STATISTICAL ANALYSIS

OF

EARTHQUAKE CATALOGS

by

JOZEF FRANS MARIA VAN DYCK

Ir. Bouwkunde, Katholieke Universiteit Leuven
(1978)

S.M. , Massachusetts Institude of Technology
(1981)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

December, 1985

Signature of Author _____
Department of Civil Engineering, December 31, 1985

Certified by _____
Thesis supervisor

Accepted by _____
Chairman, Departmental Committee on Graduate Students

STATISTICAL ANALYSIS

OF

EARTHQUAKE CATALOGS

by

JOZEF FRANS MARIA VAN DYCK

Submitted to the Department of Civil Engineering on December 31, 1985
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy.

ABSTRACT

In the context of seismic hazard quantification, the main objective of statistically analyzing an earthquake catalog is to estimate the rate of earthquake events as a function of magnitude and geographical location. Four mayor problems that arise in such an analysis are addressed in this thesis:

1. Earthquake size is often reported in different scales and needs to be expressed in a uniform scale. Statistical techniques that account for nonlinearity of the regression of one size measure against another, the heteroscedasticity of the regression error, and the presence of outliers are proposed. Corrections for the effect of measurement errors in the data and incompleteness of the earthquake sample are derived on a theoretical basis. An approximate formula to combine several reported size measures to a single scale is also presented. Finally, a conversion formula is proposed that differs from the regression curve and corrects for bias in the estimation of the recurrence rates.

2. The earthquake sequence typically displays a high degree of clustering. Clustering must be included in the statistical model, or the original catalog must be thinned through removal of the dependent events prior to further analysis based on the Poisson assumption. A method to identify clusters in the catalog has been developed. The procedure differs from earlier ones in that it allows the extent of the cluster in space and time to vary for each main earthquake and accounts for temporal and spatial variation of the observed recurrence rates (temporal variation is caused mainly by incomplete reporting).

3. The reported data is invariably incomplete, especially for events of small magnitude and in early time periods. Several methods of varying complexity are presented to account for this incompleteness. The methods differ in a fundamental way from those currently in use: they represent

incompleteness explicitly through a probability of detection that varies with magnitude, time and spatial location and estimate this probability simultaneously with the recurrence rate from the historical data. Models in which the probability of detection accounts for the temporal and spatial distribution of population and instruments are also presented.

4. Another major novelty is the extension of the usual notion of seismogenic provinces with uniform recurrence rate to provinces with smoothly varying recurrence rates. A maximum penalized likelihood method is proposed for the estimation of the recurrence rates and allows to control the degree of smoothness through a few input parameters. The method of estimation is further developed to account for errors in the epicentral location and magnitude of the reported events.

Thesis Supervisor: Daniele Veneziano

Title: Professor of Civil Engineering

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor, Professor Daniele Veneziano, for his guidance and assistance. Working with him has been a very pleasant experience.

I also like to acknowledge the members of the thesis committee, Professors Gregory Baecher, Steven Ellis, and Nafi Toksoz, for their comments and suggestions.

The assistance of Tom O'Hara in producing some of the figures is greatly appreciated.

I like to dedicate this thesis to my wife, Bea.

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

Chapter 1

STATEMENT OF THE PROBLEM

Evaluation of seismic hazard at a given site typically relies on historical seismicity in a region around that site. The statistical inference of future events on the basis of past activity poses however several problems:

1. Earthquakes are typically reported in different magnitude scales and a conversion of these scales to a single size measure is necessary.
2. Seismic data invariably displays a considerable degree of clustering, which is contrary to the common assumption of Poisson events.
3. Historical reporting of events is incomplete, especially for low magnitudes and early time periods.
4. The historical data often does not support the hypothesis of homogeneous seismicity within extended geographical regions.

Although one could formulate a statistical model that incorporates all of the above characteristics, the statistical estimation of the parameters of such a model would be prohibitively complicated, unless drastic simplifying assumptions are made. In this thesis, it is preferred to address separately the problems of magnitude conversion, clustering, incompleteness and estimation of the recurrence rates. Contrary to current practice, the latter two problems are considered simultaneously in

this thesis. The present chapter reviews the above four problems and methods proposed in the literature. Several important features of the data that have not been previously considered will be indicated. Statistical techniques that address these issues are developed in Chapters 2, 3 and 4, for problems 1, 2, and simultaneously 3 and 4 respectively. Chapter 5 summarizes the new methods of analysis and states conclusions.

The problem of magnitude conversion has received relatively little attention in the literature. Various authors have published conversion formulas of one size measure to another (e.g. Nuttli, 1974; Street and Turcotte, 1977). However, the regression lines obtained by different authors seldom agree due to regional variations of the regression relationship and to differences in the inference method. There is a need to establish a general procedure for deriving conversion formulas from the historical data. One such procedure is developed here, which includes the following characteristics of the data set:

- the relationship between two size measures may be nonlinear
- the regression error may be a function of the regressor
- outliers may be present
- the reported size measures may include measurement errors
- some events may have more than one size measure reported

The regression line and the distribution of the residual describe the conditional distribution of the dependent variable, given the independent variable. In application to magnitude conversion, complications arise, because the learning sample (i.e. the data set that is used to estimate the regression line) may have different characteristics from the prediction sample (i.e. from the data set for which the regression is to be used). For instance, measurement errors may differ from earthquake to

earthquake and the degree of incompleteness may create further discrepancies between the two samples. The influence of measurement errors on magnitude conversion is discussed by Ganse et al. (1983). These authors assume however that the marginal distribution of the earthquake sizes are Gaussian, contrary to the usual assumption of exponentiality. The implications of assuming exponential rather than normal distributions will be discussed in Chapter 2.

Finally, it is important that the estimates of the recurrence rates be invariant with respect to the chosen magnitude scale. This is not the case if one uses the regression equation as a conversion formula. This problem is also studied in Chapter 2 and motivates a correction to the regression relationship for use in magnitude conversion.

The phenomenon of earthquake clustering has generated much interest. Several statistical models allowing for clustering of events have been proposed (e.g. Vere-Jones, 1970; Kagan and Knopoff, 1976). Various empirical relations for the occurrence of clusters and the distribution of counts in a cluster have also been developed (e.g. Utsu, 1969). The influence on seismic hazard of the dependent events within a cluster has been studied by Wally (1976) and by Merz and Cornell (1973). This thesis focuses on procedures that classify historical earthquakes as either "independent" or "dependent" events: such procedures are generally less restrictive with respect to the stochastic model that describes the earthquake sequence. They also prvide information that facilitates estimation of at least of one such model, the Neymann-Scott model, according to which the earthquake sequence is a superposition of two processes. The first process is composed of earthquakes with independent locations, times of occurrence and magnitudes ("independent" events),

whereas the second process is triggered by the first and includes all "dependent" events. The dependent events are further assumed to be of magnitude not larger than that of the associated independent event.

A procedure to identify clusters in earthquake sequences has been recently proposed by Prozorov and Dziewonski (1981). In their study, the degree of closeness between two earthquakes that is considered significant for clustering is obtained from a statistical comparison of the earthquake count within a certain window (in geographical location, time and magnitude space) relative to the count generated by a Poisson process. The procedure fails however to account for the event-to-event variation of the size of the cluster windows, which has been noted by various authors (for instance, Simpson and Richards, 1981). A procedure that allows for such variations is developed in Chapter 3. In applying the method to actual earthquake data, the shape and size of the clusters are indeed found to be highly variable.

Incompleteness of the the earthquake catalog is of mayor concern in the estimation of recurrence rates. Incompleteness not only may introduce bias but also confounds the spatial variation of seismicity, if incompleteness itself varies in space. Current procedures therefore limit the estimation of recurrence rates to data in the most recent periods of the catalog which are judged to be complete. Such "periods of completeness" depend on the magnitude of the events and are typically based on knowledge of the detection capability of people and instruments and on the the historical data. The estimation of spatial variation of incompleteness is similarly based in part on judgement, in part on data. Apart from the subjectivity, there are other serious limitations to current procedures:

- earthquakes of small magnitude may be incomplete even today and should therefore not be considered in the analysis, if only the complete portions of the catalog are to be used.
- only part of the data is used, while even incomplete data are informative, e.g. on the relative spatial distribution of recurrence rates if incompleteness is spatially constant
- the estimation of the recurrence rates and incompleteness are coupled problems. For instance, the assumption that the recurrence rates vary exponentially with magnitude is informative on incompleteness, given the historical data. Such information is not considered in present analyses.

The approach developed in this thesis consists of using all the historical data to simultaneously estimate recurrence rates and incompleteness. To do so, incompleteness is represented through the probability of detection, which varies as a function of time, magnitude and geographical location. The notion of a probability of detection has been used earlier by Brillinger (1979) and by Kelly and Lacoss (1969). None of these authors models incompleteness to the degree of detail proposed in this thesis.

In Chapter 4, four models of varying complexity are examined for rates and incompleteness. In all models, the estimation of incompleteness is primarily data based. In two of them, the temporal and spatial variation of population and seismic instruments is explicitly accounted for, leading to a refined spatial description of incompleteness. Several new ideas are also presented for the estimation of the spatial variation of the recurrence rates. In current practice, it is typically assumed that the recurrence rates are constant within specified regions. Such

regions are not easily determined on the basis of seismicity or physical information and may indeed not even exist. A more general nonparametric description of the spatial variation of seismicity is proposed here, which includes the case of homogeneous earthquake sources as a special case. Different techniques, such as maximum penalized likelihood and kernel methods, are considered in Chapter 4 for the estimation of the parameters of this model. The estimation procedure is further extended to account for measurement errors on the earthquake location and size and methods to validate the model and calculate uncertainty on the estimates are developed.

Chapter 2

MAGNITUDE CONVERSION

2.1 INTRODUCTION

A typical entry in an earthquake catalog reports the time of occurrence, t, the epicentral location $\underline{x}$ and one or several size measures $\underline{m}$. In principle, one could model such data as a marked point process in time and space with a random size vector $\underline{m}$ associated with each point. Such a multivariate representation of earthquake size is however impractical in the analysis of clustering, incompleteness and recurrence rates. A more convenient alternative is to convert the set of reported size measures to a single scale prior to further analysis. Published conversion formulas are usually in the form of involving just two size measures (e.g. for the Eastern U.S., Chiburis, 1981; Nuttli, 1974; Street and Turcotte, 1977; WGC, 1982). As illustrated in Figure 2.1, differences among published regression lines can be considerable and it is not always evident which relationship one should use for a given set of earthquake data. Differences may be attributed to several causes, e.g. regional dependence (Chung and Bernreuter, 1980), the use of different estimation methods (in particular, the different degree of trimming to exclude incomplete data) and differences in the data for different catalogs. In view of these variations, it is often desirable to estimate conversion rules directly from the catalog under consideration. The problem of how to best estimate magnitude conversion rules has received little attention in the literature beyond the level of fitting simple linear regressions to the data. The approach taken in this chapter is novel in the following respects:

1. Two nonlinear regression techniques, robust locally weighted regression and linear spline regression, are applied to the estimation of the chosen size measure from a single other size measure. These methods are of interest because they both accommodate nonlinearity of the regression in a flexible manner. Furthermore, they can account for heteroscedasticity of the regression error and for the presence of outliers. These techniques are discussed and exemplified in Section 2.2.
2. The influence of measurement errors and incompleteness of the data is investigated in Section 2.3. For this purpose, one needs to specify a joint distribution model for the true and estimated size measures. The model of Section 2.3 has exponential marginal distributions and normal conditional distributions. These are common assumptions in magnitude-conversion analysis.
3. A simple approximate formula is derived in Section 2.4 to estimate the chosen size measure from several other size measures. General methods such a multiple regression are typically not applicable because too few data are available with the same set of size measures reported and because the multiple regression may be nonlinear.
4. The interaction between magnitude conversion and the estimation of recurrence parameters is discussed in Section 2.5. The problem here is that the recurrence rate should be the same if one uses earthquakes for which the chosen measure of size has been converted or directly estimated. This condition is not satisfied if the regression (the conditional mean) is used to convert other size measures to the chosen scale and uncertainty

around the regression is neglected. Distribution properties of sizes obtained through various conversion rules are derived and compared in Section 2.5. One of these rules produces unbiased estimates of the recurrence rate. The need for a correction of this type is not recognized in practice or in the literature, although the correction is substantial if uncertainty around the regression is large. A correction is needed also for direct estimates in the chosen scale, if one wishes to express results in terms of actual rather than reported earthquake size.

2.2 REGRESSION OF A SIZE MEASURE AGAINST A SINGLE OTHER SIZE MEASURE

Frequently, only two size measures are reported in an earthquake catalog. For instance, the size of early events may be measured on an empirical scale, whereas recent events are usually instrumentally recorded. The problem addressed in this section is how to estimate regression relationships between the two scales. For example, Figure 2.2 shows a scatterplot of data from the Chiburis catalog. It should be emphasized that, although formal statistical techniques are proposed in this section to estimate the regression, these techniques are not a substitute for careful inspection of the data. In particular, the following issues should be considered:

- the composition of the catalog. Is the catalog a mixture of two or more catalogs with possibly different estimators of each size measure?
- independence of the reported size measures. Are some of the reported size measures obtained through conversion from other size measures?

- dependence of the relationship of interest on covariates such as geographical location, time and focal depth.

Some of these issues are illustrated by the Chiburis data: In Figure 2.3 each observed value of magnitude is represented by a number indicating the decade since 1900 when the earthquake occurred. Only values of I_0 that are accurately reported in the catalog are shown (some of the earthquakes in the catalog have alternative values of I_0 indicated). Italics are used to indicate earthquakes observed in Canada. From this plot, it appears that for the Canadian data the regression is steeper and higher than for the U.S. data. This is consistent with Figure 2.2, where different symbols are used for different geographical regions. The plot of Figure 2.3 presents more clearly the marginal distribution of M for given I_0 and emphasizes the presence of outliers, heteroscedasticity and grouping of the data. The latter phenomenon is possibly an indication of dependence among groups of data. Variation of the regression with time is not very clear from Figure 2.3, but this variation is more evident in Figures 2.4a and 2.4b, where the data are separated according to time. The question whether a given data set is dependent or whether a particular datapoint is erroneous, is not formally addressed in this section, because it requires detailed information about the operation of the seismic network and the estimation of each earthquake size, and falls outside the scope of this work. The techniques presented in this section do however account for nonlinearity of the regression, outliers, and heteroscedasticity of the regression error and are thus a considerable improvement over simple linear least-squares regression. Estimates of uncertainty on the regression are also obtained from both methods. These estimates allow one to

judge whether the differences between the two regressions are significant or not.

2.2.1 Robust Locally Weighted Least Squares

A robust locally-weighted least-squares method (RLWLS) of regression has been proposed by Cleveland (Cleveland, 1979; Cleveland and McGill, 1984). This is an iterative non-parametric regression technique. During the first iteration, the regression is estimated at each point by fitting a straight line to the local data, as illustrated in Figure 2.5. In subsequent iterations, each datapoint is weighted, depending on its distance from the estimated regression line. Weighting reduces the influence of outliers and hence robustifies the estimated regression line. Because of possible heteroscedasticity of the error, a local estimate of the variance is also needed. In Cleveland's paper, such an estimate was not provided. The present method is therefore briefly reviewed in Appendix A. Although the derivation and notation is slightly different from that of Cleveland, the results are the same, except for the estimation of the variance of a local regression error. In applying the method, two additional modifications are made with respect to Cleveland's study: First, a local window of fixed length is choosen as opposed to a window with length varying according to the distance to the k-nearest neighbor. In application to magnitude conversion, a fixed length is preferred because of the grouping of the data and because it allows to control more easily the influence of low size measures on the regression at high size measures. For a window with variable length, this is difficult because the number of earthquakes reported in this range can be very small. Second, a normal density is used for the weighting function

that defines the local data, as opposed to the trisquare function used by Cleveland (see Equation A.25 in Appendix A.2). The difference should be small and the interpretation of the controlling parameter h is in our case simpler: h is the "standard deviation" of the weighting function and the interval of non-zero weights is restricted to 4h on each side of the estimation point (Figure 2.5).

2.2.2 Linear Splines

A linear spline is simply a piecewise linear function; in regression, it is used to model linear dependence over disjoint intervals of the predictor variable and to impose continuity of the regression at the boundaries of the intervals, so-called knotpoints. A convenient parametrization of the linear spline is in terms of the changes in slope, β_k, at each of the K knotpoints. In this case, the regression $E[y|x]$ has form

$$E[y|x] = \sum_{k=0}^{m(x)} \beta_k (x-t_k) \qquad (2.1)$$

where m(x) is the maximum index m for which $t_m \leq x$. Note also that an additional parameter β_0 is introduced to estimate the intercept at the first knotpoint t_1. The corresponding knotpoint t_0 can be chosen any number smaller than t_1. To allow for heteroscedasticity of the regression error, $\sigma^2_{y|x}$ is assumed constant in each segment (t_k, t_{k+1}) and denoted by σ^2_k. Since the regression is linear in the parameters β_k, estimation of β_k is a straight-forward application of weighted least-squares (Montgomerey and Peck, 1982). The selection of the number and position of the knotpoints is not such an easy problem. A formal approach would be to

base the selection of the knotpoints on the improvement of a goodness-of-fit statistic, such as the chi-square statistic, or to test the significance of the change in slope at the knotpoints. As will be illustrated in Section 2.2.3, a less formal formal approach is used here, based on visual inspection of the fitted regressions and on these statistics.

2.2.3 Examples

Both the RLWLS-method and the linear splines are applied to data from the Chiburis catalog that have both magnitude M and Modified Mercalli Intensity I_0 reported. Figures 2.6a-d show RLWLS regressions obtained with increasing window sizes h. Uncertainty about the regression and on the estimated regression is indicated through ± one-standard-deviation bands. The estimate of the standard deviation of the regression error is constrained to be larger than 0.3. This constraint is active for high valules of I_0 if the window h is small (Figure 2.6a). As explained in Section 2.2.1, the fitted regression is iteratively determined to robustify the estimate with respect to outliers. In each of the figures, the estimated regression is plotted for the first three iterations. Only in the last case (h=5, Figure 2.6d) some difference is noted in the fitted regression, indicating that there is negligible effect of outliers in the present case. Notice that for small values of I_0 some nonlinearity is present also for the largest window h=5. Nonlinearity of the regression of I_0 versus M is demonstrated in Figure 2.7. Also in this case, there is little effect of outliers. The nonlinearity seen in Figures 2.6 and 2.7 is likely attributed to incompleteness of the data, in M and I_0, for earthquakes of small size. The shape of the regression lines when only

data above certain cut-off values of M and I_0 are reported will be studied in the next section.

Whereas RLWLS regression is more useful as an exploratory tool, the linear spline model produces more practical conversion rules. Shown in Figure 2.8a are a simple least-squares fit (a spline with only one knotpoint) and two linear splines with knots at I_0=4 and I_0=5, respectively. Using more knotpoints was found to non-significantly improve the goodness-of-fit. As shown in Figure 2.8a, a knotpoint at I_0=5 gives the best fit and a highly significant change in slope. Comparison with Figure 2.4 shows that for values of I_0 larger than 5, the major part of the data occurred prior to 1960, whereas for values of I_0 smaller than 5 the data occurred since 1960. The large change in slope is therefore attributed to the fact that continuity at I_0 equal to 5 is enforced by the linear spline, while the regression lines for the two subsamples are shifted. In application of the RLWLS method (see Figure 2.6.a) the same shift produces a sharp bend in the regression curve for values of I_0 around 6. Based on this limited analysis, it would therefore appear most appropriate to separate the data prior to and since 1960. Uncertainty on and around the regression is illustrated in Figure 2.8.b for the case when a knotpoint is used at I_0=4. The influence of this uncertainty on the conversion from I_0 to M will be discussed for this data sample in Section 2.5.

The ability of the linear spline to model nonlinearity of the regression curve is illustrated more clearly in Figure 2.8.c, which shows the regression curve of bodywave magnitude m_b on the natural logarithm of felt area, ℓnFA. Data for this regression are taken from a catalog covering most of the northeastern U.S. (Epri, 1985). Knotpoints were

chosen at ℓnFA equal to 6,10,11 and 12. Uncertainty on the estimated regression is also shown by lines of ± 1 standard deviation.

2.3 CORRECTIONS TO THE REGRESSION FOR MEASUREMENT ERRORS AND INCOMLETENESS

The statistical techniques discussed in the previous section account for nonlinearity of the regression, heteroscedasticity of the error, and the presence of outliers. In this section, two additional problems are addressed:

- Both size measures are typically subject to estimation errors.
- The sample used in estimating the regression in incomplete.

Therefore, in the following discussion distinction is made between size measures subject to estimation error and their true values. Size measures reported in the chosen magnitude scale are denoted as Y and η. Y refers to actual observations, η refers to the corresponding unknown true values. Similarly, X and ξ are used to denote observed and true size measures that need to be converted. Distinction is also made between two samples: the learning sample, which contains all reported pairs {X,Y} and the prediction sample, which contains data with only X reported.

Following problems are to be considered:

1. One may wish to use either directly observed values Y or values η in the remainder of the analysis. If uncertainty on the observed values Y is homogeneous (e.g. measurement errors on η are iid random variables for all earthquakes), estimation of the unknown values η is not necessary. If, on the other hand, the measurement error varies for different observations Y_i, estimation of η is necessary.

2. Because observations X are subject to error, the regression of Y (or η) on X differs from that of Y(or η) on ξ. As a consequence, if the estimation error varies for different earthquakes X_i, this poses a problem in estimating the regression from the learning sample and applying it to the prediction sample. Furthermore, the difference in the regressions $E[Y|\xi]$ and $E[Y|X]$ may depend on the marginal distribution of X, as pointed out by Ganse et al. (1983). This is illustrated in Figure 2.9: The figure at the top shows the regression of Y on ξ for the entire population (the prediction and learning sample together). The figure below illustrates how the regression of Y on X differs for both samples, when X is subject to a homogeneous measurement error. It follows that a correction to the learning-sample regression may be necessary, before applying it to the prediction sample. The difference between the distribution of X in the learning and prediction sample is illustrated in Figure 2.10 for the Chiburis catalog. Here X corresponds to I_0, Y corresponds to M. One may note that the difference in the distribution of I_0 for the two samples is not very large, except at small values of I_0. In addition, it will be shown later in this section that for an exponential marginal distribution of X (here I_0) in both samples, no correction is necessary. In Figure 2.10, the assumption of exponentiality appears to hold approximately for values of $I_0>4$.
3. The learning sample is typically incomplete at low values of Y. For instance, in Figure 2.10 it is evident that data are missing at low values of M. As a consequence, the estimated regression line is a nonlinear one. Nonlinearity of this type should be

corrected when applied to the prediction sample, if one assumes that this sample is complete in Y.

To address these problems, the relations among the different regressions between variables Y, X, η and ξ are studied in this section. First, the influence of measurement errors in the estimation of the regression between two variables that have bivariate normal distributions is briefly reviewed. This case has received considerable attention in the statistical literature (Mandansky, 1959; Kendall and Stuart, 1973; Reilly and Patino-Leal, 1981) and has also led to some controversy, especially in the domain of calibration theory (Aitchinson and Dunsmore, 1975; Levin and Maritz, 1982; Hunter and Lamboy, 1981). Some of the causes for disagreement are briefly indicated in Section 2.3.1 and results are reviewed for the simplest case when the measurement errors have Gaussian distribution with a-priori known variance and the regression error is non-zero. The assumption of bivariate normality of the observed values X and Y, or of the corresponding exact values ξ and η, contradicts the usual assumption that the marginal distribution of a size measure is exponentially distributed. A statistical model that is consistent with this assumption is studied in Section 2.3.2. To assess the influence of incompleteness on the regression, Section 2.3.2 also considers the case when size measures X and Y below the respective cut-off values x_0 or y_0 are not reported. Under this assumption, it is possible to derive the regression in the incomplete data set given that the true regression is linear. These results are helpful in judging whether nonlinearity of the regression may be attributed to incompleteness. The corrections suggested by theoretical analysis to account for incompleteness and estimation errors are summarized in Section 2.3.3. Application to results for the

Chiburis catalog is illustrated in Section 2.5.2, where an additional correction will be made to account for the uncertainty around the regression.

2.3.1 Effect of Measurement Error

A comprehensive review of the influence of measurement errors on linear regression is in Kendall and Stuart (1973). The purpose of the present section is to establish the notation to be used in later sections, to indicate the fundamental problems in considering measurement errors and to summarize results for the relatively simple case when the two size measures are from a bivariate normal distribution and the distribution of the measurement errors is normal with known variance.

Denote by x_i, y_i the measured values of x and y for the i'th datapoint and by ξ_i, η_i the corresponding unknown true values. Assume that the measurement errors u_i and v_i are mutually independent Gaussian variables. Independence between u_i and η_i and between v_i and ξ_i is also assumed. Therefore:

$$x_i = \xi_i + u_i, \text{ where } u_i \sim N(0,\sigma_u^2) \tag{2.2}$$

$$y_i = \eta_i + v_i, \text{ where } v_i \sim N(0,\sigma_v^2) \tag{2.3}$$

Suppose further that ξ_i and η_i are random variables, independently drawn from a population with distribution $f_{\xi\eta}$. For a fixed value of ξ_i, the random variable $\eta_i | \xi_i$ is assumed to have a Gaussian distribution whose mean value is a linear function of ξ_i and whose variance is constant, i.e.,

$$\eta_i | \xi_i \sim N(\beta_0 + \beta_1 \xi_i, \sigma_e^2) \tag{2.4}$$

Equivalently,

$$\eta_i = \beta_0 + \beta_1 \xi_i + e_i, \quad \text{where } e_i \sim N(0, \sigma_e^2) \tag{2.5}$$

In the literature, distinction is made between the case where $\sigma_e^2 = 0$, i.e. η and ξ are functionally related as would be the case in a physical law, and the case where σ_e^2 is unknown, i.e. standard regression applies. In addition one needs to distinguish the case where the error variances σ_u^2 and σ_v^2 are known or must be estimated from the data. In the present application, one would certainly not expect two size measures to be functionally related. In what follows, it is also assumed that σ_u^2 and σ_v^2 are known. For instrumentally recorded values, such information could be derived on basis of the accuracy of the recording instruments, the variability of the records at different sites and the number of reports. For empirical size measures, the distribution of the measurement error should be based on knowledge of the reporting procedures and of the amount of available information.

The objective now is to derive the true regression coefficients β_0 and β_1 from the observed data $\{x,y\}$. In terms of the observed values x_i and y_i, Equation 2.5 is written as:

$$y_i = \beta_0 + \beta_1 x_i + v_i - \beta_1 u_i + e_i \tag{2.6}$$

If $f_{\xi\eta}$ is assumed to be bivariate normal, then ML estimates of the parameters of the distribution of X on Y can be obtained directly by equating the sample and population second order moments (Kendall and Stuart, Vol. 2, pp. 379). Omitting the derivation, following estimates of β_0 and β_1 are found:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{2.7}$$

$$\hat{\beta}_1 = \frac{s_{xy}}{(s_x^2 - \sigma_u^2)} \tag{2.8}$$

where $\bar{x}$ and $\bar{y}$ are the mean values, s_x^2 is the sample variance of x and s_{xy} the covariance for the sample {x,y}.

Equation 2.8 indicates that, due to the measurement error u, the estimate s_{xy}/s_x^2 of β_1 in standard regression is biased. Geometrically, the estimates of β_0 and β_1 correspond to a rotation of the regression line around the sample average point.

Various complications may arise. For instance, the corresponding ML estimate of the variance of the regression error may be less than zero:

$$\hat{\sigma}_e^2 = s_y^2 - \sigma_v^2 - \beta_1(s_x^2 - \sigma_u^2) \tag{2.9}$$

If this is so, one may show that the constrained ML estimate of σ_e^2 is zero, which implies that the relation between η and ξ is estimated to be a functional one. The ML estimates of β_0 and β_1 differ in this case from those in Equations 2.7 and 2.8 and must be found using explicitly the likelihood function (see, Kendall and Stuart, 1973). In the present application it is unlikely to find that two intrinsically different size measures are functionally related, because they measure different properties of the same earthquake. A more reasonable interpretation of the functional relation is that one size measure has been functionally derived from the other when assembling the catalog. In that case, the functionally derived size measures should be eliminated from the prediction sample. Another difficulty arises if one relaxes the

assumption that pairs (ξ_i, η_i) are iid random variables. For example, ξ_i might be sampled from a distribution whose mean μ_i depends on i. In this case, (n-1) additional unknowns are introduced and need to be estimated. One can show that in this case the maximum likelihood solution breaks down, in the sense that the solution is not consistent (i.e. the estimates do not converge to the true values with probability 1 when the sample size $n \to \infty$). A discussion of this problem can be found in Kendall and Stuart (1973). The same problem appears in a somewhat more general form in the calibration of instruments. Here, it is necessary to very carefully specify the experimental conditions under which the calibration data are gathered to decide on the appropriate model for the learning sample: For instance, are any of the variables ξ, η, x or y controlled, or can one assume an a-priori distribution the data are selected from? This complexity and the multiplicity of cases has led to much confusion and controversy in the literature (Hunter and Lamboy, 1981). For our present application, there is little discussion that the size measures ξ and η can be considered as random variables. The assumption that the underlying population $f_{\xi\eta}$ does not dependent on i is of course an approximation and neglects the fact that the recurrence rate may depend on epicentral location and time of occurrence. Whereas this simplification seems justified, the basic assumptions that $f_{\xi\eta}$ corresponds to a bivariate normal distribution and that all x and y are reported, are questionable. Both issues are discussed in the next subsection. The influence of u_i and v_i having a distribution that depends on i is considered in Section 2.3.3.

