte, D., 1998. Documentation for the statistical seismology library. School of Mathematical and Computing Sciences Research Report No. 98–10. Victoria University of Wellington.

Hawkes, A.G., 1971. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58, 83–90.

Hawkes, A.G., Oakes, D., 1974. A cluster representation of a self-exciting process. Journal of Applied Probability 11, 493–503.

Jeffreys, H., 1973. Scientific Inference, 3rd edition. Cambridge University Press, Cambridge, UK.

Kagan, Y.Y., 1982. Stochastic model of earthquake fault geometry. Geophysical Journal of the Royal Astronomical Society 71, 659–691.

Kagan, Y.Y., 1991. Likelihood analysis of earthquake catalogues. Geophysics Journal International 106, 135–148.

Kagan, Y.Y., 1999. Universality of the seismic moment-frequency relation. Pure and Applied Geophysics 155, 537–573.

Kagan, Y.A., Vere-Jones, D., 1996. Problems in the modelling and statistical analysis of earthquakes. In: Heyde, C.C., Prohorov, Yu.V., Pyke, R., Rachev, S.T. (Eds.), Athens Conference on Applied Probability and Time Series, Applied Probability, vol. 1, pp. 398–425.

Liu, J., Vere-Jones, D., Ma, L., Shi, Y., Zhuang, J., 1998. The principle of coupled stress release model and its applications. Acta Seismologica Sinica bf 11, 273–281.

Ogata, Y., 1988. Statistical models for earthquake occurrence and residual analysis for point processes. Journal of the American Statistical Society 83, 9–27.

Ogata, Y., 1998. Space–time point process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics 50, 379–402.

Ogata, Y., 2004. Space–time model for regional seismicity and detection of crustal stress changes. Journal of Geophysical Research 109 (B3), B03308. doi:10.1029/2003JB002621.

Ogata, Y., Akaike, H., 1982. On linear intensity models for mixed doubly stochastic Poisson and self-exciting point processes. Journal Royal Satistical Society, Series B 44, 102–107.

Ogata, Y., Vere-Jones, D., 1984. Inference for earthquake models: a self-correcting model. Stochastic Processes and Applications 17, 337–347.

Ogata, Y., Jones, L., Toda, S., 2003. When and where the aftershock activity was depressed: contrasting decay patterns of the proximate large earthquakes in southern California. Journal of Geophysical Research 108 (B6), 2318. doi:10.1029/2002JB002009.

Ozaki, T., 1979. Maximum likelihood estimation of Hawkes' self-exciting point process. Annals Institute of Statistical Mathematics, Series B 31, 145–155.

Rhoades, D.A., Evison, F.F., 1979. Long-range earthquake forecasting based on a single precursor. Geophysical Journal of the Royal Astronomical Society 59, 43–56.

Schoenberg, F., Bolt, B., 2000. Short-term exciting, long-term correcting models for earthquake catalogs. Bulletin of the Seismological Society of America 90, 849–858.

Shi, Y.L., Liu, J., Chen, Y., Vere-Jones, D., 1999. Coupled stress release models for time-dependent seismicity. Journal of Pure and Applied Geophysics 155, 649–667.
Vere-Jones, D., 1970. Stochastic models for earthquake occurrence. Journal of Royal Statistical Society, Series B 32, 1–62.
Vere-Jones, D., 1976. A branching model for crack propagation. Pure and Applied Geophysics 114, 711–725.
Vere-Jones, D., 1977. Statistical theories of crack propagation. Mathematical Geology 9, 455–481.
Vere-Jones, D., 1978. Earthquake prediction — a statistician's view. Journal of the Physics of the Earth 25, 129–146.
Vere-Jones, D., 1995. Forecasting earthquakes and earthquake risk. International Journal of Forecasting 11, 503–538.
Vere-Jones, D., Davies, R.B., 1966. A statistical study of earthquakes in the main seismic area of New Zealand. Part II: Time series analyses. New Zealand Journal of Geology and Geophysics 9, 251–284.
Vere-Jones, D., Smith, E.G.C., 1981. Communications in Statistics - Theory and Methodology A10, 1559–1585.
Vere-Jones, D., Robinson, R., Yang, W., 2001. Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation. Geophysical Journal International 144, 517–531.
Zheng, X., Vere-Jones, D., 1994. Further applications of stress release models to historical earthquake data. Tectonophysics 229, 101–121.
Zhuang, J., Ogata, Y., Vere-Jones, D., 2002. Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Society 97, 369–380.
Zhuang, J., Ogata, Y., Vere-Jones, D., 2004. Analyzing earthquake clustering features by using stochastic reconstruction. Journal of Geophysical Research 109, B05301. doi:10.1029/2003JB002879.
Zhuang, J.C., Vere-Jones, D., Guan, H., Ogata, Y., Ma, L., 2005. Preliminary Analysis of the Precursory Information in the Observations on the Ultra-Low Frequency Electric Field in the Beijing Region (to appear in Pageoph special issue on Statistical Seismology).