2.3.2 A Model for Exponential Earthquake Size Measures Observed with Error and Not Reported Below a Cut-off Value

As in Section 2.3.1, the purpose here is to derive the relation between the regression of y against x in the learning sample and the regression parameters β_0 and β_1 for the true size measures. Equations 2.2, 2.3 and 2.4 are still assumed to hold and the error variances σ_u^2 and σ_v^2 are considered known. To complete the model, only the marginal distribution of ξ needs to be specified. In the previous section, this distribution was assumed Gaussian. Such a model is however inconsistent with the usual assumption that the recurrence rate density of earthquakes varies exponentially with the size of the earthquakes. This condition is incorporated here by assuming that

$$\begin{aligned} f_\xi &= b_\xi \exp[-b_\xi(\xi-\xi_0)], && \text{for } \xi \geq \xi_0 \\ &= 0, && \text{for } \xi < \xi_0 \end{aligned} \tag{2.10}$$

Equations 2.4 and 2.10 specify the joint distribution of ξ and η and, together with Equations 2.2 and 2.3, the distribution of x and y. Various implications of this model on the regressions of y against x and of x against y are derived in this section. A graphical illustration of the difference between a bivariate normal distribution and the present model is shown in Figure 2.11. The fact that the two regression lines $E[\xi|\eta]$ and $E[\eta|\xi]$ are parallel will be shown later in this section. To arrive at a more realistic model it will be assumed also that x and y are only reported above cut-off values x_0 and y_0. The use of the theoretical results derived in this section in the estimation of the true regression of η against ξ will be discussed in Section 2.3.3.

Consider first the derivation of the joint distribution of x and y in terms of the parameters of the model. From Eqs. 2.2-2.4 and 2.10 the joint distribution of y, x and ξ can be derived to be

$$f_{x,y,\xi} = \frac{b_\xi}{2\pi(\sigma_e^2+\sigma_v^2)^{1/2}\sigma_u} \exp\left[-b_\xi(\xi-\xi_0) - \frac{(x-\xi)^2}{2\sigma_u^2} - \frac{(y-\beta_0-\beta_1\xi)^2}{2(\sigma_e^2+\sigma_v^2)}\right]$$

$$\text{for } \xi \geq \xi_0$$

$$= 0 \text{ otherwise} \tag{2.11}$$

To obtain the joint distribution of x and y, $f_{x,y,\xi}$ must be integrated with respect to ξ. For this purpose, the exponential in Equation 2.11 can be rewritten as:

$$-\frac{[\xi-f_1(x,y)]^2}{2c} + f_2(x,y) \tag{2.12}$$

$f_1(x,y)$ and c correspond to the mean value and variance of ξ for fixed value of x and y. $f_2(x,y)$ corresponds to the exponential of the joint distribution of x and y. After some algebra, one finds

$$c = \left[\frac{1}{\sigma_u^2} + \frac{\beta_1^2}{\sigma_e^2+\sigma_v^2}\right]^{-1} \tag{2.13}$$

$$f_1(x,y) = c\left[-b_\xi + \frac{x}{\sigma_u^2} + \frac{(y-\beta_0)\beta_1}{\sigma_e^2+\sigma_v^2}\right] \tag{2.14}$$

$$f_2(x,y) = b_\xi\xi_0 - \frac{x^2}{2\sigma_u^2} - \frac{(y-\beta_0)^2}{2(\sigma_e^2+\sigma_v^2)} + \frac{1}{2}c\left[b_\xi - \frac{x}{\sigma_u^2} - \frac{(y-\beta_0)\beta_1}{\sigma_e^2+\sigma_v^2}\right]^2 \tag{2.15}$$

In terms of the above functions, integration of $f_{x,y,\xi}$ with respect to ξ therefore results in following joint distribution of x and y:

$$\int_{\xi_0}^{\infty} f_{x,y,\xi} d\xi = \frac{b_\xi c^{1/2}}{[2\pi(\sigma_e^2+\sigma_v^2)]^{1/2}\sigma_u} \left(1-\Phi\left[\frac{\xi_0-f_1(x,y)}{c^{1/2}}\right]\right) \exp[f_2(x,y)] \quad (2.16)$$

where Φ (u) is the cumulative distribution function of the standardized normal random variable u.

How to derive from Equation 2.27, an analytical expression for the regression of y against x is not obvious. The joint distribution $f_{x,y}$ can be however simplified if one considers that if $f_1(x,y) >> \xi_0$

$$1-\Phi\left[\frac{\xi_0-f_1(x,y)}{c^{1/2}}\right] \approx 1 \quad (2.17)$$

In practice, size measures x and y are only reported above certain values x_0, y_0, the value of which depends on the sensitivity of the reporting devices. Since the cut-off value ξ_0 can be assumed arbitrarily low (ξ can be thought of as a non-observable variable), this simplification can be always justified over the range of observed values x and y. Using then Equation 2.17, it follows from Equation 2.16 that for values of $x > x_0$ and $y > y_0$, the joint distribution of x and y is exponential with parameter $f_2(x,y)$.

Consider next the derivation of the regression of y on x using this simplified joint density function. $f_2(x,y)$ is quadratic in terms of x and y and can therefore be rewritten as

$$-\frac{[y-\beta_0-r(x)]^2}{2c_r} + g(x) \quad (2.18)$$

where r and g are linear functions of x. After some algebra, one finds:

$$f_{x,y} \propto \exp\left[-\frac{(y-\beta_0-\beta_1 x+\beta_1\sigma_u^2 b_\xi)^2}{2(\beta_1^2\sigma_u^2+\sigma_e^2+\sigma_v^2)} - b_\xi x\right] \text{ for } x \geq x_0,\ y \geq y_0,$$

$$= 0 \text{ otherwise} \qquad (2.19)$$

Note that the distribution of Y is a Gaussian one for fixed values of x, with expected value $\beta_0+r(x)$ and variance c_r, and truncated from below at y_0. If y_0 is sufficiently small, i.e. $y_0 << E[y|x]$, then

$$E[y|x] = \beta_0+\beta_1 x-\beta_1\sigma_u^2 b_\xi \qquad (2.20)$$

$$\sigma^2_{y|x} = \beta_1^2\sigma_u^2+\sigma_e^2+\sigma_v^2 \qquad (2.21)$$

In other cases, it is necessary to consider the influence of truncation at y_0. At sufficiently high values for x, Equation 2.20 will of course always apply. Notice that in this case the regression of y against x is parallel to that of y against ξ, but that the expected value of y is lower for x fixed to a given value than for ξ fixed to the same value. Indeed, the expected value of $\xi|x$ is smaller than x, because earthquakes of smaller size are more likely to occur.

Another characteristic, which will be frequently used in the following sections, is the regression of x against y. Also this regression follows from the joint distribution $f_{x,y}$ in Equation 2.19. In this case, the exponential term $f_2(x,y)$ is rewritten as:

$$-\frac{[x-t(y-\beta_0)]^2}{2c_t} + h(y) \qquad (2.22)$$

where t and h are linear functions. Omitting the details of the derivation, one finds:

$$f_{x,y} \propto \exp\left[-\frac{[x-(y-\beta_0)/\beta_1+b_\xi(\sigma_e^2+\sigma_v^2)/\beta_1^2]^2}{2[\sigma_u^2+(\sigma_e^2+\sigma_v^2)/\beta_1^2]} - \frac{b_\xi}{\beta_1} y\right] \text{ for } x \geq x_0, y \geq y_0$$

$$= 0 \text{ otherwise} \qquad (2.23)$$

One should again distinguish between two cases: If the value of the function $t(y-\beta_0)$ is sufficiently far above x_0, then the distribution of x for fixed value of y is Gaussian; otherwise, truncation below x_0 must be considered. In the first case, the regression of x against y is parallel to that of y against x. Notice also that the marginal distribution of y is exponential. The various relations between x, y, ξ and η are summarized in Table 2.1 and illustrated in Figure 2.12 for the case when both x and y are sufficiently far from the treshold values. The effect of these treshold values on both regressions is illustrated in Figure 2.13 for three generic cases:

1. $y_0 << \beta_0 + \beta_1 x_0$
2. $y_0 >> \beta_0 + \beta_1 x_0$
3. $y_0 \approx \beta_0 + \beta_1 x_0$

It is recognized that these cases are ideal approximations of the actual effect of incompleteness on the regression: In reality, there is seldom a sharp truncation point in the distribution of a size measure; rather, one finds a progressive decline of reported values for lower values. (Incompleteness as a function of the size measure will be modelled explicitly in Chapter 4 for the purpose of estimating recurrence rates.)

Of course, such a treshold can be introduced artificially by trimming the data in the prediction sample below a given value. Another simplifying assumption in the above derivation is linearity of the regression of η against ξ.

2.3.3 Proposed Corrections and Examples of Application

Based on the theoretical model studied in Section 2.3.2, corrections to the regression of Y on X from the learning sample can be derived to account for the effect of measurement errors and incompleteness. In summary, following assumptions are made. True size measures ξ and η have marginal exponential distributions and are linearly related as

$$\eta = \beta_0+\beta_1 \xi+e, \quad \text{where } e \sim N(0,\sigma_e^2) \tag{2.24}$$

In the previous section, it was shown that such a linear relation can be satisfied for values of η, sufficiently far from $\beta_0+\beta_1 \xi_0$. Size measures in the learning sample are subject to independent measurement errors u_i, v_i with variance independent of i and normal distribution:

$$x_i = \xi_i+u_i, \quad \text{where } u_i \sim N(0,\sigma_u^2) \tag{2.25}$$

$$y_i = \eta_i+v_i, \quad \text{where } v_i \sim N(0,\sigma_v^2) \tag{2.26}$$

Error terms on the observed size measure x_i in the prediction sample are allowed to have different variances, depending on the earthquake under consideration:

$$x_i = \xi_i+u_i, \quad \text{where } u_i \sim N(0,\sigma_{u_i}^2) \tag{2.27}$$

The problem considered here is to estimate $y|x_i$ for values in the prediction sample. Alternatively, one might want to estimate $\eta|x_i$. Since, it

is assumed that the variance of the error term on η_i is constant, a direct conversion to y is simpler. If instead a conversion to η is necessary, also values of y need to be converted as will be indicated later in this section.

The influence of measurement errors on the prediction of a size measure y when only x is given is twofold: First, one must consider how to estimate the true regression coefficients β_0 and β_1 from learning sample data in terms of x and y. Second, one needs to correct the true regression coefficients to account for the measurement error $u_i=x_i-\xi_i$ in the prediction sample. It was found in Section 2.3.2 that the learning-sample regression of y_i against x_i can be written in terms of the true regression coefficients as

$$y_i = \beta_0+\beta_1(x_i-\sigma_u^2 b_\xi) + e_{y_i}, \quad \text{where } e_{y_i} \sim N(0, \beta_1^2\sigma_u^2+\sigma_e^2+\sigma_v^2) \tag{2.28}$$

In the more general case when σ_u is different for different datapoints in the learning sample, β_0 and β_1 can be estimated using a weighted least squares method for the transformed dataset $\{y_i, x_i'\}$, where

$$x_i' = x_i - \sigma_{u_i}^2 b_\xi \tag{2.29}$$

and using the following weights:

$$w_i = (b_1^2\sigma_{u_i}^2+\sigma_{e_i}^2+\sigma_{v_i}^2)^{-1/2} \tag{2.30}$$

b_1 is used here to indicate the estimated value of β_1. Because the weight w_i depends on the initially unknown values of β_1 and also on σ_e, an iterative scheme should be used. The estimation further requires a prior estimate of the recurrence parameter b_ξ. In many cases the information on the measurement error is not sufficiently detailed to differentiate the

accuracy of different observations and the error on the variables in the learning sample can be assumed to have the same distribution. In this case, the true regression coefficients for ξ and η, β_0 and β_1, are related to the learning-sample coefficients b_0 and b_1 as

$$\beta_0 = b_0 + b_1 \sigma_u^2 b_\xi \tag{2.31}$$

$$\beta_1 = b_1 \tag{2.32}$$

In converting from x_i in the prediction sample to y_i (or η_i), one must introduce one final correction to account for the error u_i with which x_i estimates ξ_i. Notice that, because the prediction sample typically spans a long timer period, the assumption of uniformity of the error variance may not hold. In calculating the expected value of y_i (this is the same as the expected value of η_i) for the i'th datapoint in the prediction sample, the regression needs to be corrected as follows:

$$E[y_i | x_i] = \beta_0 + \beta_1 (x_i - \sigma_{u_i}^2 b_\xi) \tag{2.33}$$

Substituting for the parameters β_0 and β_1 from Eqs. 2.31 and 2.32 leads to a formula in terms of the learning-sample coefficients b_0 and b_1:

$$E[y_i | x_i] = b_0 + b_1 x_i + b_1 b_\xi (\sigma_u^2 - \sigma_{u_i}^2) \tag{2.34}$$

Equation 2.34 says that the estimated regression should be adjusted downwards for those datapoints in the prediction sample that are observed less accurately than those in the learning sample. Conversely, datapoints that are observed more accurately would have a larger predicted mean value of y. No correction is necessary for datapoints with accuracy equal to that in the learning sample. It should be noted that this result is

model-dependent. For instance, a different correction is found by Ganse et al. (1983), for the case where ξ and η have normal, instead of exponential marginal distributions.

Another characteristic of the regression that is influenced by measurement errors and is needed in the conversion of magnitudes (see Section 2.5) is the uncertainty about the regression. To simplify notation, homogeneity of σ_u^2 in the prediction sample is assumed. According to Equation 2.28, the residual variance in the regression of η against ξ is related to the learning-sample residual variance $\sigma_{y|x}^2$ as

$$\sigma_e^2 = \sigma_{y|x}^2 - \beta_1 \sigma_u^2 - \sigma_v^2 \qquad (2.35)$$

If $\sigma_e^2 = 0$, then the size measures η and ξ are functionally rather than statistically related. In fact, the estimate of σ_e^2 may even be negative. An explanation of zero or negative estimates of σ_e^2 is that either some or all of the values of y in the learning sample have been functionally derived from x, or else that, contrary to what is assumed in Equation 2.33, the measurement errors u and v are positively correlated. Under this last condition, Equation 2.35 needs to be corrected to account for the covariance between the error terms u and v. From Equation 2.6,

$$\sigma_e^2 = \sigma_{y|x}^2 - \beta_1^2 \sigma_u^2 - \sigma_v^2 + 2\beta_1 \mathrm{cov}(u,v) \qquad (2.36)$$

Because estimation of the term cov(u,v) is not easy, a more pragmatic approach is suggested: it does not seem plausible that the indirect estimation of η for ξ could be more accurate than a reasonably precise, direct measurement of η. Under this assumption,

$$\sigma_e^2 \geq \sigma_v^2 \tag{2.37}$$

In combination with Equation 2.35, this leads to following heuristic estimate of σ_e^2:

$$\hat{\sigma}_e^2 = \max\{\sigma_{y|x}^2 - \beta_1^2\sigma_u^2 - \sigma_v^2, \sigma_v^2\} \tag{2.38}$$

$\hat{\sigma}_e^2$ is an estimate of the variance of the random variable η for fixed value of ξ. The variance of the predicted value of y_i should add a term to account for the variance on the estimate of the expected value, say $\hat{\sigma}_\beta^2$ (as obtained from the regression analysis), and for the corrections due to measurement error. The final expression is then:

$$\hat{\sigma}_{y|x_i} = \hat{\sigma}_\beta^2 + \hat{\sigma}_e^2 + \hat{\beta}_1^2\,\sigma_{u_i}^2 + \sigma_v^2 \tag{2.39}$$

If instead the true value η is predicted, the variance should be decreased by σ_v^2. In this case, also the direct estimates of y in the catalog should be corrected for measurement error. Under the assumptions of the statistical model in Section 2.3.2, the expected value of $(\eta|y)$ can be expressed in terms of the regression of y against η (in this case $E[y|\eta]=\eta$) and of the variance about this regression, σ_v^2. The result is

$$E[\eta|y_i] = y_i - b_\eta\sigma_v^2 \tag{2.40}$$

Derivation of Equation 2.40 is analogous to that of the conditional mean $E[x|y]$ from $E[y|x]$ from Table 2.1. In Section 2.5, it will be shown that for the magnitude conversion (as used in the clustering analysis and for the estimation of recurrence rates), neither $E[y|x]$ from $E^{-1}[x|y]$ is a good estimator of y. Similarly, neither the reported value y nor $E[\eta|y]$

should be used in the case when y is reported directly. The details of this additional correction will be explained in Section 2.5.

Equation 2.20 in the previous section gives the regression of y against x when both size measures are sufficiently far above their respective cut-off values y_0 and x_0. When the regression is close to the cut-off value of y, the learning-sample regression $E[y|x]$ is nonlinear in x, as illustrated in Figure 2.13. This type of nonlinearity is of course induced by the incompleteness in y of the learning sample. If the prediction sample is considered complete in y (a reasonable assumption since y is not reported), then the nonlinearity of the learning-sample should be corrected. One approach to this problem is to apply the results of the statistical model studied in Section 2.3.2 rigorously. For instance, the analytical expression for the joint distribution of $f_{x,y}$ could be used in a likelihood formulation to derive estimates of β_0 and β_1: For fixed value of x,y has a truncated normal distribution and, thus, the mean value and standard deviation are nonlinear in β_1 and β_0 (Johnson and Kotz, 1970). The expression for the mean value of the truncated variable y_t is

$$E[y_t|x]=E[y|x]+\sigma_{y|x}Z\left((y_0-E[y|x])/\sigma_{y|x}\right)/\left[1-\Phi\left((y_0-E[y|x])/\sigma_{y|x}\right)\right] \quad (2.41)$$

where Z and Φ are the standard normal density and cumulative distribution function, respectively. $E[y|x]$ and $\sigma_{y|x}$ refer to a non-truncated variable y, for which Equation 2.34 and 2.39 might be used.

The problem is even more complicated if one considers the possibility of true nonlinarity of the regression of η against ξ and progressive incompleteness of the sample in terms of y. Because of these difficulties, no attempt is made to incorporate incompleteness explicitly

in the estimation of the regression. Rather it is proposed to use the methods of Section 2.2 to derive the apparent regression of y against x and to compare its nonlinearity (if any) with the types of nonlinearity shown in Figure 2.13. If, based on this comparison, the nonlinearity at low values of x can be attributed to incompleteness, then nonlinearity can be eliminated by extending backwards the next linear segment in the fitted linear spline.

Strictly speaking the corrections discussed in this section only apply to the case where the conditional variable $\eta|\xi$ has a mean value that is a linear function of ξ. A theoretical treatment of the case where this condition is violated is complicated. Notice, for instance, that for exponential marginal distribution of ξ and normal distribution of the conditional random variable $\eta|\xi$, the resulting distribution of η is not exponential; which variable η or ξ should then be assumed to have exponential marginal distribution? On the other hand, it is reasonable to expect that the preceding corrections remain valid if the regression of $y|x$ is locally linear within a few standard deviations of the regression error, and if a local (with respect to ξ) estimate of the slope parameter b_ξ is used in the corrections.

A practical example of the previous corrections for measurement errors is shown in Section 2.5.2 for the Chiburis data.

2.4 ESTIMATION OF THE REGRESSION WHEN SEVERAL SIZE MEASURES ARE AVAILABLE

If in the catalog more than two size measures are used, the problem may occur of having to estimate the value of y given a vector $\underline{x}$ of other size measures. Because the number of earthquakes for which both y and $\underline{x}$

are reported is typically very small and because of possible nonlinearity of the regression, direct estimation of a multiple regression is practically impossible and approximate procedures must be considered. A natural choice for such an approximation is to use a combination of regressions of y against individual components of $\underline{x}$. This approach is further explored in the present section.

Consider first a single size measure x_i. If the assumptions of the statistical model in Section 2.3.2 hold, then

$$y|x_i \sim N(\beta_{0,i} + \beta_{1,i}x_i, \sigma^2_{e,i}) \tag{2.42}$$

$$x_i|y \sim N\left(\frac{y - \beta_{0,i} - b_y\sigma^2_{e,i}}{\beta_{1,i}}, \frac{\sigma^2_{e,i}}{\beta^2_{1,i}}\right) \tag{2.43}$$

Subscript i in the above equations refers to the size measure and not, as in the previous section, to a particular datapoint. Next, assume that the conditional random variables $x_i|y$ are mutually independent for different i. Consider then an estimator of y, say p, that is linear in the observed variables x_i:

$$p = \sum_{i=1}^{k} w_i x_i \tag{2.44}$$

Because of the assumption of independence and normality of $x_i|y$, the distribution of p for fixed value of y is

$$p|y \propto N\left(\sum_{i=1}^{k} w_i \frac{y-\beta_{0,i}-b_y\sigma^2_{e,i}}{\beta_{1,i}}, \sum_{i=1}^{k} w_i^2 \frac{\sigma^2_{e,i}}{\beta^2_{1,i}}\right) \tag{2.45}$$

Because the distribution of $p|y$ is Gaussian with $E[p|y]$ a linear function of y and $var[p|y]$ independent of y, the joint distribution of p and y must satisfy the conditions of the statistical model in Section 2.3.2 and, hence, the results derived in that section apply. In particular, the regression of $E[p|y]$ is parallel to $E[y|p]$, and $E[y|p]$ is linear in p. To obtain an unbiased estimator of y in terms of p, the slope and intercept of $E[y|p]$ must be evaluated. To facilitate further calculations, it is of interest to restrict weights w_i such that this slope is one

$$\sum_{i=1}^{k} \frac{w_i}{\beta_{1,i}} = 1 \tag{2.46}$$

Under this condition, the variances of both conditional random variables $p|y$ and $y|p$ are the same, with value

$$Var(p|y) = Var(y|p) = \sum_{i=1}^{k} w_i^2 \frac{\sigma_{e,i}^2}{\beta_{1,i}^2} \tag{2.47}$$

It is easy to derive weights w_i that minimize the variance of $y|p$ under the condition of Equation 2.46. Omitting the derivation, one finds

$$w_i = T \frac{\beta_{1,i}}{\sigma_{e,i}^2} \tag{2.48}$$

$$T = \left(\sum_{i=1}^{k} \frac{1}{\sigma_{e,i}^2} \right)^{-1} \tag{2.49}$$

Using Equations 2.45 to 2.49, the Gaussian distribution of $p|y$ is

$$p|y \propto N\left(y - T \sum_{i=1}^{k} \frac{\beta_{0,i}}{\sigma_{e,i}^2} - b_y nT,\ T\right) \tag{2.50}$$

Using the results of Table 2.1, one easily derives the distribution of $y|p$

$$y|p \propto N\left(p + T \sum_{i=1}^{k} \frac{\beta_{0,i}}{\sigma^2_{e,i}} + b_y nT - b_y T,\ T\right) \tag{2.51}$$

Finally, replacing p with its expression in terms of x_i and using the weights in Equation 2.48

$$E[y|p] = T \sum_{i=1}^{k} \frac{1}{\sigma^2_{e,i}} (\beta_{0,i} + \beta_{1,i} x_i) + b_y T(n-1) \tag{2.52}$$

The first term in the righthand side of Equation 2.52 has the intuitive interpretation of weighted average of predictors of y based on the individual size measures x_i. Each of these predictors has weight inversely proportional to the variance of y for fixed x_i. The second term corrects for the fact that individual regressions $E[y|x_i]$ are not independent.

Although Equation 2.52 is derived under the assumption that the various regressions are linear, the same formula may be used when the regression is estimated as a linear spline, or is locally approximated by a linear function, as in RLWLS. In summary, the following procedure is proposed for the estimation of y when several size measures x_i are available:

1. Estimate the individual regressions $\hat{y}_i(x_i)$ and variances $\hat{\sigma}^2(y|x_i)$ using the methods discussed in Section 2.2 and applying the corrections of Section 2.3, if necessary.

2. Combine the individual estimates using

$$\hat{\sigma}^2(y|\underline{x}) = \left[\sum_{i=1}^{k} \frac{1}{\hat{\sigma}^2(y|x_i)} \right]^{-1} \tag{2.53}$$

$$\hat{y}(\underline{x}) = \hat{\sigma}^2(y|\underline{x}) \sum_{i=1}^{k} \frac{\hat{y}_i(x_i)}{\hat{\sigma}^2(y|x_i)} + (n-1)b_y \hat{\sigma}^2(y|\underline{x}) \tag{2.54}$$

where b_y is an estimate of the slope of the exponential recurrence law for y. If y corresponds to bodywave magnitude, typical values of b_y in the New England region are in the range (1.5,2.0). If y corresponds to Modified Mercalli Intensity, the corresponding range is (0.9,1.2).

3. Compare the individual estimates with the combined estimate and flag significant differences, e.g.

$$\hat{y}(\underline{x}) - \hat{y}_i(x_i) > 3\hat{\sigma}^2(y|x_i) \tag{2.55}$$

Step 3 is added as a safeguard against anomaluous cases when the reported size measures x_i produce inconsistent predictions.

2.5 MAGNITUDE CONVERSION FOR THE ESTIMATION OF RECURRENCE PARAMETERS AND CLUSTER ANALYSIS

In current practice, the estimated regression between two size measures is used directly to convert one size measure into the other. After this conversion, no distinction is made between directly measured and converted values. Such a procedure leads to biased estimates of the recurrence rate, as will be shown in this section, and to an ordering of the earthquakes with respect to size measure that depends on the chosen size measure. Emphasis in this section is on the question of bias which

is of importance in the estimation of recurrence parameters. Invariance of the ordering of the earthquakes is of importance in a cluster analysis of the earthquakes, when only the largest earthquake within each cluster is retained as an independent event.

In the following analysis, it is assumed that all size measures are converted to a single scale m (y or η in Section 2.3.3). If true size measures (η) are used, direct observations of m (e.g. y) should be considered equivalent to measurements in an alternative scale: for instance, the regression of the true value as a function of the observed one has been derived in Section 2.3.3 (Eq. 2.40). The recurrence rate as a function of size measure m is assumed to be of the parametric form:

$$\lambda_m = a \exp(- b_m m), \text{ for } m \geq m_0 \tag{2.56}$$

where a and b are parameters that may vary with location. In the remainder of this section, recurrence parameter b will always refer to the chosen size measure and, therefore, no subscript is used.

Intuitively, one may expect that, because uncertainty about the estimate $E[m|x]$ is neglected in the conversion, the distribution of $E[m|x]$ must be narrower than that in Equation 2.56. Since $m=E[m|x]+\varepsilon$, a simple remedy is to replace the regression estimator with a simulated value m^*, such that

$$m^* = E[m|x] + \varepsilon_m^*, \quad \text{where } \varepsilon_m^* \sim N(0, \sigma^2_{m|x}) \tag{2.57}$$

$\sigma^2_{m|x}$ is the variance of m given x. However, this procedure works well only if the number of earthquakes with value equal to $E[m|x]$ is large. A more satisfactory solution to the problem of magnitude conversion for the

estimation of a and b in Equation 2.56 is given in this section. It is found that in order for $\lambda_{\hat{m}|x}$ to equal to λ_m, one needs to use an estimator of the type

$$\hat{m}|x = E[\hat{m}|x] + \frac{1}{2}\sigma^2_{m|x}b \tag{2.58}$$

The importance of the correction $\frac{1}{2}\sigma^2_{m|y}b$ in Equation 2.58 is then evaluated for the case of the Chiburis data. The influence of incompleteness and grouping of this data is also discussed.

2.5.1 Likelihood Formulations for the Estimation of a and b Parameters

The final objective of the statistical analysis of earthquake data for seismic hazard evaluation is to estimate the recurrence rate as a function of earthquake size and location. The issue considered here is how the estimation of the recurrence parameters a and b is influenced by the fact that, for different earthquakes, different size measures are reported. The following approach is taken to study this problem: A model is formulated for the joint recurrence rate density of all observed size measures. This model is consistent with the marginal recurrence rate of m in Equation 2.56. Various estimators $\hat{m}$ of m from other size measures $\underline{x}$ are then obtained by considering various likelihood approaches. An ideal property of $\hat{m}$ is that $\lambda_{\hat{m}} = \lambda_m$ irrespective of which of the variables (m or any of the x_i's) are reported in the catalog. Only one of the likelihood estimators considered here satisfies this condition. In the case of only one alternative size measure x, this estimator is in the form of Equation 2.58. Notice that $\hat{m}$ in Equation 2.58 depends on the value of b itself. Although the actual value of this parameter is of course not known at the beginning of the analysis, a reasonable initial

estimate is typically available. Alternatively, one could iterate the entire statistical analysis to revise the magnitude conversion. Iteration is not very practical and is also unwarranted considering uncertainty on the modelling assumptions underlying Equation 2.58 and the improvement such iteration could give. Other issues such as the influence of incompleteness and uncertainty about the regression are discussed in Section 2.5.2.

Consider first the case of only one alternative size measure x. In order to obtain the likelihood of an earthquake with only x reported, it is necessary to model first the joint distribution of m and x. The following assumptions are made:

- Earthquakes of different magnitudes occur with exponential rate density

$$\lambda_m = a \exp(-bm) \tag{2.59}$$

- The conditional variable $x|m$ has normal distribution with mean value linear in m and constant variance:

$$x|m \sim N(\gamma_0+\gamma_1 m, \sigma^2_{x|m}) \tag{2.60}$$

Because the above assumptions are consistent with those made in Section 2.3.2, it follows that:

- Values of x occur according to the exponential rate density

$$\lambda_x \propto \exp(-b\beta_1 x), \quad \text{for } x >> x_0 \tag{2.61}$$

- The conditional variable $m|x$ has normal distribution

$$m|x \sim N(\beta_0+\beta_1 x, \sigma^2_{m|x}), \quad \text{for } x >> x_0 \tag{2.62}$$

where

$$\gamma_0 = -\frac{\beta_0}{\beta_1} - b\frac{\sigma^2_{m|x}}{\beta_1} \tag{2.63}$$

$$\gamma_1 = \frac{1}{\beta_1} \tag{2.64}$$

$$\sigma^2_{x|m} = \frac{\sigma^2_{m|x}}{\beta_1^2} \tag{2.65}$$

Equations 2.61 and 2.62 apply for x sufficiently larger than a value x_0 given by

$$x_0 = \gamma_0 + \gamma_1 m_0 \tag{2.66}$$

where m_0 is a lower truncation value for m, e.g. only earthquakes with magnitude m larger than m_0 are assumed to occur. Except for a possible physical bound, m_0 can be assumed arbitrarily low and the observed values of x fall within the range where the above assumptions hold.

A property of special interest is the rate density of earthquakes that are reported in x only, which can be derived by integration of the joint rate density:

$$\lambda_{m,x} = \frac{a}{(2\pi)^{1/2}\sigma_{x|m}} \exp\left(-bm - \frac{(x-\gamma_0-\gamma_1 m)^2}{2\sigma^2_{x|m}}\right) \tag{2.67}$$

Assuming that m_0 is sufficiently low

$$\lambda_x = \int_{m_0}^{\infty} \lambda_{m,x}\, dm = \frac{a}{\gamma_1} \exp\left(\frac{b\gamma_0}{\gamma_1} - \frac{b}{\gamma_1} x + \frac{b^2}{2\gamma_1^2} \sigma^2_{x|m}\right) \tag{2.68}$$

In terms of the regression coefficients β_0 and β_1, this rate density is

$$\lambda_x = \beta_1 \, a \, \exp\left[- b\left(\beta_0 + \beta_1 x + \frac{1}{2} b\sigma^2_{m|x}\right)\right] \tag{2.69}$$

Based on the previous model, various likelihood procedures can now be applied to produce estimators of m given x. One possibility (e.g. Cox and Box, 1964; Plante, 1970) is to use maximum likelihood for the estimation of a and b, and of the unknown value m_i for each earthquake i with only x reported: the ML estimate of m_i must be such that the function

$$L(b,m|x) \propto f_{x|m} \exp(-bm)$$

$$\propto \exp\left(- bm - \frac{(x-\gamma_0-\gamma_1 m)^2}{2\sigma^2_{x|m}}\right) \tag{2.70}$$

is maximum. For a fixed value of b and x, the likelihood is also proportional to the conditional distribution of (m|x). In accordance with Equation 2.62 this distribution is normal, so that the ML estimate m_1^* corresponds to the regression value,

$$m_1^* = \beta_0 + \beta_1 \, x \tag{2.71}$$

As pointed out earlier, the problem in using m_1^* to convert from x to m, is that the resulting estimator of a is biased. The amount of bias can be calculated by deriving the recurrence rate of m_1^*. From Equations 2.69 and 2.71,

$$\lambda_{m_1^*} = \frac{1}{\beta_1} \lambda_x \left(\frac{m_1^* - \beta_0}{\beta_0}\right)$$

$$= a \exp\left(-bm_1^* - \frac{1}{2} b^2 \sigma^2_{m|x}\right) \tag{2.72}$$

This function is exponential, with the same decay parameter b as λ_m, but

with different a. The difference is in the term $-\frac{1}{2} b^2 \sigma^2_{m|x}$, which does not depend on the sample size n and makes the ML estimation of a biased, also asymptoticaly as $n \to \infty$. In the presence of nuisance parameters, there are other well-known cases when ML estimation is asymptotically biased; e.g. Kendall and Stuart (1967), Chapter 29.

As an alternative to maximizing the likelihood, Fraser (1976) and Andrews (1983) among others have proposed maximization of the marginal likelihood function of the parameters of interest (here a and b). This marginal likelihood results from integrating out the nuisance parameters from the total likelihood. Since the total likelihood is proportional to the joint rate density, its integration with respect to the magnitude m_i produces the marginal rate density λ_x in Equation 2.69. Therefore, the marginal likelihood is given by

$$L(a,b|x) \propto \beta_1 \, a \exp\left(-b(\beta_0 + \beta_1 x + \frac{1}{2} b \, \sigma^2_{m|x})\right) \tag{2.73}$$

The same likelihood function is found if the values of m_i are assumed known, with value

$$m_2^* = \beta_0 + \beta_1 x + \frac{1}{2} b \, \sigma^2_{m|x} \tag{2.74}$$

From Equation 2.69, it is easily verified that the recurrence rate of m_2^* is the same as that of m. It is also interesting to note that m_2^* corresponds to the average between the regression of m against x and the regression of x against m. Because the conversion rule depends on the value of b, which is initially unknown, conversion should be applied iteratively. As an approximation, Equation 2.74 could be used with an initial estimate of b.

In summary, for the estimation of recurrence rates and for cluster analysis it is recommended that m_2^* be used, based on the following properties:

- Under the modelling assumptions of Equations 2.59 and 2.60 and if values in x are complete, m_2^* is the only estimator with recurrence rate density equal to λ_m. The estimator m_1^* underpredicts the recurrence rate density. An estimator m_3^*, based on the inverse regression $E[x|m]^{-1}$, would overpredict the recurrence rate density.
- The use of m_2^* corresponds to the ML estimation of the recurrence parameters using a marginal likelihood formulation.
- m_2^* is invariant with respect to the chosen magnitude scale. By invariance it is meant here that the conversion rules from x to m and from m to x are the inverse of the other

$$x_2^*(m) = m_2^{*-1}(m) \tag{2.75}$$

A practical consequence of this property is that the ordering of earthquakes according to converted size measure is independent of the chosen size measure. For the estimators m_1^* and m_3^*, this property does not hold. In fact,

$$x_1^*(m) = m_3^{*-1}(m) \tag{2.76}$$

$$x_3^*(m) = m_1^{*-1}(m) \tag{2.77}$$

2.5.2 Practical Application of the Conversion Rule

In Section 2.3.2, the following regression estimates were derived using data from Chiburis (1981):

$$\hat{E}(m|I_0) = 0.87 + 0.60I_0$$

$$\hat{\sigma}(m|I_0) = 0.60$$

$$\hat{\sigma}_{\hat{E}(m|I_0)} = \left[(0.29)^2 + (0.060)^2 I_0^2 - 2(0.976)(0.060)I_0\right]^{1/2}$$

These regression estimates apply for $I_0 > 4$. Because nonlinearity of the regression for lower values of I_0 appears to be attributable to incompleteness, it is appropriate to use these estimates also for lower values of I_0. In the Chiburis catalog, no indication of the measurement error on M is given; therefore, one may choose to convert to M, rather than to the true value of M, and no correction is necessary for direct estimates of M. On the other hand, the catalog provides interval estimates of I_0 of the type $[I_{0_{min}}, I_{0_{max}}]$. The width of the estimation interval, $\Delta I_0 = I_{0_{max}} - I_{0_{min}}$, varies from 0 to 2. All previous regression estimates were obtained by taking $I_0 = \frac{1}{2}(I_{0_{min}}+I_{0_{max}})$ in the learning sample. Notice that I_0 is the observed value (corresponding to x, in Section 2.3.3) rather than the expected true value (corresponding to ξ). In this sample the number of points with large ΔI_0 is small and a sensitivity study showed little difference in the regression estimates. The situation is somewhat different in the prediction sample: 1,184 data have ΔI_0=0, 164 data have ΔI_0=1 and 17 datapoints have ΔI_0=2. Most of the imprecisely defined intensities have larger values of I_0.

As explained in Section 2.3.3, the regression must be corrected for data in the prediction sample whose estimation error is larger than that

in the learning sample. For this correction one should use Equation 2.34, which required an initial estimate of the recurrence slope parameter b_{I_0}. For the New England region, a reasonable estimate of b_{I_0} is 1.1. In the M-scale, this corresponds approximately to 1.1/0.6=1.83. In addition, one needs to assign standard deviations to the measurement error on I_0 for the various cases. One may note that, because I_0 is discrete, the assumption of a continuous normal error is only an approximation. If one assumes that:

$$\text{for } \Delta I_0 = 0,\ \sigma_u = 0.25$$

$$\text{for } \Delta I_0 = 1,\ \sigma_u = 0.50$$

$$\text{for } \Delta I_0 = 2,\ \sigma_u = 1.00$$

then Equation 2.34 leads to the following corrected regression lines:

$$\hat{E}(m|I_0, \Delta I_0) = 0.87 + 0.60\ I_0$$

$$\hat{E}(m|I_0, \Delta I_1) = 0.75 + 0.60\ I_0$$

$$\hat{E}(m|I_0, \Delta I_2) = 0.25 + 0.60\ I_0$$

To calculate the variance of the predicted value, Equation 2.38 needs to be used. Since the uncertainty on the regression estimate $\hat{\sigma}_{\hat{E}(m|I_0)}$ varies with I_0, also the uncertainty about the predicted value $\hat{\sigma}_{\hat{m}}$ should vary with I_0. Examination of Figure 2.8 shows on the other hand that for high values of I_0, the variance about the regression is possibly overestimated by imposing homoscedasticity above I_0=4. Considering also the sparsity of the data, it appears reasonable to assume that, for ΔI_0=0,

the standard deviation of the predicted value of M is 0.60. For different values of ΔI_0, this estimate must be corrected by a term $\beta_1^2(\sigma^2_{u,\Delta I_0} - \sigma^2_{u,0})$. Therefore

$$\text{for } \Delta_{I_0} = 0, \ \sigma_{\hat{m}|I_0} = 0.60$$

$$\text{for } \Delta_{I_0} = 1, \ \sigma_{\hat{m}|I_0} = 0.65$$

$$\text{for } \Delta_{I_0} = 2, \ \sigma_{\hat{m}|I_0} = 0.84$$

Finally consider the correction to the regression to account for bias on the estimates of recurrence rate density as a function of M, as explained previously in this section. Applying Equation 2.74, the estimates should be increased as follows:

$$\text{for } \Delta_{I_0} = 0, \ m_2^* = \hat{E}(m|I_0, \Delta I_0 = 0) + \frac{1}{2}\frac{1.1}{0.6}(0.60)^2 = 1.20 + 0.60 I_0$$

$$\text{for } \Delta_{I_0} = 1, \ m_2^* = \hat{E}(m|I_1, \Delta I_0 = 1) + \frac{1}{2}\frac{1.1}{0.6}(0.65)^2 = 1.14 + 0.60 I_0$$

$$\text{for } \Delta_{I_0} = 2, \ m_2^* = \hat{E}(m|I_2, \Delta I_0 = 2) + \frac{1}{2}\frac{1.1}{0.6}(0.84)^2 = 0.90 + 0.60 I_0$$

The final distribution of m_2^* is shown in Figure 2.14. The number of datapoints with M reported is indicated for each 0.1 magnitude interval. The number of datapoints with only I_0 reported is indicated separately for each catagory ΔI_0 and on the same scale using the above conversion rules.

From Figure 2.14 it is clear that earthquakes are incompletely reported for small valus of I_0 and, therefore, the observed recurrence rate is non-exponential. How this incompleteness can be modelled as a function of I_0 will be discussed in Chapter 4. Here, the influence of incompleteness on the conclusions of the previous section are of concern.

If one assumes that earthquakes in the prediction sample are selected from an underlying population that satisfies the modelling assumptions in that section and, to account for incompleteness, values I_0 are reported with probability $p(I_0)$, then the marginal likelihood in Equation 2.73 can be written as

$$L(a,b|I_0) \propto \beta_1 \, a \, p(I_0) \exp(-m_2^*) \tag{2.78}$$

It follows that estimator m_2^* still corresponds to using a marginal likelihood approach. The above likelihood function only contains terms for fixed sample size and, therefore, it appears that ML estimates of a and b do not depend on $p(I_0)$. If one considers also the likelihood of N events being reported in I_0, dependence of a and b on $p(I_0)$ is clear (see Chapter 4 for details).

Another feature of the present data, which has not been discussed so far, is discreteness of the I_0 scale. After conversion of I_0 to m, the grouping of the data indicates a natural choice for discretizing the converted scale m^*, see Figure 2.14. Notice that the net effect of the correction term $\frac{1}{2} b^2 \sigma^2_{m|I_0}$ proposed in this section for the regression of M against I_0 is to shift these discretizing intervals with respect to the data. How this grouping of the data affects the estimation of the recurrence parameters is discussed by Bender (1983) and will be considered further in Chapter 4.

Finally, it should be noted that, although a correction is made that accounts for the effect of uncertainty around the regression in the estimation of the recurrence rate, the conversion rule remains a deterministic one. One consequence is that the variance of the parameters

estimated on basis of the converted sample is underrated. Another consequence is that data with different uncertainty on their size measure are treated equally in the remainder of the analysis. Treating instead the uncertainty on the predicted values explicitly in the likelihood formulation will be also considered further in Chapter 4.

Chapter 3

CLUSTERING OF EARTHQUAKES

3.1 INTRODUCTION

Sequences of earthquake events are known to display non-Poissonian patterns, mainly in the form of clusters of short duration and over small regions in space. When the events within a cluster can be causally or at least physically related to a parent earthquake, one refers to these events as foreshocks or aftershocks, depending on their time of occurrence. Other anomalies, for example swarms and longer-term variations of seismic activity, are more difficult to explain through direct causal relationships among the associated earthquakes.

Different stochastic models should be used to describe causal and non-causal dependencies among earthquakes: for example, self-exciting, clustering, and branching point processes are appropriate in the former case, doubly-stochastic processes in the latter. In the case of doubly-stochastic processes, clustering is attributed to random variations of the intensity of the process, but no distinction is made between main and dependent events.

Models of either type have been proposed and fitted to earthquake sequences by Vere-Jones (1970) and Kagan and Knopoff (1976,1978) among others, and used for seismic hazard calculation by Wally (1976) and Merz and Cornell (1973). In particular, Wally represents the earthquake sequence as a doubly-stochastic Poisson process, whereas Merz and Cornell work with a clustering model of parent and offspring events of the Neymann-Scott type.

As Matérn (1960, Chapter 3) points out, it is difficult and sometimes impossible to infer the correct form of a cluster-producing process on the basis of its realizations. Fortunately, seismic hazard is insensitive to the choice of the model among those that are compatible with the data. Therefore, if physical interpretation about the causative mechanism is not of concern, one may choose a cluster-producing model based on mathematical convenience.

The objective of the present chapter is to develop a procedure to classify earthquakes as either main events or secondary events in the context of a generalized Neymann-Scott representation. The adjective "secondary" is preferred to "dependent" as a less specific qualifier; it includes foreshocks and aftershocks, earthquakes in swarms, and possibly events that have occurred at a time and place for which reporting is unusually complete. Main events are defined as the largest earthquakes of their clusters and are assumed to occur as Poisson points in (longitude, latitude, time, magnitude)-space. The intensity μ of this Poisson process may vary on the geographical plane due to nonhomogeneity of the earthquake sources, in time due to incompleteness of the catalog, and in magnitude due to nonuniformity of the distribution of size. The variation of μ with magnitude and time may further depend on geographical location. The only condition imposed is that μ varies in time at a scale larger than that of clustering, so that non-Poissonian groups of events display enough contrast against a relatively slowly-varying background activity. The procedure proposed in this chapter can then be regarded as a filtering process that eliminates high frequency components of the variation of μ.

After their identification, clusters are analyzed to determine several statistics of interest, such as the distribution of the number, space, time, and magnitude of the dependent events for each given size of the main earthquake. This statistical analysis of the clusters and the estimation of the magnitude-recurrence relationship for main events (see Chapter 4) complete the fitting of the Neymann-Scott model.

Existing methods for the analysis of earthquake clusters are reviewed in Sec. 3.2. Their main limitation is that they do not work well when the space, time, and size characteristics of the clusters vary considerably for different main events. Most of the procedures also assume spatial homogeneity or stationarity in time and therefore perform poorly when applied to non-homogeneous or non-stationary catalogs. The method proposed in Sec. 3.3 can be used under all these circumstances. It is applied in Sec. 3.4 to three catalogs: two simulated Poisson catalogs, one nonhomogeneous and stationary and the other nonhomogeneous and nonstationary (in both cases, an ideal procedure would classify all earthquakes as main events), and the Chiburis catalog. The latter catalog contains a classification of earthquakes by seismologists as either main or secondary events. Therefore, this catalog allows one to compare the performance of a judgemental procedure with the present procedure. A sensitivity analysis when applied to this catalog further indicates that the procedure is reasonably robust with respect to variations in the input parameters.

3.2 REVIEW OF EXISTING METHODS

A striking feature of most earthquake sequences is the diversity of clustering patterns: even for main events of the same size with

epicenters in the same general area, the number and space-time-magnitude distribution of secondary events may be very different. Examples of this variation are plentiful in the literature; see for example the many contributions on seismicity patterns in Simpson and Richards (1981), and various papers by Utsu (1961,1969,1970,1971) and Vere-Jones et al. (1964,1965). This means that methods for the identification of secondary events should be flexible with respect to the spatial and temporal structure of individual clusters.

Available procedures can be classified into two groups, depending on whether their primary objective is to fit a point process to data or to classify earthquakes as main or secondary events. In the former case, certain assumptions must be made a priori about the statistical characteristics of the point process, for example about the form of the probability distribution of the number of events in each cluster and about their location in space and time relative to the main event. The need for such assumptions and sometimes the difficulty of parameter estimation are the main drawbacks of direct model-fitting procedures.

Methods of the second type achieve the same objective in two steps: first, they partition the catalog into clusters by using some type of classification criteria. Second, the sequence of main events and the identified clusters are analyzed statistically. Precise assumptions about the model type and cluster characteristics are in this case postponed until the second step. Of course, if one is interested only in the main events, then one needs not model the clusters. In the case when the non-Poissonian characteristics of a catalog are initially unknown, procedures of the latter type should be preferred to those of the former type.

A wide class of models for direct or indirect fitting results from considering a primary process of independent main events and, superposed, a secondary process of offsprings grouped into clusters. The theoretical properties of several such processes are reviewed in Vere-Jones (1970). An example is the Neymann-Scott process, which in the original form models the distribution of points in time: The sequence of main events is stationary Poisson with parameter μ and the offspring process is defined through the probability distribution of their number in a cluster, N, and by the assumption that, conditional on N, the times ΔT_i between the parent earthquake and the dependent events are iid variables with some cumulative distribution function $\Lambda(\Delta t)$. Assuming that N follows itself a Poisson distribution and that $\Lambda(\Delta t)$ is of the power-law form

$$\Lambda(\Delta t) = \begin{cases} 1-\left(\frac{C}{C+\Delta t}\right)^{\delta}, & \text{if } \Delta t > 0 \\ 0, & \text{otherwise} \end{cases} \qquad (3.1)$$

Vere-Jones (1970) fitted the parameters μ, E[N], $C>0$, and $\delta>0$ to shallow earthquakes in New Zealand by matching second-order characteristics of the data. Different stochastic models (a modified Poisson process with nonzero probability of simultaneous occurrences and a Poisson-Markov process) have been studied by Shlien and Toksöz (1970,1975).

The previous models do not consider the spatial configuration of earthquake clusters or the effect of magnitude on their structure. A more general model which incorporates these features and is amenable to maximum-likelihood estimation is described by Kagan and Knopoff (1976). In their model, the process of main events is stationary Poisson, but not

necessarily homogeneous in space and with intensity μ that depends on magnitude according to the exponential Gutenberg-Richter relation. Events of magnitude M are allowed to trigger offsprings of lower magnitude, say m, at a branching rate λ which may depend on M and (M-m). The spatial location and time of the offsprings is defined by a probability distribution which may itself be a weighted average of different functional forms. The branching nature of the process follows from the fact that offspring events may further trigger events of lower magnitude. Kagan and Knopoff have applied this model to a world-wide catalog (1976) and to several regional catalogs (1978).

Both Vere-Jones (1970) and Aki (1956) review empirically observed properties of aftershocks and notice their implications on stochastic modeling and on the underlying causal mechanism. Most of these "laws" are found to be highly debatable, except for Omori's relationship for the variation in time of the rate of aftershocks. Cases when this relationship does not apply are usually referred to as earthquake swarms. During swarms, the recurrence rate is approximately constant and higher than normal. The recurrence law of aftershocks is typically found to be an exponential function of magnitude, although possibly with decay parameter different from that of the main events.

The previous stochastic models rest on the assumption that the secondary events display some statistical regularity. The exploratory analysis of these regularities through the computation of second-order moments is discussed by Vere-Jones (1978) in time and space and by Kagan and Knopoff (1976) in time, space, and magnitude.

As an alternative to directly fitting a stochastic point process, one may attempt to first classify the historical earthquakes as main and

secondary events. As previously noticed, this may be the first of two steps that eventually lead to the fitting of clustering models. The literature in this area is relatively limited. A very simple method, which is often used in engineering application, consists of classifying as secondary events all the earthquakes that fall inside a given space and time window around another event of larger magnitude. In the application to a Southern California catalog, Gardner and Knopoff (1974) used magnitude-dependent windows with the parameters of Table 3.1. The method removed about 2/3 of the earthquakes, leaving a catalog of main events with reasonably Poisson characteristics.

A different technique to separate main shocks from secondary events is based on the likelihood of occurrence of groups of events under the Poisson assumption. A simple method of this kind is mentioned but considered unsatisfactory by Gardner and Knopoff (1974). That particular method seems however to be based on the assumption that an excessive number of events in a given time interval is indicative of clustering, irrespective of the spatial distribution of the earthquakes.

A more elaborate procedure based on time and space windows has been recently proposed by Prozorov and Dziewonski (1982). For each magnitude range of the main event, the windows are iteratively determined as follows: Initial window sizes are assumed for the first iteration. The catalog is then ordered according to decreasing magnitude and increasing time and is then processed sequentially for the identification of secondary events: when earthquake i is considered, all earthquakes with number $j > i$ that fall into its associated window are tagged as secondary events. Once tagged, earthquakes are no longer considered and hence

multiple branching is not allowed. The same procedure is then repeated for a randomized catalog in which the time of occurrence of each event is generated according to a Poisson process. New space-time windows are then defined by comparing the density in space and time of secondary events around main events in the catalogs. During subsequent iterations, the same procedure is followed, except for using the last estimated windows and for removal of secondary events from the randomized catalog.

Finally, a method that does not belong to either of the two groups disucssed above is recently proposed by Ellis (1984): Instead of explicitly modelling the statistical properties of the clusters, one can also model only the variation of the recurrence rate of main events in time and space and consider clusters to be outliers. For example, one may iteratively compare the estimated recurrence rate of main events with the recurrence rate obtained through a non-parametric, and, thus, more local method and assign robustifying weights to earthquakes that fall into regions where the two rates are very different. Ellis applied this method to estimate the parameters of Omori's law for a long aftershock sequence that occurred after the Haicheng earthquake and which itself contains several imbedded aftershock sequences. For an entire earthquake catalog, however, a simple parametric form of the variation of the recurrence rate of main events in time or space is usually not available and the estimation of the recurrence rates of main events itself is complex (see Chapter 4).

3.3 A LOCAL CLUSTERING ALGORITHM

In the method of Prozorov and Dziewonski, secondary events associated with main shocks of the same magnitude are assumed to have the

same space-time distribution. In this case, a single time-space window for each magnitude is sufficient and one needs not adapt the shape of the cluster to the observed pattern of earthquakes near each main event. This assumption of homogeneity of the clusters is common to all the methods reviewed in the last section.

By contrast, the method described here allows for variations in the clustering pattern from earthquake to earthquake and is robust with respect to nonstationarities induced by catalog incompleteness. These features result from using a strictly local analysis in which each main event is considered by itself and the significance of clustering is tested in the neighborhood of that event. It is recognized that, if indeed the secondary events were generated according to a single space-time distribution, our local method would be suboptimal with respect to a global procedure: in that case, clusters that are not very pronounced might not be significant locally but would still be detected through global analysis.

In actual catalogs, it is rare to find that clusters have the same space-time distribution around earthquakes of the same magnitude. Even then, a local method would be useful as an exploratory tool, to verify that clusters are indeed homogeneous.

The basic algorithm for local analysis is as follows: The original events in the catalog are sorted according to decreasing magnitude and ordered chronologically for each magnitude. Next, each event is considered sequentially to determine whether its neighborhood displays a significant clustering of events of lower or equal magnitude. This is done through a formal statistical test, which compares the number of

earthquakes near the main shock with the number of earthquakes inside an extended neighborhood of the same event. If clustering is significant, then the spatial and temporal extent of the cluster is estimated. All or part of the earthquakes inside the cluster are classified as secondary events and are not considered any further in the analysis. The removal of secondary events modifies the significance of clustering (see later for details) so that, after all the events in the catalog have been examined, the entire procedure is repeated until no additional earthquake is eliminated. The four steps of the method - ordering of the catalog, test of clustering, estimation of cluster boundaries, and identification of secondary events inside each cluster - are examined in more detail in the remainder of this section. Variants of the basic procedure are also discussed.

3.3.1 Ordering of the Catalog

The order in which earthquakes are considered affects the outcome of the analysis. For instance, if events of smaller magnitude were considered first, then large clusters would be broken up into several smaller clusters and the significance of the larger clusters could be destroyed, unless the already identified secondary events would be taken into account. A better way to study the inner structure of large clusters would be to first identify them and then apply the present algorithm once more to each cluster of interest. The ordering in terms of decreasing earthquake size is a logical choice, since increases of seismic activity are often causally related to the occurrence of large earthquakes.

If the magnitude scale is discrete or discretized, then further ordering is necessary within each magnitude category. The chronological order, which is that used by the present algorithm, again favors the causal interpretation of earthquake dependencies. Notice however that secondary earthquakes of smaller magnitude are allowed to preceed a main event. That is, the chronological ordering influences the interpretation of dependencies only among earthquakes of the same size.

3.3.2 Testing the Significance of Local Clustering

Since clustering consists of concentrations of earthquake events in spatial neighborhoods of main events and within relatively short time intervals, the identification of clusters can be based on a comparison between the recurrence rate inside a small space-time window around the main shock and the recurrence rate in an extended neighborhood of the same event. The extended neighborhood must still be sufficiently local that spatial nonhomogeneities of the earthquake process and nonstationarity due to incompleteness are small within that neighborhood.

This procedure differs from that of Prozorov and Dziewonski in two important respects: First, the earthquake counts used in the present test are obtained separately for each event and not summed over all the main earthquakes of a given magnitude class. Second, our procedure does not assume that the presence of earthquakes in the immediate neighborhood of a main shock necessarily implies clustering. Rather, the decision whether or not the neighboring events define a significant cluster is the result of statistical testing. For example, in earlier times when the catalog is very incomplete, just two earthquakes occurring close to each

other may constitute a significant cluster, whereas this may not be the case in more recent times.

The test is performed as follows. Two windows are defined in the neighborhood of the main shock that is being examined: a local space-time window W_l and an extended window W_e. For example, W_l and W_e might be cylinders whose radius (maximum geographical distance from the main shock) is similar and is meant to include a significant fraction of the cluster but whose height (duration) is very different; see Fig. 3.1a. The duration of W_l is decided so that this local window includes the most significant portion of the cluster, whereas the duration of W_e may extend over several decades and is mainly determined by the nonstationarity caused by incompleteness. In all cases, W_e should contain W_l. Further denote by V_l and V_e the volumes in space-time of W_l and W_e and by n_l and n_e the counts of events of magnitude not exceeding that of the main shock in the same windows. If the earthquake process in W_e were stationary and Poisson with intensity parameter μ, then the random counts N N_l and N_e inside the local and extended windows would be Poisson variables with mean values μV_l and μV_e such that $\frac{E[N_l]}{V_l} = \frac{E[N_e]}{V_e}$. We take this, with μ unknown, as our null hypothesis H_0, i.e.

$$H_0: \frac{E[N_l]}{V_l} = \frac{E[N_e]}{V_e} \tag{3.2}$$

and test H_0 against the alternative hypothesis H_l that μ is higher in the local window, i.e.

$$H_1 : \frac{E[N_1]}{V_1} > \frac{E[N_e]}{V_e} \tag{3.3}$$

A uniformly most-powerful test for this case is given in Lehmann (1959, p. 140): Under H_0 and given that $N_e = n_e$, the number of events in the local window, N_1, has binomial distribution with number of trials n_e and probability of success p equal to

$$p = \frac{V_1}{V_e} \tag{3.4}$$

Therefore, the distribution of N_1 given that $N_e = n_e$ is

$$P[N_1=n_1 | N_e=n_e] = \binom{n_e}{n_1} p^{n_1} (1-p)^{n_e-n_1} \tag{3.5}$$

for $n_1=0,1,\ldots,n_e$. At a given significance level α, the rejection level n_1^R for N_1 depends on n_e and is defined as

$$n_1^R = \min\{n : P[N_1>n | N_e=n_e] \leq \alpha\} \tag{3.6}$$

From Eq. 6 it follows that, if H_0 is rejected when $N_1 > n_1^S$, the significance level of the test is less than or equal to α. In order to obtain a test with significance level exactly α, one may use the following randomized rule:

$$\text{if } N_1 > n_1^R, \text{ then reject } H_0$$

$$\text{if } N_1 = n_1^R, \text{ then reject } H_0 \text{ with probability } \gamma, \text{ where} \tag{3.7}$$

$$\gamma = \frac{\alpha - P\left[N_1 > n_1^R \mid N_e = n_e\right]}{P\left[N_1 = n_1^R \mid N_e = n_e\right]}$$

In practice, either of these rules could be applied. Although in later applications Eq. 3.7 is used, Eq. 3.6 has the advantage that results of the testing are uniquely defined, i.e., non-random.

By choosing a small value of α, one is assured that only in a few cases (in fact, in a fraction α of cases) H_0 is rejected when H_0 is true. In order to increase the power of the test one should make W_e as large as possible by extending the window in space as far as homogeneity can be reasonably assumed. As to the extent of W_e in time, a characteristic of nonstationarity due to incompleteness that allows one to extend W_e beyond the range of reasonable stationarity is the fact that the rate of catalog events is usually monotonic. Therefore, the increase of μ after the occurrence of the main event is compensated by the decrease of μ in earlier times. What is important for the test to be valid is that the average value of μ in W_e be (approximately) the same as the value of μ at the time of the main event. Because the argument of balancing μ in the extended window does not apply to events that occurred at the beginning or at the end of the period covered by the catalog, special provisions may be needed near the "boundaries". This problem will be discussed further in Section 3.4.

For a fixed value of the count in the local and extended window, the maximum value of p for which clustering is detected can be calculated from Equation 3.6. For high values of n_e and low values of n_1 a Poisson approximation can be used for the binomial distribution of n_1. Under those conditions, the maximum value of P, P_{max}, for which clustering is detected by the test is defined as:

$$\sum_{n=0}^{n_1} \frac{1}{n!} (P_{max}\, n_e)^{n_1} e^{-P_{max} n_e} = 1 - \alpha \qquad (3.8)$$

Notice that P_{max} is inversely proportional to n_e and, hence, the size of the local window where clustering is found to be significant is inversely proportional to n_e for a fixed value of n_1 and V_e. For instance, if $n_1 = 1$, the test finds all local windows with size $V_1 \leq \ell n(1-\alpha)\, V_e/n_e$ to be significantly clustered. Figure 3.2 shows the value of P_{max}, numerically calculated from Equation 3.6, as a function of n_e and for different values of n_1. Two significance levels, 0.02 and 0.05, are used. From the linearity of these curves on a log P_{max} versus log n_e scale, it follows that the Poisson approximation in Equation 3.8 is accurate, except at very high values of n_1 and low values of n_e. Those figures illustrate how the size of significant local windows increases with 1. increasing significance levels α, 2. increasing local count n_1 and 3. decreasing global count n_e respectively.

It is emphasized that the present test is based on the rather mild assumption of <u>local</u> stationarity and homogeneity of the Poisson process of main events. Also notice that the local window W_1 needs not exactly contain the entire cluster and for this reason may be taken to be the same for all earthquakes of the same magnitude.

A two-dimensional representation of the local and extended windows is made in Fig. 3.1b, where R and Δt denote respectively distance from the epicenter and time since the occurrence of the main event. The same figure illustrates a generalization of the previous test, which is useful in the case of clusters that extend mostly beyond window W_1. In this case, the test as previously described looses power because many cluster events are located in the portion of the extended window outside W_1. To prevent this from happening, one may define a buffer window W_b which, with high confidence, contains most of the cluster. If V_b and n_b denote respectively the volume of and the number of events in W_b, then the previous test is made after replacing V_e with $(V_e - V_b + V_1)$ and n_e with $(n_e - n_b + n_1)$.

Fig. 3.1c illustrates still another concept: For small clusters, it may happen that the local window does not display significant clustering because it is too large. In order to detect these clusters (which, as will be shown in Sec. 3.4, are a considerable fraction of all the clusters), a second test of significance is made in those cases when the first test results in acceptance of H_0. The second test uses a contracted local window W_c, which has the same spatial dimension as the original local window W_1, but extends backwards and forward in time by only a fraction q of the original extent. Values of q of the order of 0.1-0.2 have been found to be appropriate by variation of this parameter and considering the additional amount of clustered events.

Estimation of the shape and extent of the cluster for each main earthquake is a separate task which, for the cases when H_0 is rejected in the first test, is performed as described next.

3.3.3 Estimation of Cluster Shape and Size

The fact that the rate of earthquakes in a neighborhood of a main event significantly exceeds the rate in a larger background does not mean that the cluster is entirely contained in that neighborhood. The next task is to find a connected region near the main earthquake that indeed contains all the secondary events associated with that earthquake. This "cluster region" should be as small as possible in both time and space, in order to prevent the erroneous removal of secondary events far from the main shock and, more important, to avoid confounding among clusters that are close to one another in space and time. The actual identification of secondary events is still another task, which will be considered later in Section 3.3.4.

The cluster region is identified through a sequence of statistical tests, each of the type in Section 3.3.2, performed on extensions of the region already recognized as hosting the cluster. Of course, such a region and the earthquakes it contains are ignored when testing for significance of the extensions.

There are several ways in which the "extension regions" and a stopping rule can be defined. A simple possibility is to consider regions with fixed spatial configuration (a disc of given radius centered at the epicenter of the main event), obtained by partitioning the time axis before and after the initial significant window (Fig. 3.3). If clusters consist mainly of foreshock-aftershock sequences, then extensions backwards in time should probably be of a smaller size than foreward extensions.

The significance level α for extending the cluster region needs not be the same as in the initial test: if a cluster is known to exist, then one may want to extend it further on the basis of less evidence of increased seismic activity. In this case one should use larger values of α during the extension process. Extension in either direction terminates when the last region considered in that direction passes the homogeneity test of Section 3.3.2.

In the case of extensions only along the time axis, the radius of the disc in space must be relatively large, so that there is high confidence that the cluster is all contained into the "cluster region". A more satisfactory but also more complicated procedure is to consider extensions according to a rectangular grid in the distance-time reference of the main event (Fig. 3.4). Because space is compressed into a single distance axis, each extension region has in this case an annular form around the epicenter of the main event. There are several variants of the 2D procedure, depending on the order in which the various extension regions are tested, on the stopping rule, and on the "postprocessing" of the cluster region. Two schemes are illustrated in Fig. 3.4:

In the first scheme (top figure), regions tested for significant clustering are those with at least one side in common with the region found already to be significant. The procedure terminates when all the candidate extensions are non-significant. This applies forward as well as backward in time. If the final cluster region is multiply connected, as in the case of Fig. 3.4b, then the region is enlarged to include all the non-significant inner cells. Another possibility is to take a cylindrical envelope in space and time (Fig. 3.4c).

In the second scheme, one orders the cells according to increasing geographical distance from the epicenter and to the time elapsed since the main shock (Fig. 3.4d). One then proceeds "row by row". Each "row" is analyzed as in the 1D case, stopping as soon as a non-significant cell is encountered. The procedure terminates when the first cell of the next row is nonsignificant, both forward and backwards in time. The cluster regions obtained by this second method are simply-connected and are contained in the regions identified by the first method. In spite of the more regular shape of the cluster regions, one may still want to simplify their geometry by using cylidrical envelopes. This second method, with cylindrical envelopes, will be used in Sec. 3.4 to obtain numerical results.

Further extension to a three-dimensional scheme in space and time is clearly possible but is considered unnecessary: as will be said in the next section, spatial symmetry of the cluster region does not imply spatial symmetry of the cluster itself about the main event.

Irrespective of the extension scheme (1D, 2D, or 3D), the size of the extension regions should be not too small, in order to prevent that the procedure stops prematurely due to local decreases of the earthquake rate. Another reason why these regions should not be very small is that, when approaching the boundary of the cluster, the rate of earthquakes decreases and so does the power of the test. As mentioned previously and justified in the next section, extending the cluster region somewhat beyond the true cluster boundaries has only a small effect on the events identified as secondary.

3.3.4 Identification of Secondary Events Inside Cluster Regions

The final step of cluster analysis consists of separating main events from secondary events inside an estimated cluster region. Two procedures, one of which has several variants, can be used for this purpose: The simpler method consists of tagging as secondary all the events inside the cluster region. This method has been proposed by many authors but is unsatisfactory in two respects: 1. the boundary of the cluster region must be estimated with accuracy or else several main earthquakes will be misclassified, and 2. the procedure creates regions of no activity in the neighborhood of many events and is therefore incompatible with the assumption of Poisson main earthquakes.

A better approach is to thin the point process in the cluster region. Thinning should be such that the events not tagged as secondary occur at a rate and with a space-time distribution consistent with a homogeneous Poisson process with the intensity of the background. This can be done by simulating a Poisson point process with the target intensity inside the cluster region and by then finding the earthquakes in the catalog that are closest to the simulated ones in a certain metric (nearest-neighbor method). The nearest neighbors are considered to be main earthquakes; all the others are secondary events (see Fig. 3.5). This process is implemented separately for each magnitude range to allow for differential thinning depending on earthquake size. It is clear that, if the cluster region extends beyond the actual cluster, then most of the thinning will occur where the density of points is higher. This is the reason why the actual shape of the cluster does not depend much on the shape and size of the host region, provided that the region includes it.

The distance measure used here to identify nearest neighbors is based on the space and time dimension of the cluster region: if the region has maximum linear dimension D in space and T in time, then the distance d_{ij} between $(\underline{X}_i, t_i)$ and $(\underline{X}_j, t_j)$ is taken to be

$$d_{ij} = \left(\frac{\|\underline{X}_i - \underline{X}_j\|^2}{D^2} + \frac{\|t_i - t_j\|}{T^2}\right)^{1/2} \tag{3.9}$$

A number of variants can be defined, depending on the way in which the simulated process is obtained. Two possibilities that have been experimented with are:

1. The simulated catalog is obtained from the original catalog by <u>locally</u> randomizing the time of occurrence of each earthquake. Simulation is done only once. The original location of the earthquakes is left unchanged so that spatial nonhomogeneity is preserved. This procedure is simple but has the disadvantage that, at least during the first iteration, the simulated catalog has an intensity μ larger than the intensity of the Poisson background of main events. Therefore, clusters have too few earthquakes removed as dependent events. The problem is automatically corrected in the course of subsequent iterations if the clusters are relatively small and frequent (see Sec. 3.3.5), but bias may remain if seismicity in the spatial neighborhood of a main event is dominated by one or very few large clusters.
2. Another possibility is to simulate a separate Poisson catalog inside each cluster region, using the intensity of the local background, $\mu = \frac{n_e - n_b}{V_e - V_b}$. Simulation is actually repeated for each magnitude

range using a size-specific value of μ. This procedure is computationally more expensive than randomization of the historical catalog but has the advantage of being insensitive to large clusters, of being consistent with the test for clustering in Sec. 3.3.2, and of allowing one to easily correct for boundary effects (see Sec. 3.3.3) by increasing the values of μ estimated from time periods that preceed the most recent main events.

If the first method is used, then earthquakes identified as secondary are tagged both in the original and in the simulated catalog and are not considered further in the analysis. In the second method, tagging is done only in the original catalog.

Results from both methods will be presented in Sec. 3.4. Method 2 leads in general to removal of more secondary events than Method 1. In fact, in regions of moderate or low seismicity, Method 2 produces earthquake classifications that are similar to those from labeling as secondary all the events within the cluster regions.

3.3.5 Subsequent Iterations

Irrespective of the method used to thin the point process inside each cluster region, during the first application of the algorithm one is bound to underestimate the number of secondary events. This is because the background of each main event contains a mixture of main shocks and secondary earth quakes, with the consequence that the weaker clusters may not be significant. In addition, the simulated point processes used for thinning have too high intensity and therefore leave a too large fraction of main shocks inside the cluster regions. Iteration is a simple way to

remove this bias because events that are tagged as secondary are neglected in subsequent analysis. Somewhat different results are obtained depending on the way iterated analysis is implemented. Three alternatives are as follows:

1. In each iteration, the catalog that remains after removing the previously tagged earthquakes is examined in exactly the same way as in the first iteration;
2. Same as 1, except that when testing significance of clustering, the counts in the local and candidate extension windows are based on the complete catalog; excluding however, earthquakes that in previous iteration are identified as secondary to a <u>different</u> main shock.
3. Same as 2, but secondary events are tagged starting from scratch, without consideration of the tagging during previous iterations.

Among these methods, those favored are Methods 2 and 3. Method 2 has the advantage over Method 3 that convergence is easier to check (no additional earthquake recognized as secondary during one iteration). Method 1 has the undesirable feature that, when testing for significance of clustering after the first iteration, the power of the test is low for main shocks that had been associated with clusters during previous iterations. In spite of the conceptual differences among the three methods, the final results are similar. Method 2 is the one used in the applications described in the next section.

3.4 NUMERICAL APPLICATIONS

Before the method proposed in Sec. 3.3 can be reliably used, it should be tested with catalogs of known characteristics. Some testing of this type is made here by applying the method to three catalogs: one is

the Weston Observatory Catalog (Chiburis, 1981) updated to 1980 (Barosh, 1981 personal communication). The catalog contains 3022 events which occurred between 1534 and 1980 in a geographical region that extends in approximation from 63°W to 85°W and from 34°N to 50°N. Earthquakes in the catalog have already been classified as main and dependent events. Although this classification is likely the result of a composite process, it still provides a reference for the proposed automated method. It also gives us an opportunity to verify the consistency of judgemental methods of cluster analysis.

A second catalog has been obtained from the previous one by randomizing the time of occurrence of each event over the entire time interval from 1534 to 1980. Therefore, this catalog is Poissonian and stationary in time, but has the same nonhomogeneity in space as the original catalog.

A third catalog has been obtained by locally randomizing the time of occurrence of the historical earthquakes. Specifically, the times have been simulated as independent variables with uniform distribution inside intervals centered at the associated historical times t_j. The width of the simulation interval has been taken to be a function of t_j and I_o, according to Table 3.2. Truncation of the distributions has been imposed so that all simulated values are between 1534 and 1980. Compared with the historical catalog, this last catalog displays a smoother variation of seismic activity in time, while preserving the long-term nonstationarity due to incompleteness and, of course, spatial nonhomogeneity.

In all three cases, the analysis has been made in terms of epicentral Modified Mercalli Intensity I_o instead of magnitude m. For

events with no reported epicentral intensity, I_o is estimated using the deterministic conversion (Chiburis, 1981)

$$I_o = (m-1)/0.6 \qquad (3.10)$$

rounded off to the closest integer. Although this conversion rule is slightly different from that proposed in Section 2.5.2 and does not consider uncertainty on the conversion or on the reported values of I_o (for a detailed discussion, see Chapter 2), the results of the present clustering method would differ litte under reasonable variations of Eq. 3.10. After elimination of earthquakes with calculated intensity less than 1, each catalog contains a total of 2860 events. A plot of the events according to the original catalog is shown in Fig. 3.5.

3.4.1 Simulated Catalogs

The stationary catalog has been analyzed using the input parameters of Table 3.3. Notice in particular the sizes of the local and extended windows for the test of clustering, the value 0.1 of the factor q that defines the contracted window W_c, and the number of allowed extensions in space (2) and backward and foreward in time (4). The extension method chosen here and in all subsequent numerical calculations is that illustrated in Figs. 3.4d and 3.4e, with a cylindrical envelope. The buffer window W_b (Fig. 3.1b) is chosen as the largest cluster region allowed by the analysis; for example, in the case of I_o=4, W_b extends from (60x4) = 240 days before to (200x4) = 800 days after the main event and has a radius of (0.22x2) = 0.44 degrees. Two iterations are allowed, using Method 2 in Sec. 3.3.5.

A summary of results is given in Table 3.4, in terms of the number of main and secondary earthquakes and of main events with associated clusters. What is perhaps most interesting to consider in the case of a Poisson catalog is the fraction of main events that the algorithm associates with clusters. This fraction, denoted by η_{CLUS}, is given by

$$\eta_{CLUS} = \frac{\text{No. of clusters}}{\text{No. of main events}}$$

$$= \frac{52}{2860}$$

$$= 0.018 \qquad (3.10)$$

The fact that η_{CLUS} is very close to α=0.02 indicates that the present procedure does not confound clustering with spatial nonhomogeneity of seismicity.

In only two cases did the procedure find the contracted window W_c to be significantly clustered when W_1 was not. W_1 was itself extended in one case in time and in three cases in space. The small number of extensions is easily explained by the fact that "Poisson clusters" are small and local, especially if the intensity of the process is low; this is also demonstrated by the small average number of secondary events per cluster, which is 67/52=1.3.

The same parameters have been used in the analysis of the nonstationary catalog, except that ΔR has been doubled and n_R set to 1 for all I_o. Results in Table 3.5 indicate that the fraction of main events associated with clusters has increased to $\eta_{CLUS} = \frac{111}{2630} = 4.2\%$ and that the average cluster size has increased to $\frac{230}{111} = 2.1$. The main reason for these increases is that local randomization of the occurrence

times does not entirely eliminate the high-frequency variations of the earthquake rate, which the algorithm interprets as clusters. For the nonstationary randomized catalog and for the Weston Observatory Catalog (see next section), a correction is used for boundary effects in the estimation of the background rate. The correction consists of taking the maximum between the average background rate and the average of the rates for the portions of background that preceed and follow the main event.

3.4.2 Weston Observatory Catalog

The previous runs indicate that, for Poisson catalogs, the procedure of Sec. 3.3 classifies as secondary only a small fraction of the events. This is true also in the presence of nonhomogeneity in space and nonstationarity in time, of the type caused by incompleteness.

In actual catalogs, clusters are quite diverse in their time-space configuration; whereas some have a duration of only a few weeks or months, others may extend over several years. In order to properly identify clusters of different shape and size, one must allow for a large number of extensions of the initial test window W_1. On the other hand, one should avoid unnecessarily large buffer windows W_b, not to excessively reduce the volume of W_e and thus decrease the power of the clustering test. Because the space-time configuration of the clusters is initially not known, it is good practice to use the procedure twice: the first time one should allow a large number of extensions and obtain a rough estimate of the cluster regions, whereas the second time one should use a number of extensions just sufficient to envelope the largest cluster in each intensity class. In the case of the Weston Observatory Catalog, input parameters for the latter analysis are shown in Table 3.6.

Notice the large number of extensions in time allowed for main events with intensity between 4 and 8. This is the result of having detected, during previous preliminary analyses, large clusters associated with main events of these intensities.

Table 3.7 is analogous to Tables 3.4 and 3.5, except that it includes a breakdown of secondary events according to their classification by the seismologists. Specifically, the last three columns give the number of earthquakes tagged as secondary by our procedure only, by the seismologists only, and by both. The fraction of events that we tag as secondary decreases with increasing I_o, with an average value of 28%. Although the automatic method identifies a larger number of secondary events than the seismologists do, agreement between the two classifications appears to be satisfactory. For example, 91% of the events classified as secondary by the seismologists are also tagged as secondary by our method. As to the earthquakes that only our method detects as secondary, we believe that in many cases they should not be considered as main events (see later in this section).

A breakdown of the secondary events according to their intensity and to the intensity I_o of the main event is given in Table 3.8. These results appear to contradict the relationships proposed by Utsu (1961) and Bath (see Richter, 1958), which give the maximum intensity of aftershocks, I_M, that follow a main event of intensity I_o. According to Utsu, the difference between I_o and I_M (more precisely, between the associated magnitudes) increases with decreasing I_o, whereas according to Bath the difference in magnitudes is constant and equal to 1.2. By contrast, Table 3.8 indicates that, especially for $I_o \leq 5$, there is a

significant probability that $I_M=I_o$. Data is too limited and incomplete to allow one to confirm or disprove the frequent claim that aftershock intensities have truncated exponential distribution, with decay parameter that depends on I_o. Of course, some of our findings may be influenced by the present definition of secondary events.

Other statistics related to extensions in space and extensions and contractions in time are given in Table 3.9. The latter operations are performed each in about 15% of the cases, whereas spatial extension of the cluster region beyond the values of ΔR in Table 6 is made for only 8% of the clusters.

A more direct representation of the results is given through plots: Fig. 3.7 shows the empirical earthquake rate (number of events of any magnitude in one year), separately for the complete catalog, for only the events classified by the present procedure or secondary, and for only the main events. Note the large clusters associated with the 1727 Cape Ann and with the 1976 St. Simeon earthquakes. Also notice how the removal of secondary events smoothes the empirical rate of main earthquakes.

The spatial distributions of secondary and main events are shown in Figs. 3.8a and 3.8b. These partial plots of seismicity should be compared with the combined plot in Fig. 3.6.

A separation of clusters by intensity of the main event is made in Fig. 3.9. For each I_o, two plots are shown using the local reference of the main event in each cluster: the horizontal axis gives the time in days since the main event and the vertical axis gives the squared epicentral distance in degrees. The second plot of each pair contains only the secondary events of the clusters, represented with different

symbols depending on their classification by the seismologists. The first plot displays the same events against the local "background" of main earthquakes of intensity at most I_o. Background events are also plotted with different symbols according to their classification by the seismologists. Symbols are as follows:

Δ - earthquakes classified as secondary by both procedures (present method and seismologists);

□ - earthquakes classified as secondary only by the present method;

◇ - earthquakes classified as main events by both procedures;

o - earthquakes classified as main events only by the present method.

The reason why squared distance is used instead of simply distance is that, for a spatially homogeneous Poisson process, the density of points is constant in the former representation. This facilitates the visual identification of clusters. Because clusters with main events of Intensity 9 and 10 are very few, they are combined in a single plot.

One might find it strange that, in the case of intermediate intensities, the algorithm classifies as secondary events earthquakes that are far away from the main shock and are embedded in a dense background. This apparent contradiction is explained by the fact that the plots of Fig. 3.9 are the result of mixing many different clusters and their neighborhoods. In reality, the intensity of the background varies significantly from cluster to cluster. In order to show this, the most prominent clusters for I_o=5,6 and 7 are plotted in isolation in Fig. 3.10 using again the format of Fig. 3.9. No cluster dominates for I_o=4; therefore, clusters with main intensity equal to 4 have been separated on the basis of size ($n<4$ and $n\geq 4$). It is clear from Fig. 3.10 that each

cluster (each cluster group in the case of I_o=4) is quite distinct from its own background. An extreme case is the cluster of the 1727 Cape Ann earthquake, whose background is empty.

The Cape Ann earthquake can be used also to illustrate the reason why, in the analysis, uncertainty on the geographical location of the historical epicenters has been neglected. If the errors in the determination of the epicenters were mutually independent random variables, then earthquake clusters would appear "blurred" in the catalog. As exemplified by the 1727 Cape Ann cluster, this is not the case, especially for the earlier events. The reason is dependence among the errors: although there is considerable uncertainty on the actual location of the Cape Ann earthquake and its aftershocks, the fact that these events are part of the same cluster has made the seismologists assign the same epicentral coordinates to all. It would be difficult to obtain parameters for a model with dependent errors, and the analysis would become very complicated. In addition, we believe that the final classification of earthquakes with errors modeled would be virtually identical to that with errors neglected.

Oddly enough, the seismologists have not identified as secondary three of the Cape Ann aftershocks and many events within the cluster of Fig. 3.10b: although one could make a variety of assumptions about cause-effect relationships among the earthquakes of Fig. 3.10b, the sparsity of the background makes it difficult to believe that most of these events occurred independently of one another.

With large clusters removed, the plots of Fig. 3.9 would show high concentrations of secondary events very near the origin of the axes, embedded in rather uniform backgrounds.

Table 3.9 and the previous figures give little statistical support to the hypothesis that cluster dimensions in time and space increase systematically with the intensity of the main event. The reason may very well be that the statistical sample is too small. On the other hand, Fig. 3.11 shows some evidence of dependence of cluster dimension on cluster size n. In this case, secondary events are plotted separately for n=1, n=2-6, and n>6. Also the latter dependence should however be interpreted with caution, because it is due in part to the testing procedure, which is unable to detect clusters that contain only very few and widely separated earthquakes.

The plots of Fig. 3.12 show the distribution in space of the secondary events relative to the main earthquakes. The one-tenth-degree accuracy in the reported coordinates produces a grid pattern and obscures somewhat the true space distribution, due to multiple occurrences at some locations. Yet, a NE-SW trend is apparent in the clusters, except for very small and very large values of I_o. This trend is even more evident if one groups clusters according to the number of secondary events, as shown in Fig. 3.13.

3.4.3 Sensitivity Analysis

Eight variants of the input parameters in Table 3.6 have been considered. The variants are described in Table 3.10 and summary results are given in Tables 3.11 and 3.12, respectively for the number of clusters and the number of secondary events. The percentages in the bottom row of Table 3.11 are calculated by dividing the number of clusters by the difference between the total number of earthquakes in the catalog (2860) and the total number of secondary events from Table 3.12.

None of the changes in the input has a significant effect on the final classification of earthquakes, except for halving the space dimension of the local and extension windows, ΔR. The consequent reduction in the number of clusters and secondary events is not unexpected: in the limiting case as $\Delta R \to 0$, the procedure breaks down and no secondary event can be detected. Hence, ΔR should be chosen such that, in the region of clustering, several events are expected to fall inside the local and each of the extension windows. Interestingly, the solution remains almost the same if one doubles the values of ΔR in Table 3.6. The small increase in the number of secondary events is due to the fact that, beyond the cluster regions identified using the parameters of Table 3.6, there is still a modest amount of clustering. This clustering is not significant at the 0.02 level but is removed by increasing the size of the window.

Changing the levels of significance (α for the local window, α_{ext} for the extensions) or the size of the extended windows in space (Case 5) or time (Case 6) has only a minor effect on the classification of earthquakes. Modifying the procedure of earthquake classification inside the cluster region (last two cases) also produces small changes in the results. This is especially true if the new rule is to classify all the events in the cluster regions as secondary (Case 7) and thus to create "holes" in the immediate neighborhood of the main events. The reason for lack of sensitivity is weakness of the background. Tagging earthquakes by Method 1 of Sec. 3.3.4 (last sensitivity case) leads to a reduction in the number of secondary events, as a consequence of the bias described previously in that section.

Overall, sensitivity analysis shows that the proposed method is robust with respect to the input parameters. The only exception is ΔR, which should be chosen to be not much smaller than the expected radius of the clusters. Use of the cluster-extension procedure in Figs. 3.4a-3.4c would reduce sensitivity to this parameter.

3.5 EXPLORATORY ANALYSIS OF THE CLUSTERING RESULTS

Because of the size of the earthquake data set and the many variables involved, such as location, time of occurrence and earthquake size, it is not simple to conduct an exploratory analysis of the clustering results. For this purpose, the displays in Figure 3.14 are found to be useful and will be discussed in this section.

About 93% of the catalog data falls within the region from 38 to 54 degrees North and from 60 to 80 degrees West. To maximize the spatial resolution of the figures, only events inside this region are presented. Furthermore, the time period of the catalog is divided into six intervals, each containing almost the same number of events. For each time period, four plots are produced, showing 1. all events in the catalog, 2. the clusters detected by the algorithm, 3. earthquakes classified as main events by the algorithm, and 4. earthquakes indicated as aftershocks in the original catalog (judgemental aftershocks). Each of the plots shows the spatial distribution of the earthquakes (latitude versus longitude), and latitude and longitude versus time. The size of the symbols is used to indicate the intensity of the events. For the cluster plots, two symbols are used: squares indicate the main event associated with each cluster, crosses indicate the aftershocks. For the

judgemental aftershock plots, squares are used to indicate earthquakes that the present method classified as main events.

The number of events in each of the plots is shown in Table 3.13. Also shown in the table is a comparison of the percentage of secondary events in the catalog according to the judgemental and automatic classifications. Notice that the cluster analysis tends to classify more earthquakes as secondary. The first time period, from 1500 to 1800, contains a relatively large number of aftershocks according to both classifications. Some of these clusters may actually be due to the on-off pattern of reporting, as one can see from Figure 3.14a.1. The large cluster of events following the Cape Ann earthquake of 1727 also partly explains the increased number of aftershocks. During the last two time periods, the present analysis finds a relatively low and a relatively high number of aftershocks, respectively. In part, this may be a consequence of overestimating and underestimating the background recurrence rate in those respective time periods: In the analysis, a time period of 15 years is used for the background window (see Table 3.3), which extends 10 years backwards and 5 years forwards. An asymmetric window has been used to correct for the increased activity in the last time period. However, the counts in Table 3.13 indicate that the yearly recurrence rate over the period from 1974 to 1981 is about three times that from 1960 to 1974. Therefore, results in the last two time periods may be somewhat biased. The percentage of clusters, which is also calculated in Table 3.13, is however remarkably stable over all time periods, except the first one. It follows that the average cluster size during the last time period is substantially larger, presumably due

to increased reporting of events of small magnitude. In the remainder of this section, Figures 3.14a-3.14f will be further discussed with respect to the performance of the clustering algorithm, the pattern of aftershock sequences and the pattern of main shocks.

3.5.1 Performance of the Cluster Analysis

For a visual verification of the clustering procedure, it is of interest to compare the clusters identified by the analysis with the pattern of judgementally identified aftershocks. First, one may note that almost all aftershocks identified in the catalog are also identified by the present method (see the small number of boxes in the plots of aftershocks). Aftershocks not identified by the present analysis are mainly associated with one of the following two effects: 1. the present analysis does not always extend the window over the entire sequence of events, if the sequence is very long or is distributed over a large geographical region; rather, it breaks the sequence into two or more parts (see for example the Cape Ann sequence), 2. in a cluster with events of equal size, the present analysis defines the earlier event as the main shock and the later event as an aftershock. In the catalog classification, this relation is often inversed. Several examples of this type can be seen in the period from 1850 to 1860.

Second, one may note that cluster analysis identifies more aftershocks than the judgemental procedure. From the cluster plots, it is indeed clear that several dependent events have been "missed" in the judgemental classification. For instance, many events that practically coincide in time and space with other events of equal or higher size are not labeled as dependent (e.g., time period 1920-1925).

Another way to judge the performance of the method is to attempt a visual identification of the clusters using the plots of all events. For early time periods, where events are sparse, this is reasonably easy. In later time periods, more detailed plots are needed, but even then such a task seems prohibitively time-consuming and imprecise. In any case, the proposed method is not very different from the reasoning one would likely use during such a process and results are expected to be similar.

3.5.2 Pattern of Aftershock Sequences

Examination of the cluster plots is of interest to formulate a statistical model for the aftershock sequences. Such a model is however only of secondary importance in seismic hazard analysis and its study falls outside the scope of this thesis. A statistical model of the clusters would be however of interest to seismologists and to risk engineers in the context of earthquake prediction. One may notice some secondary clustering of "primary clusters", for instance during the period around 1880. The geographical distribution of cluster centers is also reasonably consistent with that of the main shocks. The spatial resolution of these plots is insufficient to examine in detail any spatial pattern of the secondary events around the main events. On the other hand, the figures illustrate clearly the large variations in the time span of the clusters, also for main events of the same size. There is no clear evidence of geographical dependence of these time spans. Finally, one should note that a major problem in a formal statistical analysis of the clusters is posed by incompleteness: from the figures of main events only, it is evident that early periods are highly incomplete, especially for events of small size. As a consequence, one may expect

that also clusters identified in those periods are only partially identified and, thus, less representative of the actual cluster shape and size. In addition, uncertainty on the earthquake parameters (epicentral location, size of the events, and time of occurrence) may differ significantly for early and more recent events. Therefore, a statistical analysis of the clusters should perhaps focus only on the data that are reliable and clearly delineated from the background. Unfortunately, this may lead to a data set which is too small to produce definitive conclusions.

3.5.3 Pattern of Main Shocks

Estimation of the recurrence rate of main shocks as a function of time, spatial location and size will be discussed in detail in Chapter 4. Here, the two most striking features of the main-shock sequence, incompleteness and "non-Poissonian" patterns, are discussed informally.

Figures 3.14a.3 and 3.14b.3 indicate clearly that, prior to 1870, the catalog is extremely incomplete, except for events of large size ($I_o \geq 4$). Chiburis (1981) suggested that the sudden increase in seismicity around 1870 is associated with an increased probability of reporting. One should note that, around this time, newspapers and magazines become major sources of earthquake reports. On the other hand, the decrease of seismic activity that follows does not seem to confirm such an hypothesis; possibly, part of the earthquake sequence during the more intense period should be classified as a swarm. In more recent time periods, the relative proportion of reported events of small size increases gradually (see the histograms of I_o in the figures). Better

management of the seismic network can possibly explain the jump in seismic activity after 1925. Note also that, if one accepts the hypothesis of an exponential decay of the recurrence rate with earthquake size, then one should conclude that also during the last time period (1974-1981) events with $I_o < 3$ are incompletely reported.

Close inspection of the data after 1870 also seems to indicate several non-Poissonian characteristics or short-term and relatively local variations of the recurrence rate, which are not explained by incompleteness. In particular, it appears that crustal stress at a given location is released in time-lapses, rather than continuously, and shifts from one location to another. The latter pattern is most clearly observed in the last time period. Definite conclusions are not easy to reach based on these figures because of the confounding effect of incompleteness. In Chpater 4, a model is proposed that attempts to quantify incompleteness of earthquake reporting. Examination of the difference between observed and predicted seismicty is a better way to enhance nonstationary episodes and non-Poisson anomalies.

3.6 RESEARCH DIRECTIONS

Although the present method is considered satisfactory for the identification of secondary events of the foreshock and aftershock type, some potential improvements are worth mentioning.

1. The a-priori choice of the background window size is somewhat arbitrary and can possibly introduce bias. Alternatives one might consider are the internal estimation of the extent of the background window (e.g. based on a K-nearest neighbor method) or a

non-parametric estimation of the "local" background recurrence rate within a fixed window (e.g., fitting a locally linear, monotonically increasing recurrence rate inside the background window).

2. Detection of the cluster shape is presently based on the scheme of Figures 3.4d and 3.4e. As pointed out before, other extension schemes are possible and worth investigating. A possibility which has not been mentioned yet, is to determine the extent of the cluster by moving from neighbor to neighbor, using either some heuristic rule to simulate visual identification or statistical tests based on nearest neighbor distance.

3. No measure of how well the clusters are separated from the other events is presently calculated. In particular, it would be useful to obtain an estimate of the misclassification errors, i.e., of the probability that a window found to be significant actually contains only main events and, vice versa, that a window found to be non-significant, actually contains one or more aftershocks. Estimation of such probabilities could possibly proceed along the following lines:

   a. Estimation of the distribution for the ratio R between the recurrence rates in the local and extended windows, using catalog data.

   b. Calculation for each window in the analysis of the likelihood of the "local count" for given "extended count" and window sizes, as a function of R.

   c. Calculation for each window of the a-posteriori probability that $R \leq 1$ (no clustering) and $R > 1$ (clustering). Summing the

probability that $R < 1$ for all windows labeled as significant in the analysis gives an estimate of the misclassification error for clusters.

Note however that Step (a) requires further investigation, because it is not clear whether the assumption of a single distribution of R for all windows is reasonable. One may expect for instance that for backgrounds with higher seismicity (during more recent periods or at more active locations), high values of R are less probable, if one assumes that the size of a cluster inside the initial window is less sensitive to incompleteness or to seismic activity than the background rate.

Chapter 4

ESTIMATION OF INCOMPLETENESS AND RECURRENCE RATES

4.1 INTRODUCTION

After conversion of the different size measures to a single scale m (see Chapter 2) and the removal of dependent events (Chapter 3), the earthquakes in the catalog can be thought of as points in a multi-dimensional space $(\underline{x},t,m)$: for earthquake i, $\underline{x}_i$ is the geographical location, t_i is the time of occurrence and m_i is a unique size measure. The problem discussed in this chapter is how to estimate the rate density function $\nu(\underline{x},m)$ from the historical data. This function is defined such that $\nu(\underline{x},m)\, d\underline{x}dm$ is the expected count of earthquakes in the $(d\underline{x},dm)$-neighborhood of $(\underline{x},m)$. Two basic assumptions will be used throughout this chapter:

1. The earthquake sequence is a realization of a Poisson process, i.e. points in $(\underline{x},t,m)$ space are independently located.
2. Nonstationarity of the observed earthquake sequence is attributed to incomplete reporting, whereas the seismicity generating process is stationary. Therefore, the rate density of <u>reported</u> events can be written as

$$\lambda(\underline{x},t,m) = P_D(\underline{x},t,m)\ \nu(\underline{x},m) \qquad (4.1)$$

where $P_D(\underline{x},t,m)$ is the probability that an earthquake of size m, and at location $\underline{x}$ and time t is reported. It is further assumed that detection/no-detection of different earthquakes are independent events.

Both the assumptions of independence and stationarity of the seismic process are debatable, especially over short time periods as illustrated by the exploratory analysis of the New England data in Section 3.4. These assumptions are maintained here for three reasons: First, there is little physical basis to establish a model explaining the micro-variations of seismicity. Second, because of computational constraints and the lack of sufficient data, a statistical model that is more complex with respect to nonstationarity or non-Poissonian characteristics would have to introduce other simplifying assumptions, for instance about the spatial variation of seismicity or about the incompleteness of the catalog. Finally, deviations of the historical data from the proposed model can be detected a-posteriori, i.e. by comparison of the predicted and observed recurrence rates. If such deviations are significant, local corrections to the model could be made, for instance, using judgement or formal Bayesian updating.

Current procedures for the estimation of the recurrence rates usually employ several additional assumptions, such as

1. $\nu(\underline{x},m)$ is spatially constant within given regions Ω_k, usually referred to as seismogenic provinces; hence

$$\nu(\underline{x},m) = \nu_k(m) \qquad \text{for } \underline{x} \in \Omega_k \tag{4.2}$$

2. The rate density inside province k, ν_k, varies exponentially with m, i.e.

$$\ell n\ \nu_k(m) = a_k - b_k\, m \qquad m_0 \leq m \leq m_1 \tag{4.3}$$

where a_k and b_k are unknown parameters, m_0 is a lower bound of interest and m_1 is a physical upper bound, which may vary from province to province.

3. Inside prespecified regions S_ℓ, the catalog is complete for magnitude m within the last $T_\ell(m)$ years (so-called periods of completeness), so that

$$P_D(\underline{x},t,m) \equiv 1 \qquad \text{for } \underline{x} \in S_\ell \text{ and } t \geq t_o - T_\ell(m) \tag{4.4}$$

where t_o is the most recent time of observation included in the catalog. Notice that the seismogenic provinces Ω_k are not necessarily the same as the completeness regions S_ℓ, the latter being characterized by uniform detection capability rather than uniform seismicity.

Under the above assumptions, estimation of the parameters a_k and b_k in each province is relatively straight-forward if only earthquake data within the periods of completeness are used. A technique which is currently used for doing so will be reviewed in detail in Section 4.2.

In the present chapter, four statistical models, A to D, are presented, which extend one or more of the assumptions in Eqs. 4.2, 4.3 and 4.4. These models differ fundamentally from earlier ones in the sense that the probability of detection P_D and the seismicity rate $\nu(\underline{x},m)$ are simultaneously estimated from the data. Doing so allows one to utilize a larger part (possibly all) of the historical data and provides means to objectively quantify the completeness of the catalog. Depending on which of the four models is used, information on P_D will be derived only from the nonstationarity and non-exponentiality of the observed recurrence rates (model C-D) or also from the distribution of population and seismic instruments in time and space (models A and B). Other extensions that are considered in the various models are with respect to the spatial variation of seismicity, the relation among the slope parameters b_k for

different provinces, the uncertainty on the location $\underline{x}_i$ and size measure m_i of the historical earthquakes, and the assumed exponentiality of the the recurrence law in Equation 4.3. Techniques to examine the goodness-of-fit of the models and to obtain estimates of uncertainty on the parameters are also discussed.

Although models A to D were developed cronologically in an attempt to improve their performance, each has its own merits and sheds light into the problem of estimating recurrence rates and incompleteness. Before going into technical details, it is useful to consider the work presented here from a more global perspective. For this purpose and after reviewing a traditional technique for the analysis of the catalog data in Section 4.2, Section 4.3 describes the conceptual basis of the models and motivates different assumptions or techniques that are used. Section 4.4 analyzes qualitatively the different causes of incompleteness and describes available data. Because the different models have much in common (all of them use some form of discretization in the multi-dimensional space of $\underline{x}$, t and m and a maximum-likelihood method to estimate the parameters), the numerical procedures are developed in parallel in Sections 4.5 to 4.9. The likelihood formulation will be introduced in its simplest form in Section 4.2, while reviewing techniques currently used for the estimation of recurrence rates. Section 4.5 considers various representations of the variation of P_D with the time, geographical lodation and size of the earthquakes. The extended maximum likelihood equations, accounting for the probability of detection, are developed first in Section 4.6 for the case when no prior information is available on the parameters, and then in Section 4.7 for the case when a-priori information needs to be considered. Numerical procedures used to solve these equations are also

discussed. In Section 4.8, the maximum likelihood formulation is further developed to allow for uncertainty on the data and a numerical solution technique is presented. Section 4.9 discusses methods to check the goodness-of-fit of the model and to quantify uncertainty on the estimated parameters. Application of the models to actual data is presented in Sections 4.10 to 4.13, one section for each of the models. The data used are those of the Chiburis catalog, presented earlier, and of a catalog for northern Italy (Friuli region, ENEA, 1984). Conclusions and recommendations for further research are given in Section 4.14.

4.2 MAXIMUM LIKELIHOOD FORMULATION FOR A SEISMOGENIC-PROVINCE MODEL WITH PERIODS OF COMPLETE REPORTING

The purpose of this section is twofold: The first objective is to exemplify the estimation of recurrence rates on the basis of Equations 4.2, 4.3 and 4.4. For convenience, such a technique will be referred to as Stepp-Weichert-Seismogenic-Province method (SWSP method). The second objective is to introduce for this simple model the likelihood formulation used extensively later in this chapter and to present typical uncertainties on the parameters. No attempt is made to present an exhaustive review of all methods which have been used for the estimation of earthquake recurrence rates. Suffice it to say that, with minor variants, the SWSP method is very widely used for the purpose of calculating seismic hazard. A broad discussion of previous models of earthquake occurrences can be found in Basu (1977) with emphasis on seismic hazard and in Savage (1975) with emphasis on geophysical aspects.

4.2.1 The Stepp-Weichert-Seismogenic-Province Method

The first step in a SWSP method is to partition the geographical plane into regions Ω_k that can be assumed homogeneous with respect to seismic activity (see Equation 4.2). Unfortunately, in the Eastern United States as well as in many other regions, there is no strong physical association between seismicity and tectonic, geological or geomorphological variables, on the basis of which one might identify such earthquake sources. An extensive study by Barstow et al. (1981) has concluded that, although certain physical anomalies often occur in regions of strong seismicity, earthquake activity is not always present where such anomolies are found. In addition, the historical data rarely indicate abrupt changes of seismicity at certain boundaries. As a consequence, the specification of seismogenic provinces is somewhat controversial. In most seismic hazard studies (e.g. WGC, 1983) it is therefore common practice to analyze several alternative seismic source configurations. Such configurations can be judgementally determined on the basis of geophysical data or be derived from the historical data. Several examples of source zones for the New England area and the Eastern U.S. are found in WGC (1983) and EPRI (1985) respectively. For instance, Figure 4.1 shows a proposed source configuration within a region, which will be studied later in the application of the models.

The second step of the SWSP method is to determine periods of completeness for the region of interest (see Equation 4.4). The underlying notion is that nonstationarity of the events in the catalog is due to incomplete reporting of the earthquakes. The problem of missing data is especially severe for earlier time periods, for sparsely populated areas and for events of smaller size. Apparent nonstationarity due to

incompleteness is quite evident in Fig. 4.2a, where the empirical recurrence rate in the region of Fig. 4.1 is plotted for each intensity against the period of observation. Stepp (1972) has proposed to estimate the periods of completeness $T_\ell(m)$, for magnitude m and within region S_ℓ, based on the stability of the empirical recurrence rate and to use only data within these time intervals in the estimation of recurrence rates. The method requires a certain degree of judgement, especially at very low and very high intensities, due to statistical variability of the empirical rates and to the fact that, for small size measures, the catalog may be incomplete even today. The difficulty of estimating $T_\ell(m)$ is even greater if one analyzes each province indicated in Fig. 4.2a separately, e.g. to account for differences in population density and instrumentation; see for example, Provinces 1,3,6 and 7 in Figs. 4.2b. The fact that the recurrence rates in each province should follow the parametric relationship in Eq. 4.3 adds one more level of complexity, because the exponential parametrization couples the estimation of the periods of completeness with that of the recurrence rates.

The final step in the SWSP method consist of estimating the recurrence parameters a and b in Eq. 4.3, from the given periods of completeness $T_\ell(m)$ and the associated historical recurrence rates. Weichert (1980) has shown that such estimates can be obtained by a maximum likelihood method, which accounts for the unequal periods of observations for various magnitudes. Weichert has also derived an expression for the asymptotic variance on the estimated slope parameter b_ℓ, which extends earlier results by Aki (1965), Utsu (1966) and Page (1968). More recently, Bender (1983) has derived numerically the distribution of the

maximum-likelihood estimator of b_ℓ for small sample sizes, for the case of equal observation periods. Possibly large discrepancies of the maximum likelihood estimates with alternative estimates, such as the least-squares values based on the empirical density or cumulative distribution function, have been reported by Utsu (1966), Weichert (1980) and Bender (1983).

As an introduction to the likelihood formulation used in this chapter, maximum likelihood estimation of the recurrence parameters in Equation 4.3 is presented in detail in the next subsection. The derivation differs from that of Weichert or Bender, who fix the sample size and estimate only the b-parameter. Under the condition of fixed sample size, the earthquake counts in discrete magnitude intervals follow a multinomial distribution, as opposed to a Poisson distribution. One can show that the Poisson and multinomial sampling scheme lead to the same maximum-likelihood estimate for the distribution of the counts (Bishop et al., 1975). However, if one wants to study the distribution properties of the estimators $\hat{a}$ and $\hat{b}$, then the appropriate model is the Poisson not the multinomial.

4.2.2 Maximum Likelihood Estimation of a and b Parameters in Equation 4.3

To derive the maximum likelihood estimates of the recurrence parameters in Equation 4.3, it is convenient to omit the subscript ℓ, which refers to seismogenic province. On the other hand, to indicate the dependence on magnitude of earthquake counts, recurrence rates and periods of completeness a subscript m will be used. It is assumed earthquake magnitude is discretized into intervals of equal width. Recurrence rates of earthquakes with different discrete magnitude m follow the parametric relation similar to that of Equation 4.3. We write such relation as

$$\nu_m = \exp\{a - b\, m\} \qquad m_0 \leq m \leq m_1 \tag{4.5}$$

It should be noted that the various bias corrections found in the literature for magnitude discretization or upper-bound magnitude are corrections to estimates obtained by maximizing a likelihood that does not consider those characteristics of the distribution, i.e. that is not the correct likelihood. If the likelihood is correctly formulated, the maximum- likelihood estimates are asymptotically unbiased under very general conditions (Cox and Hinkley, 1974). For instance, Weichert (1980) showed how various bias corrections in the literature are implicit in the maximum likelihood equation.

If the historical magnitudes are uncertain, the question arises of how to assign each earthquake to a discrete magnitude interval. The problem of uncertainty on earthquake size was addressed earlier in Chapter 2, where a deterministic bias correction was proposed. In model C and D, uncertainty on the size measures will be explicitly incorporated into the likelihood formulation (see Section 4.8).

Consider next the derivation of the likelihood. For a Poisson process with recurrence rate ν_m, the probability of observing n_m earthquakes over a period T_m has Poisson distribution:

$$f_{N_m}(n_m) = \frac{(\nu_m T_m)^{n_m}}{n_m!} \exp\{-\nu_m T_m\} \tag{4.6}$$

Therefore, the likelihood of the earthquake counts $\{n_m\}$ over the magnitude range $[m_0, m_1]$ depends on the unknown recurrence rates ν_m as

$$\ell\left(\{\nu_m\}\,|\,\{n_m,T_m\}\right) = \prod_{m=m_0}^{m_1} f_{N_m}(n_m) \tag{4.7}$$

Using the relation in Equation 4.5, the likelihood may be expressed as a function of the parameters a and b as follows:

$$\ell\left(a,b\,|\,\{n_m,T_m\}\right) \propto \prod_{m=m_0}^{m_1} \exp\{n_m(a-bm)\}$$

$$\exp\Big\{-\sum_{m=m_0}^{m_1} T_m \exp(a-bm)\Big\} \tag{4.8}$$

The log-likelihood is of the form:

$$\ell n\ \ell\left(a,b\,|\,\{n_m,T_m\}\right) \propto a \sum_{m=m_0}^{m_1} n_m - b \sum_{m=m_0}^{m_1} m\, n_m$$

$$- \sum_{m=m_0}^{m_1} T_m \exp\{a-bm\} \tag{4.9}$$

Notice that the likelihood depends on the earthquake counts only through the total count N and the total magnitude M,

$$N = \sum_{m=m_0}^{m_1} n_m \tag{4.10}$$

$$M = \sum_{m=m_0}^{m_1} m\, n_m \tag{4.11}$$

Therefore, N and M are sufficient statistics and the log-likelihood function simplifies to:

$$\ell n\ \ell\ (a,b|N,M) \propto aN - bM - \sum_m T_m \exp\{a-bm\} \tag{4.12}$$

The corresponding maximum likelihood equations are found by setting to zero the partial derivates of Equation 4.12 with respect to the unknown parameters a and b. This gives

$$N - \sum_m T_m \exp\{a-bm\} = 0 \tag{4.13}$$

$$-M + \sum_m m\ T_m \exp\{a-bm\} = 0 \tag{4.14}$$

There is a simple interpretation for these equations: Equation 4.13 implies that the expected count should equal the observed count, whereas Equation 4.14 requires equality of the expected and observed total magnitude. Uniqueness of the maximum-likelihood estimates can be shown by demonstrating that the Jacobian of Equations 4.13 and 4.14 is negative definite, so that $\ell n\ \ell$ is a concave function with a single maximum. The Jacobian has the form

$$J = \begin{vmatrix} -\sum_m T_m \exp\{a-bm\} & +\sum_m m\ T_m \exp\{a-bm\} \\ +\sum_m m\ T_m \exp\{a-bm\} & -\sum_m m^2\ T_m \exp\{a-bm \end{vmatrix} \tag{4.15}$$

with negative diagonal terms for all a and b. The determinant $|J|$ is given by:

$$|J| = \sum_m p_m^2 \sum_m q_m^2 - \left(\sum_m p_m q_m\right)^2 \tag{4.16}$$

where $p_m = m[T_m \exp(a-bm)]^{1/2}$

and $q_m = [T_m \exp(a-bm)]^{1/2}$

From the Cauchy-Schwarz inequality, $|J|$ is always larger than zero. This condition and the negativity of the diagonal terms in Eq. 4.15 indicate that the Jacobian is a negative definite matrix.

Equations 4.13 and 4.14 can be efficiently solved using Newton's method. At the k'th iteration, estimates of a and b are found from:

$$\begin{bmatrix} a^k \\ b^k \end{bmatrix} = \begin{bmatrix} a^{k-1} \\ b^{k-1} \end{bmatrix} - J^{-1} \begin{bmatrix} \Delta f_a^{k-1} \\ \Delta f_b^{k-1} \end{bmatrix} \qquad (4.17)$$

where Δf_a^{k-1} and Δf_b^{k-1} are imbalances at the (k-1)'th iteration, respectively in Eq. 4.13 and 4.14. Study of the higher derivatives further shows that convergence is monotonic if Δf_a^0 and Δf_b^0 are respectively positive and negative, i.e. if the initial estimates predict a total count and a total magnitude which are too high. If this condition is not satisfied, the values of a and b in the next iteration may significantly overshoot the solution and produce numerical problems in the calculation of the exponential terms. This problem is easily corrected for by limiting the value of the increments to a and b in each iteration step. One should also note that for $N \neq 0$, $M=0$ (i.e. all counts fall in the lowest magnitude interval, which is assigned by convention the value $m=0$), the maximum-likelihood estimate of b is infinite, whereas for $N=0$ and $M=0$, the parameter a must equal $-\infty$ and b is undefined. If only finite values of a and b are allowed, this problem must be resolved by constraining the solution.

An approximation to the asymptotic covariance matrix of the estimates can be found from the matrix of second derivatives of the log-likelihood with respect to the parameters, i.e. from $-J^{-1}$ (Cox and Hinkley, 1974).

For this purpose, it is useful to introduce following additional variables

$$\mu_0 = \sum_m T_m \exp\{a-bm\} \tag{4.18a}$$

$$\mu_1 = \sum_m m\, T_m \exp\{a-bm\} \tag{4.18b}$$

$$\mu_2 = \sum_m m^2 T_m \exp\{a-bm\} \tag{4.18c}$$

which correspond to increasing moments of the exponential recurrence law (scaled by the periods of complete reporting T_m) and depend of course on the parameters a and b. Using this notation, the negative inverse of the Jacobian equals

$$-J^{-1} = \frac{1}{\mu_1^2 - \mu_0\mu_2} \begin{bmatrix} \mu_2 & \mu_1 \\ \mu_1 & \mu_0 \end{bmatrix} \tag{4.19}$$

Equation 4.19 can be used to derive asymptotic expressions for the variance on the maximum likelihood estimators a and b or any linear combination of a and b. In particular, one may derive the variance of the estimated rate of earthquakes with mangitude in interval m. This variance is

$$\sigma^2_{\hat{\nu}_m} = \sigma^2_{\hat{a}-\hat{b}m} = \frac{\mu_2 - 2\mu_1 m + \mu_0 m^2}{\mu_0\mu_2 - \mu_1^2} \tag{4.20}$$

and is minimum for $m = \mu_1/\mu_0$, which is the expected magnitude of the distribution. For such magnitude, the variance is simply

$$\sigma^2_{\nu_{\mu_1/\mu_0}} = \frac{1}{\mu_0} \tag{4.21}$$

Equations 4.19, 4.20 or 4.21 can be used to approximate the variance on the rate estimator for large sample sizes, by calculating the moments μ_0, μ_1, μ_2 at the maximum-likelihood point. Bender (1983) calculated numerically the estimated slope parameter b for small but fixed sample sizes and equal periods of observation. One should note, that if the sample size is fixed, only a finite number of b-estimates are possible, whereas for a fixed period of observation and given recurrence rates the estimator of b may have any value.

To supplement the results of Bender, the following simulation study has been made: For given periods of observation T_m and given values of a and b, maximum-likelihood estimates a and b are obtained in 500 artificially generated samples. Since it is generally expected that values of b fall within a [0.5,2.0] range based on unit Modified Mercalli Intensity intervals, estimates of b have been restricted to this range. In addition, artificially generated samples with zero count have been excluded from the simulation. The true value of b is assumed to be 1.0 and the expected number of events in the lowest magnitude interval m_0 is varied between 1 and 100 (per year). Eight magnitude intervals are used and results are presented for two sets of completeness periods T_m:

- Case A : T_m = [1, 5, 10, 50, 80, 120, 200, 250] years
- Case B : T_m = [0, 0, 0, 50, 80, 120, 200, 250] years

Fig. 4.3.a shows the distribution of the estimated values of a and b in Eq. 4.5 for ν_0 = 1, 10 and 100. For ν_0 = 100, the distribution of both parameters are nearly Gaussian. For lower values of ν_0, the distribution of a is clearly skewed towards smaller values. This is not surprising, since a is closely related to the logarithm of the total sample size: for

small sample sizes, the logarithmic transformation occasionally produces very low estimates of a. On the other hand, the distribution of b remains nearly Gaussian for all ν_0. For $\nu_0 = 1$, the effect of constraining b to the interval [0.5,2.0] is clear and produces peaks at each boundary. Fig. 4.3b and Fig. 4.3.c summarize the results of both simulations. The figures at the top present the sample average, the sample average plus and minus two sample standard deviations, and the sample minimum and maximum of a and b. For ease of interpretation, the exponential value of the various a statistics are plotted rather than a itself. The figures at the bottom show the sample median and the 10 and 90 % percentiles for the cumulative rates.

Note that because b is constrained to the interval [0.5,2.0], the uncertainty band defined by ± two standard deviations exceeds the sample minimum and maximum for small values of ν_0. The most striking feature of these plots is that uncertainty on the cumulative rates is substantially smaller than one would expect by considering uncertainty on a and b to be independent. This feature is a consequence of the correlation between a and b, and is better understood if one calculates the expected counts in T_m for each magnitude interval m. These expected counts are,

- Case A : $n_m = \nu_0$ [1., 1.84, 1.35, 2.49, 1.47, 0.81, 0.50, 0.23]
- Case B : $n_m = \nu_0$ [0., 0. , 0. , 2.49, 1.47, 0.81, 0.50, 0.23]

It follows from these counts that the expected total sample size is 9.69 ν_0 and 5.50 ν_0 for case A and B, respectively. The corresponding expected average magnitude value is 3.7 and 5.0 for each case. As shown earlier (Eq. 4.20) the uncertainty on the estimated rates ν_m is minimum for this value. A similar variation of uncertainty on the estimated

values of the cumulative rates as a function of m is noted in the figures: In case A, the uncertainty is lowest for m between 3 and 4, and in Case B for m between 4 and 5. This shift of the average value of m explains why, for high values of m, the uncertainty on the cumulative rates in case B is not much larger than in Case A.

4.3 OVERVIEW OF PROPOSED MODELS FOR INCOMPLETENESS AND RATES

In this chapter, four statistical models are presented which relax one or more of the assumptions made in Equations 4.2, 4.3 and 4.4. The purpose is to give a global overview of the models, with emphasis on their relative merits and the motivations behind their respective assumptions.

Model A originated from considerations regarding the treatment of incompleteness in current practice (Equation 4.4): 1. The identification of regions S_ℓ where reporting of the events can be assumed uniform is not evident, 2. One would expect a smooth variation of the period of completeness as a function of location, rather than sudden changes along the boundaries of the regions S_ℓ, 3. As illustrated in Section 4.2, estimation of $T_\ell(m)$ is often difficult, and 4. Only the complete part of the catalog data is used for seismicity estimation.

As an alternative, Model A utilizes all the data in the historical catalog, by replacing the notion of period of completeness in Equation 4.4 with that of a probability of detection in Eq. 4.1. A similar approach was used by Lee and Brillinger (1979) in analyzing the incompleteness of a Chinese earthquake catalog. Model A is however fundamentally different from that of Lee and Brillinger in that the probability of detection is estimated from the data, rather than assigned judgementally. Moreover, P_D

is allowed to vary with several main causes of incompleteness: population density in the neighborhood of the epicenter, distance to the nearest seismic instrument, size of the event, and time of occurrence. A further constraint on the variation of P_D with its parameters comes from the assumed exponentiality of the recurrence rates, the stationarity of the earthquake process, the prior information on P_D for recent times and from imposing smoothness conditions on the variation of P_D with earthquake size, time of occurrence, population density and distance to the nearest instrument. On the basis of the type of size measure reported in the catalog (for instance, Modified Mercalli Intensity or instrumental magnitude) one may also infer how many earthquakes have been detected only by people, only by instruments or by both instruments and people. If the reporting of earthquakes by either source is independent, this information alone can be used to estimate the probability of detection (Bishop et al., 1975).

Another novelty of Model A is the treatment of the slope parameters b_k in Equation 4.3 for different provinces. Instead of treating these parameters as completely unrelated, the options are provided to consider the parameters as independent realizations of the same random variable with unknown mean value and variance, or to be identical. Introducing dependence among the parameters b_k is of interest, because uncertainty on the independent estimates can be rather large for small provinces and because spatially smooth values of b are usually expected.

Finally, since the assumption of exponentiality is not always well satisfied over the entire magnitude range, a weighted likelihood formulation is used in Model A to produce better fitting of the earthquake

counts for the large size measures.

In applying model A to the analysis of the Chiburis catalog (see Section 4.9), various limitations were noted:

1. The method produces estimates of incompleteness and recurrence rates for a given set of seismogenic provinces. However, the boundary of provinces with homogeneous activity may not be initially known; in fact, homogeneous provices may not even exist. Thus, it would be desirable to estimate incompleteness and local actvity rates without reference to seismogenic provinces.
2. Although P_D is modelled as a function of the main exploratory variables (time, size, population and instruments), differences in the effect of time for earthquakes reported by people or by instruments were not allowed. For instance, it is reasonable to assume that, for a given population density, the percentage of reported earthquakes does not change over the last 80 years. The same assumption is however unlikely to hold for a fixed distance to the nearest seismic instrument, since the quality of these instruments and the operation of the seismic network has improved significantly in the recent past.
3. In Model A, the variation of P_D with earthquake size m is non-parametric. On the other hand, it would seem that the influence of m on the probability of detection could be inferred on physical ground, for instance, by accounting for the variation of population exposed to ground motion and of site intensity at the location of the nearest instrument.

Model B adresses the above concerns. First, the assumption of seismogenic provinces in Equation 4.2 is replaced by that of smoothly

varying recurrence rates on a discretized spatial grid. The degrees of smoothness of a and b can be controlled separately, so that, depending on the degrees choosen, a range of solutions is produced. Later on, this idea was further developed to allow for piecewise smooth variation of the estimates within specified regions, thus effectively extending the concept of seismogenic provinces. Such an extension is useful because the identification of seismogenic provinces is often a difficult and controversial operation (see Section 4.2). By allowing for a partially data-based, partially judgemental modelling of seismicity, fewer and larger seismic sources could be specified reflecting geological information independent of the historical earthquake data. Second, the variation of P_D with its arguments was changed in accordance with the conclusions of Model A. In particular, model B incorporates a physical representation of the dependence of P_D on the earthquake size. Also, the effect of time on the reporting probability is allowed to be different for population and instruments.

Finally, more consideration is given under Model B to validation of the statistical model. Because of the large number of parameters involved, the sparseness of the earthquake count, and, most of all, the prior information used in the solution, usual goodness-of-fit statistics such as χ^2 are not very useful (e.g. the number of degrees of freedom is not well defined). As an alternative, use is made of an exploratory analysis of the residuals for different subsets of the data (e.g. by comparing predicted and actual counts in different space-time cells).

Application of model B to the Chiburis data lead to the following conclusions:

1. The redefinition of population density to eliminate the magnitude as an independent factor for P_D is not always appropriate. For instance, some of the earliest large earthquakes occur in very sparsely populated regions (even after accounting for the larger felt area). One might speculate that, for such damaging earthquakes, the presence of even a small number of people is sufficient to obtain historical records of the event.
2. The model assumes that reporting of earthquakes by people and instruments are independent events, given the location, size, and time of occurrence. It appears however that in recent years, attention has focused on recording instrumental size measures. For instance, for earthquakes that are detected by instruments, no report of an empirical size measure is usually found, even for large events with epicenter in densily populated areas.
3. Estimation of spatially smooth values of the recurrence parameters was found in some cases to be computationally demanding.

Neither of the previous models addresses the fact that the reported location and size measure of the earthquakes may be uncertain which may be a problem, especially for the very early events. This issue was found to be important in a preliminary analysis of Italian earthquake data (Friuli catalog). A measure of uncertainty on location is given in this catalog, whereas this is not the case for the Chiburis catalog. A different model was therefore developed for the analysis of the Friuli catalog. Model C has the following distinct features:

1. Because the region being analyzed is relatively small, it is reasonable to assume that the probability of detection does not vary

in space and thus that the population density is non-informative. Since most of the historical data have only an empirical size measure reported, also the location of seismic instruments is not considered in this model. Variation of P_D with time and magnitude is inferred from the nonstationarity and non-exponentiality of the empirical recurrence rates. Because time periods where P_D is very small add little information on the seismicity parameters, Model C incorporates the option of analyzing only the part of the data which falls inside a time interval, which may vary with earthquake magnitude. In the special case where P_D is fixed to 1 inside these intervals, the method is equivalent to using given periods of completeness.

2. Because incompleteness is not allowed to vary in space, smoothness of the seismicity parameters a and b is directly related to spatial smoothness of the observed counts. This characteristic allows one to consider nonparametric estimation techniques other than the maximum penalized likelihood criterion of Model B. In fact, Model C uses a kernel-estimation technique, which is computationally more efficient.

3. The location and size of the earthquakes are treated as random variables with known prior distribution. Two approaches are then possible. One is to estimate the parameters of the model as well as the unknown location and size by maximizing the total likelihood. Alternatively, only parameters of the model are estimated by maximizing the expected likelihood, where expectation is with respect to the unknown size and location of each historical event. Difficulties of the total likelihood approach have been discussed earlier in Section 2.5.1 in the context of magnitude conversion.

Because of these difficulties, the second approach is used in Model C.

4. More consideration is also given in Model C to determine uncertainty on the parameter estimates. This problem is not an easy one because of the large number of parameters, the smoothing and other prior information used in the model and uncertainty on the location and size of the historical earthquakes. Model C uses a simple bootstrapping technique, which creates artifical samples from the estimated model (without considering uncertainty on the generated earthquake magnitude and location). This approach should provide a lower-bound to actual uncertainty.

Application of Model C to the Friuli data proved successful and suggested a similar approach to the New England data. However, the spatial variation of incompleteness, especially for early periods of the catalog is too obvious in New England to be neglected. Moreover, if such spatial variation is allowed, a kernel-estimation of the recurrence parameters does not seem feasible. As a result, the last model (Model D) combines elements from all previous models. It also includes some new elements:

1. P_D is determined as in Model C, but regions with different completeness characteristics can be specified. P_D is then estimated separately for each such region.
2. Spatial variation of the seismicity parameters is determined through maximum penalized likelihood, as in Model B. However, a somewhat different form is used for the penalty term to improve convergence of the solution.

3. The bootstrapping technique introduced in Model C is used more extensively to determine uncertainty of the estimators of the seismicity parameters.

The four statistical models cover a wide range of assumptions and present various degree of computational complexity. Conclusions about the validity of the assumptions and the possibility of simplifying the analysis will be presented in Section 4.14.

4.4 INCOMPLETENESS: CAUSES AND DATA

Before developing a statistical model, it is useful to analyze the main reasons why an earthquake of size m, epicentral location $\underline{X}$, and time of occurence t may not appear in the catalog. The process that leads to enlisting an earthquake in the catalog comprises three steps: observation, recording and transmittal.

The probability of observing an earthquake clearly depends on population density and seismic instrumentation near the epicenter $\underline{x}$ at time t. Knowing the sensitivity of each type of observer - an individual or an instrument - and knowing the attenuation law which relates site intensity to epicentral intensity, the probability of detection by each observer can be calculated. Observer sensitivity may be a function of time. This is especially true for instruments, as a consequence of technological innovations, but also for humans, e.g. due to increased awareness and to the growing number of tall buildings producing amplification of the ground motion. Because of the spatial correlation of earthquake attenuation, one may expect earthquake detections by observers at nearby locations to be probabilistically dependent events.

Recording of an earthquake is an even more complicated process. Most of the early entries in the Chiburis catalog, say before 1780, are based on earthquake accounts in missionary reports, personal diaries, and town histories. After 1780, records are usually found in newspapers and magazines. One may conclude that the probability of recording is mainly a function of population density in the epicentral area and of time of occurrence: time of occurrence determines the mode of recording, whereas population density is clearly correlated with the number of earthquake accounts (diaries, newspapers, etc). Site intensity is another important variable, because more destructive earthquakes are usually more extensively documented.

Imperfect transmittal includes the loss of documents and the possibility that existing earthquake records may have remained undiscovered. Therefore, the probability of transmittal is mainly a function of time and of effort in the search for relevant documents.

In summary, the major factors that influence incompleteness are: time of occurrence, population density especially in the epicentral area, and seismic instrumentation. The effect of each factor further depends on epicentral intensity.

In order to estimate the dependence of the detection probability P_D on population density and seismic instrumentation, maps have been compiled which describe the evolution in time of demography and instrumentation in the region of interest. Boundary effects have been eliminated by extending the region one degree in each direction; see Fig. 4.1.

Population maps for the U.S. are given in Friis (1960) for years prior to 1790 and in Lord and Lord (1953) for more recent years.

Demographic data for Eastern Canada is found in the National Atlas of Canada (1974). The format of the maps varies for different sources: The maps of Lord and Lord use a discretization of population density according to the six categories in Table 4.1. Those of Friis indicate the location of each 200 rural inhabitants and the location of cities with a population of 3000 or more. The National Atlas of Canada indicates in a similar way groups of 1000 rural inhabitants and cities of size 10,000 or more. Prior to 1850, the last reference gives only the date of arrival of early settlers in cities whose population in 1961 exceeded 10,000.
For the present analysis, maps are needed on a common population-density scale and over a common geographical grid. The scale of Table 4.1 is an appropriate one: it has a high resolution at low population densities, which is where the probability of detection is most variable. Other maps have been converted to the same format, using judgement when a precise conversion of scale could not be established. The discretization grid has been defined by meridians and parallels within the region of Fig. 4.1, with a quarter-degree spacing in each direction.

Twelve population maps have been compiled on this discrete grid for the period from 1625 to 1950 (Fig. 4.4). The time interval between consecutive maps is approximately 25 years before 1780 and approximately 40 years afterwards. After 1950, the population is assumed to have remained stable. Although the latter is the period when more accurate demographic information is available, any increase of population above the 1950 level would only produce insignificant changes in the estimated probability of detection: Higher completeness of the catalog in recent years is due almost exclusively to more reliable recording and transmittal

and to the installation of a denser seismograph network. Fig. 4.5 shows the fraction of total area associated with each population category as a function of time. Notice that, because the population map of 1950 is characterized by a sharp contrast between rural and urban population, population Category 4 almost disappears in recent years. The persistence of very low population density at this time is due mainly to the fact that some provinces extend over the Atlantic Ocean.

Information on the evolution in time of seismic instrumentation is found in several sources: A comprehensive list of seismic stations in the United States, their location, operating dates, and instrument characteristics has been assembled by Poppe (1979). Early stations, both in the United States and in Canada, are also described in Stevens (1980). Information for the more recent Canadian stations is given by Halliday et al. (1977,1981). Based on this data, a list of operating seismic instruments in the region has been compiled for each year. Fig. 4.6 shows the total number of stations as a function of time and indicates a noticeable improvement of the network during the early 1970's.

The probability of detection of an earthquake depends on the configuration of the seismograph network near the epicenter. In order to account for instrument location, the distance to the nearest operating station has been calculated for each cell of the spatial grid and for each year from 1910 to 1980. Distances have then been classified into the five categories of Table 4.2. According to intensity attenuation models developed for the Eastern United States, the distance intervals in Table 4.2 correspond to approximately unit changes of site intensity. A representative sample of the resulting instrumentation maps is given in Fig. 4.7 for a few selected years.

4.5 MODELS FOR THE PROBABILITY OF DETECTION

4.5.1 Introduction and Notation

A major novelty of the present analysis is that both the probability of detection P_D and the recurrence rate ν in Eq. 4.1 are estimated from the catalog data.

The only published work on methods of this type is that by Kelly and Lacoss (1969) and Brillinger (1976). Kelly and Lacoss assume that, for instrumentally reported events, P_D has the form of an error function:

$$P_D(m) = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{m} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \qquad (4.22)$$

where μ and σ are unknown parameters and m refers to body wave magnitude. Assuming that the true recurrence rate is exponential, they estimate by maximum likelihood the parameters μ and σ as well as the recurrence parameters a and b in Eq. 4.3, for the first 2000 events reported by USCGS in 1968. The estimates obtained are $\mu = 5.1$ and $\sigma = 0.415$, i.e. the probability of detection at that time is found to be 0.5 for events with body wave magnitude equal to 5.1. Brillinger (1976) discusses from a theoretical point of view how a probability of detection that varies only in time can be estimated from an incomplete realization of a point process.

The models used in this chapter consider that P_D either varies with time and magnitude (model C) or with time, magnitude and location (models A, B and D). In Model D, variation of P_D in space is a-priori specified, whereas in Models A and B spatial variation of P_D is estimated from the data and information on the population density and seismic instrumentation

in the neighborhood of the epicenter. Models A and B also differentiate between the probability of detection by people and by instruments. The following notation is useful in that respect:

- z is a bivariate indicator variable, whose three possible values define the mode of detection as follows:

  $z = \{1,0\}$ for events detected by people only

  $z = \{0,1\}$ for events detected by instruments only (4.23)

  $z = \{1,1\}$ for events detected by both people and instruments

- p denotes a measure of population density at a given time and location and will be defined more precisely when considering each model.
- similarly, d denotes a measure of the distance to the nearest seismic instrument.

As before, dependence on the the explanatory variables t,m,p and d will be indicated by subscripts. z is used as a superscript for probabilities that vary with the mode of detection. In both models A and B, it will be assumed that reporting by instruments and reporting by people are independent events, given t, m, p and d. The symbol P_D with no superscript refers to the probability of detecting an earthquake by either people or instruments and can be written as:

$$P_D = P_D^{(0,1)} + P_D^{(1,0)} + P_D^{(1,1)} \tag{4.24}$$

Prior to the installation of seismic instruments, $P_D^{(0,1)}$ and $P_D^{(1,1)}$ are evidently equal to zero and P_D equals $P_D^{(1,0)}$.

4.5.2 Common Features

Consider next the problem of modelling P_D as a function of the variables affecting incompleteness, which are (p,t,m) and (d,t,m) respectively for people and instruments. Features that are fundamental to the analysis and common to all models are discussed first, whereas implementation details for the various models will be given later in the section.

First, all variables are discretized: discretization is essential to arrive at a practical solution of the maximum-likelihood problem because this solution involves repeated calculation of an integral of the recurrence rate over the domain of interest in ($\underline{x}$,t,m,p,q)-space (see Section 4.6). Examples of such discretizations will be shown in the application of the models in Sections 4.10 to 4.13. To avoid laborious notation, the names of the discretized variables are left unchanged. Hence, for example, t refers to time intervals rather than continuous time.

Second, a nonparametric representation of P_D is preferred to an analytical form such as that in Equation 4.22. Although parametric models have the advantage that monotonicity or smoothness can be implicitly imposed, estimation of the parameters is often more difficult and validity of the parametric assumption may be dubious. If the ordering of the explanatory variables is neglected, the problem of modelling P_D is clearly related to that of model selection in categorical data analysis (Bishop et al., 1975; Fienberg and Holland, 1980).

Techniques for categorical data analysis on ordered variables are presented by Agresti (1984). Notice however that the present problem is a

very particular one, because the "time" over which the earthquake process is observed in each category (p,t,m) and (d,t,m) may vary. For instance, in recent time-periods categories with low population density occupy a much smaller area than those with high population density (see Fig. 4.5). The limiting case when the time of observation is zero for a given category corresponds to the presence of a "structural zero" in a categorical table. Such cases have been treated extensively in the literature. No discussion of the present case of a Poisson sampling scheme with period of observation that varies from category to category has been found in the literature.

Another complication is that, if detection by instruments and detection by people are separated, an additional category, the mode of detection z, must be considered. By definition, the categorical table is incomplete for the missing counts, i.e. categories with z=(0,0) are not observed. Bishop et al. (1975) discuss this case as the "capture-recapture" problem for the usual Poisson sampling scheme and show that, given some assumption about the structure of the model with respect to z (e.g. independence of reporting for z=(1,0) and z=(0,1)), the probability of being in class (0,0) can be estimated. This is of importance, since it implies that, under the assumption of independent detection by people and instruments, missing counts can be estimated without additional information. This property does not hold if the mode of detection is not considered. In the latter case, only the relative variation of P_D with its explanatory variables can be inferred, while the absolute value of P_D is not identifiable.

4.5.3 Model A

Model A assumes the following simple structure for the variation of $P_D^{(1,0)}$, $P_D^{(0,1)}$ and $P_D^{(1,1)}$ with t, m , p and d:

$$P_D^{(1,0)} = \beta_{tm}\ \alpha_{pm}(1-\gamma_{dm}) \tag{4.25a}$$

$$P_D^{(0,1)} = \beta_{tm}\ \gamma_{dm}(1-\alpha_{pm}) \tag{4.25b}$$

$$P_D^{(1,1)} = \beta_{tm}\ \gamma_{dm}\ \alpha_{pm} \tag{4.25a}$$

where α_{pm}, β_{tm} and γ_{dm} are unknown probabilities.
Notice that α_{pm} and γ_{pm} are treated as independent probabilities, while β_{tm} is used as a common factor. The associated probability of detection, irrespective of detection mode, P_D, is

$$P_D = \beta_{tm}\{1 - [1-\alpha_{pm}][1-\gamma_{dm}]\} \tag{4.26}$$

The quantity β_{tm} can be thought of as the probability of transmitting a reported earthquake, whereas α_{pm} and γ_{dm} give the probability that an earthquake is recorded by people and instruments respectively. For earthquakes of given magnitude m, no interaction is assumed between p and t or between d and t. This implies that the time effect for size measure m is independent of population density and seismic instrumentation. Moreover, it is assumed that the time-effect is identical for both modes of reporting. As will be shown later in Section 4.10, this assumption may not be reasonable.

Before progressing further in the analysis, a more precise definition of the explanatory variables should be given: the discretization of time should be such that the loss of records may be assumed constant inside

each time interval. Such a discretization can be determined on the basis of knowledge of the main sources of earthquake records in different periods (see Section 4.4). The quantities p and d are in Model A discretized versions of population density at the epicenter and distance to the nearest instrument as shown in Figs. 4.4 and 4.7 for different time intervals. A variant of Model A uses the maximum value of p within a distance from the epicenter that depends in a given way on the size of the event. Such a redefinition is useful if one wants to simplify the model by excluding m as an independent explanatory variable.

4.5.4 Model B

Model B assumes that, for a proper definition of p and d, magnitude m has no independent effect. In contrast to Model A, interaction effects of time and population are included and the effect of time may depend on the mode of detection. One reason for allowing interaction between t and p is that one may expect different effects of time in rural zones with low p and urban areas with high p; in the latter, time should be less influential. This leads to the following model:

$$P_D^{(1,0)} = \alpha_{tp}(1-\beta_{t^*}\gamma_d) \tag{4.27a}$$

$$P_D^{(0,1)} = \beta_{t^*}\gamma_d(1-\alpha_{tp}) \tag{4.27b}$$

$$P_D^{(1,1)} = \alpha_{tp}\beta_{t^*}\gamma_d \tag{4.27c}$$

and

$$P_D = [1-(1-\alpha_{tp})(1-\beta_{t^*}\gamma_d)] \tag{4.28}$$

where t^* is a time discretization for the detection capability of seismic instruments.

A critical choice in Model B is of course the definition of d and p, which should implicitly account for the effects of m. For d this is relatively straight-forward: since reporting of earthquakes by instruments should depend mainly on the local intensity of the earthquake, a reasonable choice for d is the site intensity at the nearest instrument. The appropriate definition of p is less evident. Model B uses the following form:

$$p(\underline{x},m,t) = \frac{1}{m^{*r}} \int_{\Omega(\underline{x})} q(\underline{x},t)\ \hat{m}^r(\underline{x},m)\ d\underline{x} \qquad (4.29)$$

where $\Omega(\underline{x})$ is a large but fixed neighborhood around the epicenter $\underline{x}$, q is the actual population density, $\hat{m}$ is the estimated intensity at site $\underline{x}$, m^* is an arbitrary reference site intensity and r is a constant.
Note that, for r=0, p corresponds to the total population in $\Omega(\underline{x})$ and does not depend on m. As r increases, the kernel function $\hat{m}^r(\underline{x},m)/m^{*r}$ becomes narrower and p depends more on earthquake intensity and population near the epicenter. For intermediate values of r, p is a weighted average of the population distribution, with weights that depend on site intensity. The choice of the coefficient r and of the discretization of p will be commented upon in the application of the model in Section 4.11.

4.5.5 Models C and D

Models C and D do not consider the mode of detection or the distribution in space of population and instruments. As a result, the model of probability of detection must be applied to a region that is

sufficiently homogeneous with respect to incompleteness, within the period of time of the analysis. Accordingly, only variation of P_D with t and m is considered.

Models C and D assume that

$$P_D = \alpha_{tm} \tag{4.30}$$

and allow for interaction effects between magnitude and time. In categorical data analysis, a model of this type is said to be fully saturated, because without further constraints, α_{tm} can be choosen to exactly match the observed count in each cell. It is immediately clear that, if the recurrence rates are unknown, P_D can be determined only up to a proportionality factor. For instance, one can scale P_D down and the recurrence rates up without modifying the expected count in each category. Various forms of constraints that allow to determine the actual values of P_D will be discussed in Section 4.7.

4.6 MAXIMUM LIKELIHOOD ESTIMATION OF PROBABILITY OF DETECTION AND RECURRENCE RATE: NO ERRORS IN THE DATA

4.6.1 Introduction and Notation

This section derives the likelihood function and maximum likelihood estimates (m.l.e.) of seismicity and incompleteness parameters. It is assumed that the magnitude, location and time of occurrence of the historical events are known without error. The case when errors on the reported values of m, $\underline{x}$ and t need to be considered, will be discussed in Section 4.8.

The analysis is an extension of that presented in Section 4.2 for the Stepp-Weichert-Seismogenic-Province method. In order to make the presentation concise, a general form of the m.l. equations is derived first, and then the equations are specialized for the various incompleteness models. Modifications to the maximum likelihood to include prior information on the parameters will be discussed separately in Section 4.7.

For the general formulation, it is convenient to consider P_D as a generic function of p,q,t and m, with unknown parametervector $\underline{\theta}$. Also, geographical location $\underline{x}$ may refer here to any partition of the region, including seismogenic provinces or cells of a regular rid, such that the recurrence parameters $a_{\underline{x}}$ and $b_{\underline{x}}$ are constant within the region $\Omega(\underline{x})$ associated with $\underline{x}$. Since the recurrence rates $a_{\underline{x}}$ refer to a unit area, unit time interval and unit magnitude interval, it is necessary to calculate the "volume" occupied by each category $c = (\underline{x},t,m,p,q)$. In accordance with earlier notation in Section 4.2, these volumes will be referred to as periods of observation and denoted by T_c. One should note that p and q vary with geographical location and time, and may even vary within $\Omega(\underline{x})$ or the time interval t. Because calculation of T_c is tedious, the periods of observation are calculated only once and stored. Reduction in computation time and amount of storage is also the reason why discrete variables are used throughout the analysis rather than continuous variables.

As before, the variables on which a parameter or recurrence rate depends are indicated by subscripts, whereas the mode of detection z is indicated

by a superscript. Remember that z has only three possible values, i.e. missing counts are not considered. For easy reference, the relevant parameters are summarized here:

- $a_{\underline{x}}, b_{\underline{x}}$ are recurrence relation parameters, as in Eq. 4.31 below
- $\nu_{\underline{x},m}$ is the actual ("true") recurrence rate, i.e. the rate if all earthquakes were detected
- λ_c^z is the recurrence rate of earthquakes for mode of detection z and category c
- P_D^z is the mode of detection and varies with detection category D = (tmpq).

Relations among these parameters are as follows:

$$\nu_{\underline{x}m} = \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \tag{4.31}$$

$$\lambda_c^z = P_D^z \, \nu_{\underline{x}m} \tag{4.32}$$

Observed, expected observed, and expected "true" counts will be denoted by n, $\bar{n}$ and n^* respectively. These quantities depend on category and n and $\bar{n}$ depend also on the mode of detection. The counts $\bar{n}$ and n^* are related to the previous parameters as:

$$\bar{n}_c^z = T_c \, \lambda_c^z \tag{4.33}$$

$$n_c^* = T_c \, \nu_{\underline{x}m} \tag{4.34}$$

and n_c^* and $\bar{n}_c^z$ are related as

$$\bar{n}_c^z = p_D^z \, n_c^* \tag{4.35}$$

In the derivation that follows, counts and periods of observation need to be summed frequently over a subset of the categories (c,z). The convention is then introduced that, if a count or period of observation is summed over a certain category to calculate a marginal value, the corresponding subscript is omitted. For instance, the expected reported count in category c, irrespective of mode of detection, is denoted by $\bar{n}_c$, where:

$$\bar{n}_c = \bar{n}_c^{(0,1)} + \bar{n}_c^{(1,0)} + \bar{n}_c^{(1,1)} \tag{4.36}$$

Similarly, the total observed count at location $\underline{x}$ is denoted by $n_{\underline{x}}$ and is given by

$$n_{\underline{x}} = \sum_{t,m,p,q,z} n_c^z \tag{4.37}$$

4.6.2 General Form of the Likelihood Function

Under the Poisson assumption, counts in different categories c and detection modes z are independent and follow a Poisson distribution with parameter $\bar{n}_c^z$, i.e.

$$f(n_c^z) = \frac{(\bar{n}_c^z)^{n_c^z}}{n_c^z\,!} \exp\{-\bar{n}_c^z\} \tag{4.38}$$

This probability mass and all the following likelihood functions depend of course on the parameters $a_{\underline{x}}, b_{\underline{x}}$ and $\underline{\theta}$. Because of the Poisson assumption, the total likelihood ℓ of the counts $\{N_c^z\}$ for categories c and z is found by multiplication of the probabilities in Eq. 4.38. Omitting terms that do not depend on the parameters, one finds

$$\ell \propto \prod_{(c,z)} \left[(\bar{n}_c^z)^{n_c^z} \exp\{-\bar{n}_c^z\} \right] \tag{4.39}$$

The log-likelihood can then be written as

$$\ell n\ \ell \propto \sum_{c,z} n_c^z\ \ell n\ \bar{n}_c^z - \sum_{c,z} n_z^c \tag{4.40}$$

From Equations 4.32 and 4.33, it follows that

$$\bar{n}_z^c = T_c\ P_D^z \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \tag{4.41}$$

and

$$\ell n\ \bar{n}_z^c = \ell n\ T_c + \ell n\ P_D^z + a_{\underline{x}} - b_{\underline{x}} m \tag{4.42}$$

Using the convention of eliminating subscripts for counts that are summed over a given set of categories, the first term in Equation 4.40 becomes

$$\sum_{c,z} n_c^z\ \ell n\ \bar{n}_c^z = \sum_{c,z} n_c^z\ \ell n\ T_c + \sum_{D,z} n_D^z\ \ell n\ P_D^z + \sum_{\underline{x}} n_{\underline{x}} a_{\underline{x}} - \sum_{\underline{x}m} m\ n_{\underline{x}}\ b_{\underline{x}} \tag{4.43}$$

Further denoting by $m_{\underline{x}}$ the total observed magnitude at location $\underline{x}$,

$$m_{\underline{x}} = \sum_m mn_{\underline{x}} \tag{4.44}$$

and usings Equations 4.41 and 4.43 in the log-likelihood expression, one obtains

$$\ell n \ell(\underline{a}_x, \underline{b}_x, \underline{\theta} | \{n_c^z\}) \propto \sum_c \sum_z n_c^z \ell n P_D^z + \sum_{\underline{x}} n_{\underline{x}} a_{\underline{x}} - \sum_{\underline{x}} m_{\underline{x}} b_{\underline{x}} - \sum_c T_c P_D \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \quad (4.45)$$

Maximum likelihood equations can be found by computing the partial derivatives of $\ell n\ \ell$ with respect to each of the parameters $a_{\underline{x}}$, $b_{\underline{x}}$ and $\underline{\theta}$. It is instructive to do this in two steps: First, the partial derivatives with respect to $a_{\underline{x}}$, and $b_{\underline{x}}$ are found and then the maximum likelihood equations for $\underline{\theta}$ are derived.

4.6.3 Maximum likelihood equations for $a_{\underline{x}}$ and $b_{\underline{x}}$

Comparison of the log-likelihood in Equation 4.45 with the expression derived earlier in Equation 4.12 for the Stepp-Weichert-Seismogenic-Province model shows that, for given P_D^z, the two expressions are similar. If fact, one may define $T^*_{\underline{x}m}$ as a time period

$$T^*_{\underline{x}m} = \sum_{tmpq} T_c P_D \quad (4.46)$$

and rewrite the log-likelihood

$$\ell n \ell(\underline{a}_x, \underline{b}_x, | \theta\{n_c^z\}) \propto \sum_{\underline{x}} [n_{\underline{x}} a_{\underline{x}} - m_{\underline{x}} b_{\underline{x}} - \sum_m T^*_{\underline{x}m} \exp\{a_{\underline{x}} - b_{\underline{x}} m\}] \quad (4.47)$$

Since log-likelihood contributions from different values of $\underline{x}$ are additive and involve only the local parameters $(a_{\underline{x}}, b_{\underline{x}})$, the maximum-likelihood estimates of $a_{\underline{x}}$ and $b_{\underline{x}}$ are independent for different $\underline{x}$. Because of

similarity with Equation 4.12, $T^*_{\underline{x},m}$ can be thought of as an equivalent period of completeness. Contrary to the usual period of completeness, $T^*_{\underline{x}m}$ combines the entire time span of the catalog by weighting each time interval by the associated probability of detection. Maximum likelihood equations are easily derived from Equation 4.47 and correspond to those found earlier in Eqs. 4.13 and 4.14, i.e.

$$n_{\underline{x}} - \sum_m T^*_{\underline{x}m} \exp\{a_{\underline{x}} - b_{\underline{x}m}\} = 0 \qquad \text{for each } \underline{x} \tag{4.48}$$

$$-m_{\underline{x}} - \sum_m T^*_{\underline{x}m} \; m \exp\{a_{\underline{x}} - b_{\underline{x}m}\} = 0 \qquad \text{for each } \underline{x} \tag{4.49}$$

Considerations made in Section 4.2 on Eqs. 4.13 and 4.14 remain valid and the same iteration scheme can be used to estimate the parameters $a_{\underline{x}}$ and $b_{\underline{x}}$ at location $\underline{x}$. The expression for the asymptotic covariance matrix (Equation 4.19) is still valid, conditionally on given $\underline{\theta}$.

4.6.4 Maximum likelihood equations for $\underline{\theta}$

Consider next the log-likelihood as a function of the parameter values $\underline{\theta}$ used in modelling probability of detection. For given values of $a_{\underline{x}}$ and $b_{\underline{x}}$, this function can be written as

$$\ell n \ell\left(\underline{\theta} \middle| a_{\underline{x}}, b_{\underline{x}}, \{n_C^z\}\right) \propto \sum_D \sum_z n_D^z \; \ell n \; P_D^z - \sum_D n_D^* \; P_D \tag{4.50}$$

Remember that n_D^* is defined as the expected total count (including missing events) in detection category D and depends on $a_{\underline{x}}, b_{\underline{x}}$. For the k'th parameter θ_k, the maximum likelihood equation is found by partial

differentiation of $\ell n\ \ell$ with respect to θ_k. This gives

$$\sum_{D_k} \left[\sum_{z_k} n_D^z \frac{Q_{k,D}^z}{P_D^z} - n_D^* Q_{k,D} \right] = 0 \tag{4.51}$$

where D_k and z_k are the subsets of D and z for which P_D depends on θ_k. $Q_{k,D}^z$ is the partial derivative of P_D^z with respect to θ_k. Interpretation of Equation 4.51 is more evident if one notes that n_D^z/P_D^z is an estimator of the total count in category (c,z). Thus, the first term in Equation 4.51 is a measure of the change in the estimated observed count for each class z as θ_k changes and the last term corresponds to the change in the expected observed count for all z as θ_k changes. Equation 4.51 implies that, for the m.l.e. of θ_k, the two values should be the same when summed over D_k. Parametrization of P_D for the models used here is such that P_D^z is proportional to θ_k or to $1-\theta_k$, depending on the mode of detection. (Model B is an exception to this rule for parameters β_{t*} and γ_d). Modes of detection z for which P_D is proportional to θ_k or to $1-\theta_k$ are denoted respectively z_{k+} and z_{k-}. It follows that

$$\frac{Q_{k,D}^z}{P_D^z} = \begin{cases} \frac{1}{\theta_k} & \text{for } z \in z_{k+} \\ -\frac{1}{1-\theta_k} & \text{for } z \in z_{k-} \end{cases} \tag{4.52}$$

Therefore, the maximum likelihood equation for parameter θ_k simplifies to

$$\frac{n_{\theta_k}^{z_{k+}}}{\theta_k} - \frac{n_{\theta_k}^{z_{k-}}}{1-\theta_k} - \sum_{D_k} n_D^* Q_{k,D} = 0 \tag{4.53}$$

where $n_{\theta_k}^{z_{k+}}$ and $n_{\theta_k}^{z_{k-}}$ denote the total observed count in categories (D_k, z_{k+})

and (D_k, z_{k-}), respectively, and are sufficient statistics. Since P_D is at most linear in θ_k, the derivative $Q_{k,D}$ is constant. Eq. 4.53 is then of second degree in θ_k and can be easily solved. Because the partial derivative of the left side of Eq. 4.53 with respect to θ_k is always negative, there cannot be multiple solutions. Specialized forms of Eq. 4.53 for the various models of P_D will be given at the end of this section.

4.6.5 Solution of the Maximum Likelihood Equations and Specialized Forms

Maximum likelihood equations 4.48, 4.49 and 4.53 can be solved simultaneously for a ,b and $\underline{\theta}$ by iteration: First one solves for $a_{\underline{x}}$ and $b_{\underline{x}}$ for given $\underline{\theta}$, and then one fixes $a_{\underline{x}}$ and $b_{\underline{x}}$ and solves Equation 4.53 for each θ_k. These operations are performed iteratively until convergence. If the derivative $Q_{k,D}$ in these equations depends on components of $\underline{\theta}$, other than θ_k, then additional iterations are necessary. Since the likelihood increases monotonically in each of the iteration steps, convergence must be reached. It is less clear that the solution is unique, i.e. that the unconditional likelihood function has only one maximum. Haberman (1973) has shown that this is true for the case of loglinear models in categorical data analysis. Also, in all numerical applications, the solutions have been found to be independent of the initial values. In the remainder of this section, the specialized forms of Equation 4.53 are given for the models proposed in Section 4.5.

Model A

In model A, the parameters $\underline{\theta}$ correspond to β_{tm}, α_{pm} and γ_{dm} (see Eqs. 4.25). The maximum likelihood equation for β_{tm} is derived as follows: Note that, irrespective of the mode of detection z, P_D^z is proportional to β_{tm}. Therefore, the second term in Equation 4.53 is zero and the summation in the first term extends over all z. The partial derivative of P_D with respect to β_{tm} (variable $Q_{k,D}$ in Eq. 4.53) is found from Equation 4.26. The maximum likelihood equation associated with β_{tm} is then:

$$\frac{n_{tm}}{\beta_{tm}} - \sum_{pd} n_D^* \{1-[1-\alpha_{pm}][1-\gamma_{dm}]\} = 0 \qquad \text{for each (tm)} \tag{4.54a}$$

Maximum likelihood equations for α_{pm} and γ_{dm} are derived similarly. In this case, the second term in Equation 4.53 is however not zero. For instance, the probability of detection is proportional to α_{pm} for z=(1,0) and z=(1,1), and to $1-\alpha_{pm}$ for z=(0,1). The maximum likelihood equations for α_{pm} and γ_{dm} are

$$\frac{n_{pm}^{(0,1)} + n_{pm}^{(1,1)}}{\alpha_{pm}} - \frac{n_{pm}^{(0,1)}}{1-\alpha_{pm}} - \sum_{td} n_D^* \beta_{tm}[1-\gamma_{dm}] = 0 \qquad \text{for each (pm)} \tag{4.54b}$$

$$\frac{n_{pm}^{(1,0)} + n_{pm}^{(1,1)}}{\gamma_{dm}} - \frac{n_{dm}^{(0,1)}}{1-\alpha_{dm}} - \sum_{tp} n_D^* \beta_{tm}[1-\alpha_{pm}] = 0 \qquad \text{for each (dm)} \tag{4.54c}$$

Equation 4.54 together with equations 4.48 and 4.49 define the values of $a_{\underline{x}}$, $b_{\underline{x}}$, β_{tm}, α_{pm} and γ_{dm} for which the likelihood is maximum. It is clear from these equations that the likelihood is invariant to scaling up all recurrence rates while scaling down β_{tm} by the same factor. This indicates that, without any further assumption, only relative completeness can be determined as a function of time.

Model B

In the case of model B, the maximum likelihood equations for α_{tp}, β_{t*} and γ_d are:

$$\frac{n_{tp}^{(1,0)} + n_{tp}^{(1,1)}}{\alpha_{tp}} - \frac{n_{tp}^{(0,1)}}{1-\alpha_{tp}} - \sum_{t^*d} n_D^* [1-\beta_{t^*}\gamma_d] = 0 \qquad \text{for each (tp)} \qquad (4.55a)$$

$$\frac{n_{t^*}^{(0,1)} + n_{t^*}^{(1,1)}}{\beta_{t^*}} - \sum_{d} \frac{n_{t^*d}^{(1,0)} \gamma_d}{1-\beta_{t^*}\gamma_d} - \sum_{dpt} n_D^* \gamma_d [1-\alpha_{tp}] = 0 \qquad \text{for each } (t^*) \qquad (4.55b)$$

$$\frac{n_d^{(0,1)} + n_d^{(1,1)}}{\gamma_d} - \sum_{t^*} \frac{n_{t^*d}^{(1,0)} \beta_{t^*}}{1-\beta_{t^*}\gamma_d} - \sum_{ptt^*D} n_D^* \beta_{t^*} [1-\alpha_{tp}] = 0 \qquad \text{for each (d)} \qquad (4.55c)$$

Note that the maximum likelihood equations for β_{t^*} and γ_d are slightly different, because interactions between t^* and d are excluded in this model. As a result, estimation of β_{t^*} and γ_d is more complicated. When all other parameters are fixed, the values of β_{t^*} and γ_d can be calculated by noting that these values are inside the interval [0,1] and the maximum likelihood equation is monotonic. A solution is then easily found by iteratively refining this interval.

Model C

For the characterization of incompleteness, Model C uses only the parameters α_{tm} and the corresponding maximum likelihood equation are:

$$\frac{n_{tm}}{\alpha_{tm}} - n_{t,m}^* = 0 \qquad \text{for each (tm)} \qquad (4.56)$$

Estimation of α_{tm} for given n_{tm} is such that the observed count in each

(tm) category is matched. Evidently such a model is not well defined, since the recurrence rates can be scaled up and the detection probability α_{tm} can be scaled down without affecting the likelihood. In addition, one can vary the slope parameters $b_{\underline{x}}$ and the probabilities α_{tm} such that the likelihood remains the same. Various forms of prior information on the incompleteness parameters that may be used to stabilize the solution are discussed in the following section.

4.7 CONSTRAINTS, PENALTIES, SMOOTHING, AND A-PRIORI CONDITIONS

4.7.1 Introduction

The maximum likelihood solution derived in Section 4.6 is entirely data-based, i.e. it does not incorporate any prior beliefs about the values of the parameters. Given the small amount of earthquake data available and the number of parameters to be estimated, it is no surprise that these estimates may have large statistical uncertainty. Such uncertainty is in part due to an over-parametrization of the problem. One possibility is of course to fit a model with fewer parameters, for instance by using larger seismogenic provinces or by eliminating categories. Selection of a model with the appropriate number of parameters can also be done systematically, by comparing goodness-of-fit statistics or by calculating likelihood ratios, while considering the decrease or increase in the number of parameters. Another possibility is to use a model with many parameters, whose values are however constrained. Examples of the latter methods are kernel estimation (for a discussion, see Devroye and Györfi, 1985) and penalized maximum likelihood estimation (Tapia and Thompson, 1978). The constraints applied to the parameters may

be determined automatically using goodness-of-fit statistics, e.g. by balancing the bias against the variance of the estimators or specified a priori.

For example, it is usually assumed that in recent periods all earthquakes above a given magnitude have been reported. Similarly, using worldwide observations or other independent data one might form a prior distribution or establish bounds on the slope parameter b (a histogram of various estimated b-values is for instance given in Utsu, 1971). One would expect smooth variation of the seismicity parameters a and b, at least within certain regions, and monotonic variation of the probability of detection with time and magnitude. Such prior beliefs can be incorporated using Bayesian analysis or by appropriately constraining and penalizing the likelihood function in maximum-likelihood estimation.

Several of the above mentioned techniques have been used in the application of the models: The values of P_D are constrained for some of the detection categories. Maximum penalized likelihood estimation (MPLE) and, in one case, kernel estimation are used smooth the variation of a, b and P_D with their respective parameters (geographical location, time, magnitude, etc.). Prior belief about the b parameters is incorporated using Bayesian statistics. The Bayesian approach also provides an alternative interpretation of the MPLE method. In this section, the different forms of prior information and their effect on maximum likelihood estimation are discussed, first for the estimation of the probability of detection and then for the recurrence rates. Some of the techniques simply aim at reducing the number of parameters involved and, thus, to increase the accuracy of the estimated parameters at the possible expense of introducing bias. No formal evaluation of the trade-off

between uncertainty and bias is made here. As will be seen in Section 4.9, a quantitive assessment of uncertainty on the parameters or of goodness-of-fit of the model is difficult and computationally demanding. Instead, in application of the models to the data, values of input parameters that describe prior information are based on an informal examination of the goodness-of-fit and prior knowledge on the values of the parameters.

4.7.2 Prior Information on the Probability of Detection

As previously shown in Section 4.6, for two of the three models proposed for the estimation of P_D, the absolute value of P_D cannot be determined without additional information or constraints: In model A, the loss of reports due to imperfect transmittal remains undefined, although the probability of reporting by people and instruments can be theoretically determined from the data only. In model C-D, only the relative variation of P_D with time and magnitude m can be inferred from the data. Because the distribution of the counts as a function of m is also regulated by the parameters $b_{\underline{x}}$, it is clear that the values of P_D need to be constrained for at least two categories (t,m). Finally, although estimates in model B are uniquely defined by the data (basically through comparison for each time-magnitude category of the number of events reported by instruments, by people or by both instruments and people), uncertainty on the estimates can be large if the period of observation or the recurrence rate is small. This is true for large size measures, for early time periods (where no instruments are available for comparison with detection by people) and for some unlikely combinations of population and instrument levels (i.e. low population density and short

distance to the nearest seismic instrument). Fortunately, there is often a strong prior belief about the possibility of detection for some of the categories. For instance:

1. P_D is typically thought to be one above a given magnitude and for recent time periods.
2. All very large earthquakes are typically assumed to have been reported over most of the time span of the catalog.
3. P_D is expected to vary smoothly and monotonically as a function of time, magnitude, population density and the distance to the nearest seismic instrument.

Monotonicity of P_D has not been strictly imposed in some of the models. In fact, it is found that in recent periods the recurrence rate of events with an empirical size measure reported (here interpreted as reported by people) decreases, when an instrumental size measure is available. This is probably due to the fact that, for recent parts of the catalog, instrumental size measures have been given priority over macroseismic determinations, rather than being caused by an actual decline in the detection capability of human observers. Therefore, only the influence of fixing values of P_D or imposing smoothness on the maximum likelihood estimates is discussed next. In model D, which uses the total probability of detection, irrespective of detection mode, monotonicity has been imposed. This will be discussed separately when applying Model D in Section 4.13.

4.7.2.1 A-Priori Known Values of the Completeness Parameters

Fixing one or more of the parameters that affect the probability of detection corresponds to eliminating the corresponding maximum likelihood

equations and is therefore easily incorporated. For instance, in model A, the following is assumed: 1. There is no transmittal loss of reports of any size since 1950, 2. All events with epicentral intensity I_0=VIII on a Modified Mercalli scale are assumed reported by both people and instruments over the entire time span of the catalog, without loss of reports. In terms of the parameters of the model, this means:

$$\begin{aligned} &\alpha_{pm} = \gamma_{dm} = \beta_{tm} = 1 && \text{for all p,d and t,} \\ &&& \text{and for m = VIII} \\ &\beta_{tm} = 1 && \text{for all m and for} \\ &&& \text{time categories t after 1950} \end{aligned} \qquad (4.57)$$

Similar constraints are used in the other models and will be mentioned in the application sections. In general, constraints are imposed for the highest size measure throughout the entire time span of the catalog, because for strong events the counts are very small and, consequently, the estimates are unreliable, if one does not use additional information. The earthquake magnitudes for which P_D should be fixed to 1 in recent times depends on the quality of the seismic network. Whatever assumptions one makes on P_D, such assumption should be verified against the data, for example by comparing actual with predicted counts in categories with fiexd P_D.

4.7.2.2 Smoothness Conditions on the Variation of P_D

As the number of detection categories increases, the estimates of the completeness parameters inevitably become more uncertain, because the count in each detection category decreases. This is a commonly encountered problem in the area of probability density estimation, for which numerous techniques have been developed (for a general discussion in

the context of nonparametric density estimation, see Devroye and Györfi, 1985).

One method is based on the idea, that in histogram estimation, the bandwith of the intervals should be varied such that each interval contains a sufficiently large count, without grouping together regions with widely different probability density. Although such a method works well in the one-dimensional case, problems are encountered in multi-dimensional generalizations, for which one must decide on some direction of grouping. Therefore, such an idea is used only on a qualitatively in choosing a reasonably coarse discretization.

Another method which is often used is kernel-estimation with variable width. In this case, local estimates of the density are obtained as weighted averages of the surrounding counts. The weight assigned to the neighboring cells may depend on how well the local estimate is defined by its own count and on its difference with surrounding estimates. Such a method could for instance be applied to model C-D, by replacing the local m.l.e. of $\alpha_{t,m}$ in Equation 4.56 with a kernelestimate of the form

$$\alpha_{t,m} = \frac{\sum_{t',m' \in h(t,m)} K_\alpha(|t-t'|,|m-m'|) n_{t'm'}}{\sum_{t',m' \in h(t,m)} K_\alpha(|t-t'|,|m-m'|) n^*_{t'm'}} \qquad (4.58)$$

where h(t,m) defines a neighborhood of (t,m) and K_α assigns weights to the counts in neighboring categories (t',m'), depending on the "distances" $|t-t'|$ and $|m-m'|$. This technique is however not easily extended to cases when the underlying density is partially parametrized. For instance, it is all but evident how to define kernel estimates of the parameters in models A and B (see Equations 4.54 and 4.55). A method,

which is suitable for all models, is maximum penalized likelihood estimation, MPLE (Tapia and Thompson, 1978), and is discussed next.

In MPLE, a penalty term Q is added to the log-likelihood. Q is a function of the unknown parameters and thus changes the maximum likelihood solution. Depending on its form, such a term may penalize the roughness of the solution or, more generally, may penalize deviations of the parameters from estimates obtained through a simpler model. As the sample size gets larger, the penalty term becomes less important and thus asymptotic properties of the maximum likelihood solution can be preserved. On the other hand, as the sample size becomes smaller, the influence of the penalty term increases and forces the parameter estimates to coincide with the estimates from the simpler model. Examples of MPLE can be found in Good and Gaskins (1971, 1980) and Simonoff (1983).

The form of the penalty term Q is different from model to model. The basic idea however remains the same and is to impose smoothness on the variation of P_D with parameters such as t, m, p, and d. Model A, for instance, penalizes deviations from local linear interpolations; hence in the case of the parameter β, the following penalty term is added to the log-likelihood:

$$Q_\beta = - \sum_{t,m} \left[\frac{p^t_\beta}{2} \left[\beta_{tm} - \hat{\beta}^t_{tm} \right]^2 + \frac{p^m_\beta}{2} \left[\beta_{tm} - \hat{\beta}^m_{tm} \right]^2 \right] \qquad (4.59)$$

where $\hat{\beta}^t_{tm}$ and $\hat{\beta}^m_{tm}$ are interpolated values of β_{tm} using neighboring (t,m) cells:

$$\hat{\beta}^t_{tm} = \frac{1}{2} \left[\beta_{t-1,m} + \beta_{t+1,m} \right] \qquad (4.60a)$$

$$\hat{\beta}^{m}_{tm} = \frac{1}{2}\left[\beta_{t,m-1} + \beta_{t,m+1}\right] \qquad (4.60b)$$

The coefficients P^{t}_{β} and P^{m}_{β} regulate the influence of the penalty term on the estimates. For instance, if P^{t}_{β} is large, the estimates of β will vary linearly as a function of time category t. Similar penalties are used for α_{pm} and γ_{dm} in model A. To avoid boundary effects, only penalty terms for interior points are included. Because the penalty terms introduce coupling of the parameters for different values of the subscript indices, the iteration scheme to obtain the maximum-likelihood solution must be modified. As before, in each iteration each set of parameters $(\underline{a}_{x},\underline{b}_{x})$, β_{tm} , α_{pm} and γ_{dm} is estimated for given values on the other parameters. However, due to the penalty, additional iterations are necessary to obtain estimates for each set. Consider for example the parameter $\beta_{t'm'}$. If all other parameters, including β_{tm} for $t \neq t'$ and $m \neq m'$, are fixed, then the penalty is given by Equation 4.59 with the summation limited to terms that contain $\beta_{t'm'}$. This means that the maximum likelihood equation of $\beta_{t'm'}$ is modified by an additional linear term in $\beta_{t'm'}$. In this case, solution is easy. As before, unconditional estimates are obtained by iteration.

In model B, penalty terms have been included only for α_{tp}, because the variation of β_{t*} and γ_{d} was found to be monotonic and sufficiently smooth without any penalty. An expression of the type in Equation 4.59 is used to define penalty functions of α_{tp}, except that interpolated values are calculated using the logits α' of α, i.e using

$$\alpha'_{tp} = \ell n \frac{\alpha_{tp}}{1-\alpha_{tp}} \qquad (4.60)$$

The assumption that α'_{tp} rather than α_{tp} should be linear appears reasonable since it enforces smoother variation for α close to zero or one. (In model A, estimates are of course restricted to be between 0 and 1 and, if outside this region, they are set equal to the appropriate boundary value). In this case, the partial derivative of the penalized likelihood with respect to a given parameter $\alpha_{t'p'}$ is no longer simple, because some of the interpolated values depend on the logistic transformation of $\alpha_{t'p'}$. The iteration scheme is therefore revised as follows: First, one considers not only penalties for interior points, but also for boundary points using an appropriate extrapolation formula to calculate "interpolated" values. Second, if one keeps the interpolated values fixed, the partial derivative of the penalty term Q_α with respect to $\alpha_{t'p'}$ is simply:

$$\frac{\partial Q_\alpha}{\partial \alpha_{t'p'}} = P_\alpha[\alpha_{t'p'} - \hat{\alpha}_{t'p'}] \qquad (4.60)$$

Equation 4.60 states that the original maximum likelihood equations are modified by a linear term in $\alpha_{t'p'}$ and again solution is simple. Several iterations are of course necessary to update the interpolated values $\hat{\alpha}_{t'p'}$. Notice that it is essential to include penalty terms for parameters on the boundary, since those estimates would otherwise remain unchanged. The same scheme, with minor modifications, is used in models C and D to smooth α_{tm}. In Model C, interpolated values are calculated in the logit-scale, but using a weighted average that accounts for the expected recurrence rate in each category. For $\alpha_{t'm'}$,

$$\hat{\alpha}_{t'm'} = \frac{\sum n^*_{tm}\, \alpha_{tm}}{\sum n^*_{tm}} \qquad (4.61)$$

where n^*_{tm} is the expected count in each category and the summation extends over neighboring cells. Therefore, the interpolated value accounts for the relative uncertainty of the estimates. In particular, if one of the neighboring cells has expected count equal to zero, e.g. because a certain (time,magnitude) category is not considered in the analysis, then that cell is not used in the interpolation. This is important in models C and D, because for each magnitude, only data inside a given time interval is analyzed. Because no correction is made for boundary effects, the estimates of α_{tm} become constant as the penalty gets very high. Later in the study, it was realized that Equation 4.61 is not a very reasonable one because n^*_{tm} increases with decreasing magnitude m and, hence, the weights assigned to α_{tm} in Equation 4.61 increase with decreasing m. In order to avoid this effect, Model D uses a simple local average where the summation is limited to neighboring cells with n^*_{tm} different from zero. To correct for the fact that constant values of α_{tm} are obtained for very high penalties, lower penalty coefficients are used in this model for boundary values of α_{tm}.

4.7.3 Prior Information on the Recurrence Parameters $a_{\underline{x}}$ and $b_{\underline{x}}$

Estimation of the recurrence parameters $a_{\underline{x}}$ and $b_{\underline{x}}$ is subject to the same problems as estimation of the completeness parameters: As more locations $\underline{x}$ are considered, the uncertainty on the estimates increases and prior information on the value of the parameters or some smoothness constraints become necessary (Typical values of the uncertainty on individual a-and b-estimates are shown in Section 4.2.2.). Various forms of prior information have been considered in the different models. Before

describing each form in detail, a brief overview is given first. In model A, the assumption of seismogenic provinces is used and, hence, there is less concern about uncertainty on the estimated parameters, provided that the provinces are sufficiently large. A frequent assumption in practice is that, while $a_{\underline{x}}$ varies from province to province, the slope parameter $b_{\underline{x}}$ is the same everywehere. This assumption, as well as a less restrictive alternative is included as optional choices in model A. The alternative assumption is that the parameters $b_{\underline{x}}$ are independent relatizations of a random variable with unknown mean and variance, and are therefore informative one on the others. In model B, uncertainty on $a_{\underline{x}}$ and $b_{\underline{x}}$ is a more serious concern, because a more refined spatial grid is used. The method of estimation for Model B is MPLE, i.e. a method similar to that used for the completeness parameters. In model C an alternative technique, based on direct smoothing of the counts and similar to kernel-estimation, is explored. This method is computationally much simpler, but unfortunately does not appear to generalize easily to the case where the probability of detection varies with location. A possible solution to this problem will be indicated in Section 4.7.3.5. Finally, model D uses again a MPLE method, but employs a different solution technique and a different form of the penalty. The way in which these forms of prior information are included in maximum penalized likelihood estimation is discussed next.

4.7.3.1 Identical values of $b_{\underline{x}}$

The assumption that the parameters $b_{\underline{x}}$ are constant inside seismogenic provinces S_i is easily accounted for. Because the original log-likelihood

is additive for different $\underline{x}$ and because, for each $\underline{x}$ in S_i, $b_{\underline{x}}$ is now replaced with a single parameter b_i, the partial derivative of the log-likelihood with respect to b_i is simply the sum of all partial derivatives with respect to $b_{\underline{x}}$, for $\underline{x}\varepsilon S_i$. The maximum likelihood equations for $\underline{x}\varepsilon S_i$ are then replaced with the single equation:

$$- \sum_{\underline{x}\varepsilon S_i} m_{\underline{x}} - \sum_{\underline{x}\varepsilon S_i} \sum_{m} T^*_{\underline{x}m}\, m \exp\{a_{\underline{x}} - b_i m\} = 0 \qquad (4.62)$$

Both Equation 4.62 and the maximum likelihood equations for $a_{\underline{x}}$ are coupled to b_i and nonlinear in b_i and $a_{\underline{x}}$. Their solution could again be obtained using Newton's method, but such a method involves the inversion of the Jacobian, which has dimension equal to the number of spatial cells $\underline{x}$ in S_i plus one. As a better alternative, the solution technique used in Model A is to solve each equation for one parameter in turn, while fixing all other parameters. Convergence to the maximum likelihood solution is of course somewhat slower in this case.

4.7.3.2 Parameters $b_{\underline{x}}$ that are Realizations of the Same Random Variable

Suppose that instead of being identical, the parameters $b_{\underline{x}}$ in S_i are independent realizations of a random variable with normal distribution $N(m_{B_i}, \sigma^2_{B_i})$, in which the mean value m_{B_i} and variance $\sigma^2_{B_i}$ are unknown. In this case, the catalog data can be used to estimate not only the parameters $b_{\underline{x}}$ but also the distribution parameters m_{B_i} and $\sigma^2_{B_i}$. Such a technique is called empirical Bayes, because the prior distribution of each $b_{\underline{x}}$ is determined empirically. The log-likelihood in Equation 4.45 is now modified by the following additive term:

$$- \frac{1}{2\sigma_{B_i}^2} \sum_{x \epsilon S_i} (b_{\underline{x}} - m_{B_i})^2 - \sum_{\underline{x} \epsilon S_i} \ln \sigma_{B_i} \tag{4.63}$$

It follows that the maximum-likelihood equations for $b_{\underline{x}}$ should include the additional term

$$- \frac{1}{\sigma_{B_i}^2} (b_{\underline{x}} - m_{B_i}) \tag{4.64}$$

The maximum likelihood equations of m_{B_i} and $\sigma_{B_i}^2$ are obtained by setting to zero the partial derivatives of the log-likelihood term in Equation 4.63 with respect to these parameters. Hence, the following additional equations must be satisfied:

$$\sum_{\underline{x} \epsilon S_i} b_{\underline{x}} - n_{S_i} m_{B_i} = 0 \tag{4.65}$$

$$\sum_{\underline{x} \epsilon S_i} (b_{\underline{x}} - m_{B_i})^2 - n_{S_i} \sigma_{B_i}^2 = 0 \tag{4.66}$$

where n_{S_i} is the number of discrete locations $\underline{x}$ in S_i. If $b_{\underline{x}}$ were known, then Eqs. 4.65 and 4.66 would correspond to the usual maximum-likelihood conditions for the mean and variance of a Gaussian distribution. Once again, the solution for $a_{\underline{x}}$, $b_{\underline{x}}$, m_{B_i} and $\sigma_{B_i}^2$ can be found by iteration.

It is worth mentioning, however, that in the present case the log-likelihood L has a rather peculiar behavior: Let $L(\sigma_{B_i}^2)$ denote the maximum of L for given $\sigma_{B_i}^2$. Then $L(\sigma_{B_i}^2)$ does not necessarily have a point of stationarity, implying that the previous equations may have no solution. In addition, one can easily show that the limit of L for $\sigma_{B_i}^2 \to 0$ equals ∞. Two possible situations are exemplified in Fig. 4.8.

In Figure 4.8b, a point of local maximum of $L(\sigma^2_{B_i})$ exists for $\sigma^2_{B_i} = \sigma^2_{B_i\rho}$. This value of the variance is associated with unequal estimates of the slopes $b_{\underline{x}}$, which are more clustered than the unconstrained estimates from Eqs. 4.49. Clustering is towards the group average m_{B_i} and is more pronounced for smaller $\sigma^2_{B_i}$ and for locations $\underline{x}$ with a smaller number of events. Figure 4.8a illustrates the case when $L(\sigma^2_{B_i})$ has no point of stationarity. This happens when the unrestricted estimates of $b_{\underline{x}}$ from Equations 4.49 are already close one to another, relative to their estimation variances. In this case, the solution is identical to that for $b_{\underline{x}} \equiv b_i$. Finally, if the slopes $b_{\underline{x}}$ are treated as nuisance parameters, then the marginal likelihood (i.e. the function obtained by integrating the log-likelihood with respect to $b_{\underline{x}}$) should be used to estimate m_{B_i} and $\sigma^2_{B_i}$. Individual values of $b_{\underline{x}}$ can be obtained afterwards based on the posterior density of $b_{\underline{x}}$ for given m_{B_i} and σ_{B_i} (for instance, by maximizing the posterior density). In this case all likelihood functions would be well behaved. Calculation of the marginal likelihood is however not straightforward and the former technique of directly maximizing the likelihood function is preferred here for numerical implementation.

4.7.3.3 Independent Prior Information on Values of $b_{\underline{x}}$

In some cases, independent information exists on the value of $b_{\underline{x}}$. For example, such information may reflect the distribution of b for world wide or regional earthquake data. Lower-and upper-bounds for $b_{\underline{x}}$ can be incorporated in the analysis by solving each of the maximum likelihood equations separately and, instead of using a Newton-Raphson method, by iteratively decreasing the interval which contains the maximum likelihood

estimate. This option has been included in models C and D. In the latter model it was found more convenient to use Newton's method to calculate increments of a and b inside the feasible region. If these increments predict values of b outside the feasible region, the appropriate boundary value is used and the increment of a is recalculated using Newton's method for the maximum likelihood equation of a only.

Including a prior distribution of $b_{\underline{x}}$ with given parameters $(\tilde{b}, \sigma^2_{\tilde{b}})$ is also easy. Each maximum likelihood equation should in this case include the additional term

$$-\frac{1}{\sigma^2_{\tilde{b}}}\left(b_{\underline{x}} - \tilde{b}\right) \tag{4.67}$$

Again, the conditional likelihood equations are easily solved, by either Newton's method or interval reduction.

Two problems that arise in the specification of independent priors of $b_{\underline{x}}$ should be pointed out. First, $\sigma^2_{\tilde{b}}$ is the variance of the slope $b_{\underline{x}}$ averaged within a given neighborhood of $\underline{x}$. If the area of the neighborhood varies (in the limit, $\underline{x}$ might be associated with an entire seismogenic province) then also $\sigma^2_{\tilde{b}}$ should change. If this were not the case, the prior would become very strong compared to information from the data as the area associated with each $\underline{x}$ decreases. One should also be careful not to mix two arguments: 1. $\sigma^2_{\tilde{b}}$ is the variance of an average value and changes as the area associated with $\underline{x}$ changes, 2. the influence of $\sigma^2_{\tilde{b}}$ depends on the earthquake count $n_{\underline{x}}$ used in the maximum likelihood equation for $b_{\underline{x}}$. Because of the second argument the solution $b_{\underline{x}} \equiv b$ for $\underline{x}$ associated with small areas is a correct one, if all $b_{\underline{x}}$ are estimated independently and are associated with small counts. It will be shown

later that as the spatial discretization becomes more refined, there is an increasing need to smooth $b_{\underline{x}}$ to obtain reliable estimates. Hence neighboring estimates are increasingly dependent, which reduces the influence of the prior. With respect to the first argument it is assumed in the applications that a priori the parameters $b_{\underline{x}}$ are mutually independent. This assumption is consistent with decreasing $\sigma_{\tilde{b}}^2$ proportionally to the area considered. A second characteristic of an independent prior is that the global maximum likelihood equation for b (i.e. summed over all $\underline{x}$) is no longer satisfied. Thus one may find that the total expected magnitude no longer equals the total observed count.

4.7.3.4 Penalized Maximum Likelihood Estimation

Similar to the estimation of completeness parameters, penalized maximum likelihood estimation can be used to introduce smoothness in the spatial variation of $a_{\underline{x}}$ and $b_{\underline{x}}$ in order to reduce the statistical uncertainty on individual estimates. Penalties can also be interpreted as priors on the function $a_{\underline{x}}$ and $b_{\underline{x}}$ based on a single parameter of these distributions such as the roughness. Technically, the interpretation of the penalties makes no difference. Because in application the influence of the penalties is regulated in an interactive manner, i.e. by visual examination of the results for different penalty coefficients, it is perhaps most appropriate to interpret the technique as a pragmatic way to reduce the number of degrees of freedom of the model. The basic form of the penalty term used in the models is the same and penalizes deviations of the local estimates $a_{\underline{x}}$, $b_{\underline{x}}$ from more global estimates $\hat{a}_{\underline{x}}$, $\hat{b}_{\underline{x}}$ obtained by local averaging or interpolation. Because of problems to calculate the

MPL solution for high values of the penalties, different solution techniques have been used. These techniques will be explained next, first for model B and then for model D.

The penalty term $Q_{a,b}$ which is added to the log-likelihood is of the following form in model B:

$$Q_{a,b} = -\frac{P_a}{2}\sum_{\underline{x}}(a_{\underline{x}} - \hat{a}_{\underline{x}})^2 - \frac{P_b}{2}\sum_{\underline{x}}(b_{\underline{x}} - \hat{b}_{\underline{x}})^2 \qquad (4.68)$$

where $\hat{a}_{\underline{x}}$ and $\hat{b}_{\underline{x}}$ are interpolated values of $a_{\underline{x}}$ and $b_{\underline{x}}$ respectively. The summation extends over all $\underline{x}$, including boundary cells where an appropriately modified interpolation formula needs to be used. If deviations from a locally constant level need to be penalized, interpolation can be done using locally weighted averages. If a locally linear variation of $a_{\underline{x}}$ or $b_{\underline{x}}$ is allowed, $\hat{a}_{\underline{x}}$ and $\hat{b}_{\underline{x}}$ can be calculated by fitting a local linear regression to neighboring values. The first approach is evidently simpler and, if the averages are sufficiently local, can also capture linear trends over larger region. In practical applications, it was also found that the second approach is not always stable for complicated geometries at the boundaries. The modified likelihood equations in model B are obtained by assuming that $\hat{a}_{\underline{x}}$ and $\hat{b}_{\underline{x}}$ in Equation 4.68 is fixed. Of course, iteration is then necessary to update $\hat{a}_{\underline{x}}$ and $\hat{b}_{\underline{x}}$ for changes in $a_{\underline{x}}$ and $b_{\underline{x}}$. Under those conditions, the maximum likelihood equations are of a simple form and can be easily solved. For instance, the MPL equation for $a_{\underline{x}}$ is

$$n_{\underline{x}} - \sum_{m} T^*_{\underline{x}m} \exp\{a_{\underline{x}} - b_{\underline{x}}m\} - P_a(a_{\underline{x}} - \hat{a}_{\underline{x}}) = 0 \qquad (4.69)$$

A special condition on the interpolators $a_{\underline{x}}$ should be noted. If one applies MPL to estimate the spatial variation of the recurrence realtionship inside a given region, it may be desirable that the total expected and observed counts inside that region be the same. That is

$$\sum_{\underline{x}} n_{\underline{x}} - \sum_{\underline{x}} \sum_{m} T^{*}_{\underline{x}m} \exp\{a_{\underline{x}} - b_{\underline{x}}m\} = 0 \qquad (4.70)$$

where $T^{*}_{\underline{x}m}$ is the equivalent period of completeness as derived in Equation 4.46. It follows that the interpolators should satisfy the condition

$$\sum_{\underline{x}} a_{\underline{x}} = \sum_{\underline{x}} \hat{a}_{x} \qquad (4.71)$$

The same requirement holds for the $\hat{b}_{\underline{x}}$ interpolator meaning that $\sum_{\underline{x}} b_{\underline{x}}$ should equal $\sum_{\underline{x}} \hat{b}_{\underline{x}}$.

Eq. 4.71 can be easily satisfied by calculating first the interpolated values $\hat{a}_{\underline{x}}$ from $a_{\underline{x}}$ and next by adding a constant to correct for any imbalance in Equation 4.71. This technique has been used in model B. The problem with such a technique is that for high values of P_a and P_b in Equation 4.68, convergence of the maximum likelihood algorithm is very slow if a large number of locations $\underline{x}$ are used. This is due to the fact that coupling between $\hat{a}_{\underline{x}}$ and $a_{\underline{x}}$ is not recognized in each iteration. For instance, for fixed $\hat{a}_{\underline{x}}$ and large P_a, changes to individual estimates $a_{\underline{x}}$ are extremely small, although it is possible that a relatively large global change of all $a_{\underline{x}}$ is necessary to converge to the maximum likelihood solution. This problem can be partly corrected for by using initial estimates which are constant and satisfy the global maximum likelihood equation, i.e. Equation 4.70 for $a_{\underline{x}}$. However, if a linear

trend, which receives little or no penalty, is present in the data, convergence may be very slow.

Because of this convergence problem, the solution technique in model D has been modified and made more explicit. The penalty on $a_{\underline{x}}$ and $b_{\underline{x}}$ is written in Model D as

$$Q_{a,b} = -\frac{P_a}{2}\,[a_{\underline{x}}]^T[I-H]^T[I-H][a_{\underline{x}}] - \frac{P_b}{2}\,[b_{\underline{x}}]^T[I-H]^T[I-H][b_{\underline{x}}] \qquad (4.72)$$

where $[a_{\underline{x}}]$, $[b_{\underline{x}}]$ are column vectors, superscript T indicates transposed matrices or vectors, I is the identity matrix and H is an interpolator matrix such that

$$[\hat{a}_{\underline{x}}] = [H][a_{\underline{x}}] \qquad (4.73)$$

Notice that the same interpolator is used for $\hat{b}_{\underline{x}}$ and that the degree of smoothness of $a_{\underline{x}}$, $b_{\underline{x}}$ is regulated by the penalty coefficients P_a and P_b.

Evidently, Equation 4.72 is equivalent to Equation 4.68. The likelihood equation one solves in each iteration is however quite different, if one considers $\hat{a}_{\underline{x}}$ as an explicit function of $a_{\underline{x}}$. For instance, Equation 4.69 changes to

$$n_{\underline{x}} - \sum_m T^*_{\underline{x}m} \exp\{a_{\underline{x}} - b_{\underline{x}}m\} - P_a[W]_{\underline{x}}[a_{\underline{x}}] = 0 \qquad (4.74)$$

where $[W]_{\underline{x}}$ is the $\underline{x}$'th row of the matrix $W = [I-H]^T[I-H]$. Again, Equation 4.70 needs be satisfied, which imposes the following condition on W:

$$[1]^T[W][a_{\underline{x}}] = 0 \qquad (4.75)$$

It is interesting to note that, for a proper choice of the interpolator matrix [H], Equation 4.75 is satisfied independently of the value of $[a_{\underline{x}}]$.

For instance, a natural condition for [H] is that

$$[1] = [H][1] \tag{4.76}$$

It follows then immediately that Equation 4.75 is always satisfied since

$$[1]^T[W] = [1]^T[I-H]^T[I-H] = [0] \tag{4.77}$$

This is not true in Equation 4.69 beacuse $[1]^T[a_{\underline{x}} - a_{\underline{x}}]$ is not necessarily zero when H is not symmetric for locations $\underline{x}$ on the boundary. The interpolator choosen in model D has the simple form

$$\hat{a}_{\underline{x}} = \frac{1}{k_{\underline{x}}} \sum_{\underline{x} \epsilon N(\underline{x})} a_{\underline{x}} \tag{4.78}$$

where $N(\underline{x})$ is the set of locations that are neighbors of $\underline{x}$ and $k_{\underline{x}}$ equals the number of neighbors. Equation 4.78 alllows one to express the various terms in [W] as simple functions of $k_{\underline{x}}$ for all $\underline{x}$. Omitting the details of the derivation, one finds that

$$w_{\underline{xx}} = 1 + \sum_{\underline{y} \epsilon N(\underline{x})} \left(\frac{1}{k_{\underline{y}}}\right)^2 \tag{4.79}$$

$$w_{\underline{xy}} = -\frac{1}{k_{\underline{x}}} - \frac{1}{k_{\underline{y}}} + \sum_{\underline{z} \epsilon N(\underline{x}) N(\underline{y})} \left(\frac{1}{k_{\underline{z}}}\right)^2$$

$$w_{\underline{xt}} = - \sum_{\underline{y} \epsilon N(t)} \left(\frac{1}{k_{\underline{y}}}\right)^2$$

where $\underline{y}$ indicates locations that belong to $N(\underline{x})$ and $\underline{t}$ indicates locations that belong to $N(\underline{y})$ but not to $N(\underline{x})$.

Solution of Equations 4.74 for each $\underline{x}$ must again proceed by iteration. One way to do so would be to calculate the inverse of the

Jacobian and to use Newton's method. This is however not a very practical method if the number of locations $\underline{x}$ is large and again an iterative scheme is used, which works as follows:

1. Select initial values of $a_{\underline{x}}$ and $b_{\underline{x}}$.
2. Solve separately for each $\underline{x}$ and update immediately all coupled equations to account for changes in the penalty term, e.g. $- P_a w_{\underline{yx}} \Delta a_{\underline{x}}$ and $- P_a w_{\underline{tx}} \Delta a_{\underline{x}}$.
3. After solving the equations in the entire region, calculate the total imbalance for the maximum likelihood equations and add constants Δa and Δb to all $a_{\underline{x}}$ and $b_{\underline{x}}$ to remove this imbalance.
4. Continue with 2.

Although no formal comparison is made of the solution techniques used in models B and D, the last one appears to be much more efficient. However, in some cases convergence is still found to be slow.

For small regions, it may be reasonable to penalize deviations of $a_{\underline{x}}$ and $b_{\underline{x}}$ from constant levels independent of $\underline{x}$, rather than allowing for a linear trend. In this case, the penalty in Equation 4.68 simplifies to

$$Q_{a,b} = - \frac{P_a}{2} \sum_{\underline{x}} (a_{\underline{x}} - \hat{a})^2 - \frac{P_b}{2} \sum_{\underline{x}} (b_{\underline{x}} - \hat{b})^2 \qquad (4.80)$$

where $\hat{a}$ and $\hat{b}$ are necessarily global averages of $a_{\underline{x}}$ and $b_{\underline{x}}$ so that Eq. 4.71 is satisfied. If the penalty terms are large, the penalized maximum likelihood solution converges to the solution found in a traditonal zonation method, i.e. $a_{\underline{x}} \equiv \tilde{a}$ and $b_{\underline{x}} \equiv \tilde{b}$. For low values of the penalty coefficients, a local solution is found.

4.7.3.5 Smoothing of the Counts

Because of the computational difficulties of MPL, it is of interest to consider alternative techniques to obtain spatially smooth estimates of $a_{\underline{x}}$ and $b_{\underline{x}}$. An intuitive and simple way to do so is to smooth the data, i.e. the earthquake counts, prior to the estimation. In general, this leads to estimators of a kernel-type. In the present case, the formulation of a kernelestimator is however not evident because smoothness of the a-and b-parameters is required rather than smoothness of the counts. To illustrate this problem, reconsider first the maximum likelihood equations for $a_{\underline{x}}$ and $b_{\underline{x}}$ in Equations 4.48 and 4.49. After eliminating $a_{\underline{x}}$, Equation 4.49 can be written as:

$$-m_{\underline{x}} + n_{\underline{x}} \frac{\sum T^*_{\underline{x}m}\, m \exp\{-b_{\underline{x}} m\}}{\sum T^*_{\underline{x}m}\, m \exp\{-b_{\underline{x}} m\}} = 0 \qquad (4.81)$$

It is clear that spatial smoothness of $b_{\underline{x}}$ is related to smoothness of $m_{\underline{x}}$ and $n_{\underline{x}}$, but also depends on the spatial variation of $T^*_{\underline{x}m}$. For instance, smooth estimates of $b_{\underline{x}}$ could be found by replacing $m_{\underline{x}}$, $n_{\underline{x}}$ and $T^*_{\underline{x}m}$ with smoothed values, calculated as

$$m^b_{\underline{x}} = \sum_{\underline{y}} K_b(\|\underline{x} - \underline{y}\|)\, m_{\underline{x}} \qquad (4.82a)$$

$$n^b_{\underline{x}} = \sum_{\underline{y}} K_b(\|\underline{x} - \underline{y}\|)\, n_{\underline{x}} \qquad (4.82b)$$

$$T^{*b}_{\underline{x}m} = \sum_{\underline{y}} K_b(\|\underline{x} - \underline{y}\|)\, T^*_{\underline{x}m} \qquad (4.82c)$$

where K_b is a kernel function, the value of which depends on the distance from $\underline{x}$, e.g. $\|\underline{x}-\underline{y}\|$.

The corresponding estimate of $b_{\underline{x}}$ is then found from

$$-m^b_{\underline{x}} + n^b_{\underline{x}} \frac{\sum T^{*b}_{\underline{xm}} \; m \; \exp\{-b_{\underline{x}} m\}}{\sum T^{*b}_{\underline{xm}} \; m \; \exp\{-b_{\underline{x}} m\}} = 0 \qquad (4.83)$$

A similar analysis for $a_{\underline{x}}$ immediately shows that, because $b_{\underline{x}}$ is initially unknown, there is no simple way to impose smoothness on $a_{\underline{x}}$. On the other hand, one should recognize that if one wants to impose smoothness on the spatial variation of the cumulative count $\sum_m \exp\{a_{\underline{x}}-b_{\underline{x}}m\}$, this poses no problem. In this case, a different kernel function K_a must be used to allow different smoothness of $b_{\underline{x}}$ and $a_{\underline{x}}$. If one defines

$$n^a_{\underline{x}} = \sum_{\underline{y}} K_a\left(\|\underline{x} - \underline{y}\|\right) n_{\underline{x}} \qquad (4.84a)$$

$$T^{*a}_{\underline{xm}} = \sum_{\underline{y}} K_a\left(\|\underline{x} - \underline{y}\|\right) T^*_{\underline{xm}} \qquad (4.84b)$$

then the estimate of the a-parameter is found from

$$n^a_{\underline{x}} = \sum_m T^{*a}_{\underline{xm}} \exp\{a_{\underline{x}}-b_{\underline{x}}m\} = 0 \qquad (4.85)$$

Equations 4.83 and 4.84 can be solved using the techniques discussed earlier in Section 4.2.2. The problem with such a solution is that the total expected count and the total expected magnitude in the region do not equal the corresponding observed values. Conditions on K_a and K_b to satisfy this requirement can be derived by substituting estimates $a_{\underline{x}}$, $b_{\underline{x}}$

as defined by Equations 4.83 and 4.84 into the global maximum likelihood equations. For the general case, this leads to complicated expressions and the approach is only illustrated here for the special case when $T^*_{\underline{x}m}$ does not depend on $\underline{x}$. Then Equations 4.83 and 4.84 can be rewritten as:

$$-m^b_{\underline{x}} + n^b_{\underline{x}} \frac{\sum_m T^*_m \; m \exp\{-b_{\underline{x}} m\}}{\sum_m T^*_m \exp\{-b_{\underline{x}} m\}} = 0 \tag{4.86}$$

$$n^a_{\underline{x}} - \sum_m T^*_m \exp\{a_{\underline{x}} - b_{\underline{x}} m\} = 0 \tag{4.87}$$

Conditions for the global maximum likelihood equations are in this case

$$-\sum_{\underline{x}} m_{\underline{x}} + \sum_{\underline{x}} \sum_m T^*_m \; m \exp\{a_{\underline{x}} - b_{\underline{x}} m\} = 0 \tag{4.88}$$

$$-\sum_{\underline{x}} n_{\underline{x}} + \sum_{\underline{x}} \sum_m T^*_m \exp\{a_{\underline{x}} - b_{\underline{x}} m\} = 0 \tag{4.89}$$

Equations 4.87 and 4.89 lead to the condition that

$$\sum_{\underline{x}} n^a_{\underline{x}} = \sum_{\underline{x}} n_{\underline{x}} \tag{4.90}$$

and Equations 4.86, 4.87 and 4.88 impose in addition that

$$\sum_{\underline{x}} \frac{m^b_{\underline{x}}}{n^b_{\underline{x}}} n^a_{\underline{x}} = \sum_{\underline{x}} m_{\underline{x}} \tag{4.91}$$

Eq. 4.90 simply requires that after smoothing the total count should be preserved. Equation 4.91 implies that the weighted sum of $m^b_{\underline{x}}$, with weights $n^a_{\underline{x}}/n^b_{\underline{x}}$, should equal the total observed magnitude, and is less

intuitive. Both requirements can be easily satisfied by adding a constant term to each $n^a_{\underline{x}}$ and $m^b_{\underline{x}}/n^b_{\underline{x}}$. This technique is used in model C. Generalization of Equations 4.90 and 4.91 to the case when $T^*_{\underline{x}m}$ is not independent of $\underline{x}$ is not evident.

An alternative solution to satisfy Equations 4.88 and 4.89 is to consider an additional variable a and b, such that

$$a_{\underline{x}} = a_{\underline{x}} + a \tag{4.92}$$

$$b_{\underline{x}} = b_{\underline{x}} + b \tag{4.93}$$

In that case, a and b can be determined such that Equations 4.88 and 4.89 are always satisfied, without changing the relative smoothness of the solution.

4.8 MAXIMUM LIKELIHOOD ESTIMATION OF PROBABILITY OF DETECTION AND RECURRENCE RATES INCLUDING ERRORS IN THE DATA

4.8.1 Introduction

So far, no attention has been paid to the fact that, in reality, the values $(\underline{x}_i, t_i, m_i)$ for each earthquake i are uncertain. Whereas the time of occurrence t_i is usually sufficiently accurate for the present purpose, geographical location $\underline{x}_i$ and size measure m_i may be subject to large errors, especially for early events. The importance of this problem is well illustrated by earthquake data for the Friuli region in Northern Italy (Figure 4.9): In this catalog, location uncertainty for each earthquake has been indicated through a categorical variable, which is associated with a certain maximum radius of uncertainty as shown in Table

4.3. The size measure I_o of practically all earthquakes in the catalog is reported in the Modified Mercalli Intensity scale and, as in the Chiburis catalog, two alternative values are given. The difference between the two values can be used as a measure of uncertainty on I_o. To represent the data in figures and tables, a single value of I_o is choosen as:

$$I_o = \text{nearest integer } [(I_{o1} + I_{o2})/2] \tag{4.94}$$

The data have also been analyzed for magnitude conversion and clustering. It would lead us to far to comment on this particular application and further information can be found in reports by Veneziano and Van Dyck (1985a, 1985b). Here, only the distribution of main events will be discussed. Figure 4.10 presents an exploratory analysis of the catalog data and illustrates the significance of location uncertainty. Similar to the plots used in the exploratory analysis of the Chiburis data (Section 3.5), Fig. 4.10a shows two-dimensional scatter diagrams of $(\underline{x}_i, t_i, m_i)$ for all earthquakes. Figures 4.10b to 4.10e present similar plots, each for a different value of the uncertainty on location i_L. Evidently, accurately located earthquakes (Fig. 4.10b) are very few and are found only in recent time periods. For recent earthquakes, the most common value of i_L is 3 (Fig. 4.10c), which corresponds to a maximum radius of uncertainty less than 20 km. Few such earthquakes are found prior to 1850. Figures 4.10d and 4.10e indicate clearly that location uncertainty for earthquakes in early periods of the catalog is substantial. Moreover, one may notice that many of the earthquakes occur at particular locations. This is not by accident! Comparison with Fig. 4.9 shows that several locations correspond to major cities. Other popular locations correspond to rounded-off values of latitude and longitude. Going back to Fig. 4.10a,

one can distinguish roughly four periods: For the first two and a half centuries, activity is reported exclusively in the southwest region of Venice, Padova, and Vicenza, whereas between 1250 and 1700 activity is reported also in the north, near the town of Gemona. The third period, from 1700 to about 1870, is one of transition: seismicity spreads more evenly in space, with a trend of the larger events to migrate to the north. Finally, after about 1870, reported seismicity has been essentially confined to latitudes north of 45.45 N. There are several possible explanations for the redistribution of events in space: One is that seismicity in Friuli is highly nonstationary, with strong migratory episodes over periods of one or very few centuries. An alternative explanation is that seismicity is (approximately) stationary and the observed spatial and temporal patterns are due to catalog incompleteness. The latter hypothesis would explain the increase of reported activity in the northern mountain area, but not the recent reduction of activity in the plains. A third and more plausible explanation is that the spatial pattern of reported events reflects more the location of "observers" near the epicenters than the location of the epicenters themselves. This would explain why, in earlier times, earthquakes are reported to have occurred at the site of large cities. Errors in the location of epicenters and the reduction of such errors in recent times explain both the increase of activity in the north and the simultaneous decrease of activity in the southwest, hence the apparent migration of epicenters in Figs. 4.10b to 4.10e.

The importance of uncertainty on the size measures, as well as that of location uncertainty can be also judged from Table 4.4. This table

shows, for the different seismic sources indicated in Fig. 4.10f, the total earthquake counts cross-classified according to time t, location uncertainty (denoted by UL), the difference between the two estimates of intensity dI_o, and the average intensity I_o. Source 1 has few earthquakes and is not important. Source 2 corresponds to a region that has been recently more active. Notice that the number of earthquakes with $dI_o \neq 0$ is quite large, also in recent times. As one might have expected, large values of dI_o tend to be associated with large values of UL. Source 3 contains a major part of the early, inaccurately located earthquakes; this is shown by the large number of events with high values of UL. Also, the fraction of earthquakes with $dI_o \neq 0$ is larger than in Source 2.

It follows from the previous discussion that, without consideration of uncertainty on location and the size measure, predicted recurrence rates may be substantially biased. Earlier in Section 2.5, a correction to account for uncertainty on I_o was derived, which basically replaces I_o with the expected value of its a-posteriori distribution when the slope parameter of the exponential recurrence relation is known. Such a correction is not easily extended to the uncertainty on earthquake location. In this section, a more general and theoretically satisfactory treatment of uncertainty on data is given, based on an extension of the maximum likelihood formulation of Section 4.6. In Section 4.8.2 the necessary modification is derived in a general form and a practical solution technique is discussed. Section 4.8.3 discusses the modification to the maximum likelihood solution when the prior distribution of $\underline{x}$ or m falls outside the domain of interest in the analysis. Application of

these techniques will be presented later in Sections 4.12 and 4.13 for Model C and Model D, respectively.

4.8.2 Maximum Likelihood Formulation Considering Errors in the Data

From a statistical point of view, the present problem is similar to that of estimating Poisson rates for the cells of a multi-way contingency table when the data is erroneously classified. Problems of this general type arise often in practice and have been studied in the statistical literature under the name of "missing categorical data". However, only in a few studies is the misclassification probability allowed to vary from observation to observation; examples are Press (1968), Pregibon (1977), Little (1982), and Nordheim (1984). This is clearly the case in our problem, because uncertainty on the correct category c varies from earthquake to earthquake. In some analyses (Pregibon, Nordheim) the misclassification probabilities are assumed to depend on the true class c to which the individual (here, the earthquake) belongs, while in others (Press) the same probabilities may vary from individual to individual. The formulation given in this section is fundamentally similar to that of Press (1968), except that Press estimates cell probabilities rather than Poisson rates and his model is a saturated one.

To derive a general formulation of the maximum likelihood accounting for errors in the data, the notation of Section 4.6 will be used. In the present case, one should consider however that the category (c_i, z_i) to which the i'th earthquake belongs is initially unknown and needs to be estimated. It is assumed that based on information other than regional seismicity a prior distribution $P_c^{i,z}$ is given for each earthquake i such that

$$P[(c_i = c) \quad (z_i = z)] = p_c^{i,z} \tag{4.95}$$

where z_i is the mode of detection of the i'th earthquake.

When summed over all categories, this probability should equal one, i.e.

$$\sum_{c,z} p_c^{i,z} = 1 \tag{4.96}$$

To incorporate this information into the likelihood, the log-likelihood function in Eq. 4.45 should be written first in terms of the set of unknown categories $\{c_i, z_i\}$. To do so, it is convenient to introduce an indicator variable $\delta_c^{i,z}$ for each earthquake such that

$$\begin{aligned} \delta_c^{i,z} &= 1 \quad \text{for } c = c_i,\ z = z_i \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{4.97}$$

The various counts used in Equation 4.45 are easily related to $\{\delta_c^{i,z}\}$ as follows:

$$n_c^z = \sum_i \delta_c^{i,z} \tag{4.98a}$$

$$n_{\underline{x}} = \sum_i \delta_{\underline{x}}^i \tag{4.98b}$$

$$m_{\underline{x}} = \sum_i \sum_m m\ \delta_{\underline{x},m}^i \tag{4.98c}$$

where the usual convention is used that omission of subscripts or superscripts indicates summation over the missing indices. For instance

$$\delta_{\underline{x}}^i = \sum_{z,D,m} \delta_c^{i,z} \tag{4.99}$$

Substituting Equation 4.98 into Equation 4.45 and using also Equation 4.97 leads to the following intuitive expression of the log-likelihood for <u>given</u> true locations of the earthquakes

$$\ell n \ell(a_{\underline{x}}, b_{\underline{x}}, \theta | \{c_i, z_i\}) = \sum_i [\ell n \, P_{D_i}^{z_i} + a_{\underline{x}_i} - m_i \, b_{\underline{x}_i}]$$

$$- \sum_c T_c P_D \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \qquad (4.100)$$

The problem now is to modify this likelihood expression to account for the fact that the classes $\{c_i, z_i\}$ are unknown, with prior distribution given by Equation 4.95. Since the contribution of the i'th earthquake in Equation 4.100 is of the form

$$\ell_i(a_{\underline{x}}, b_{\underline{x}}, \theta | c_i, z_i) \propto P_{D_i}^{z_i} \exp\{a_{\underline{x}_i} - m_i b_{\underline{x}_i}\} \qquad (4.101)$$

these terms should be modified as

$$\ell_i(a_{\underline{x}}, b_{\underline{x}}, \theta, c_i, z_i | P_c^{i,z}) \propto P_{c_i}^{i,z_i} \, P_{D_i}^{z_i} \exp\{a_{\underline{x}_i} - b_{\underline{x}_i} m\} \qquad (4.102)$$

It is clear that the likelihood in the above form is useless for the estimation of the unknown parameters, since the number of parameters is larger than the number of data. Notice also that, for given values of $a_{\underline{x}}$, $b_{\underline{x}}$ and θ, Equation 4.102 is proportional to the posterior density of (c_i, z_i) in a Bayesian interpretation. Since the interest of the present analysis is in the estimation of $a_{\underline{x}}$, $b_{\underline{x}}$ and θ, a more useful form of the likelihood can be derived by treating (c_i, z_i) as nuisance parameters and therefore calculating the marginal likelihood ℓ_i^m as a function of $(a_{\underline{x}}, b_{\underline{x}}, \theta)$ only. From equation 4.102, it follows immediately that

$$\ell_i^m(a_{\underline{x}}, b_{\underline{x}}, \theta | P_c^{i,z}) = \sum_{c,z} q_c^{i,z} \qquad (4.103)$$

where $q_c^{i,z}$ is proportional to the posterior density of (c_i, z_i) and is defined by Equation 4.102 as:

$$q_c^{i,z} \alpha p_c^{i,z} P_D^z \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \tag{4.104}$$

The final expression for the marginal log-likelihood is obtained by combining the marginal likelihoods ℓ_i^m for all earthquakes and is given by

$$\ell n \ell^m(a_{\underline{x}}, b_{\underline{x}}, \theta | \{P_c^{iz}\}) = \sum_i \ell n\Big(\sum_{c,z} q_c^{i,z}\Big) - \sum_c T_c P_D \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \tag{4.105}$$

The corresponding maximum likelihood equations can be found by calculating the partial derivatives of the marginal log-likelihood with respect to each of its parameters. This has been done previously in Section 4.6 for the second term in Equation 4.105 and only the first term requires further study. From Equation 4.104 it follows that

$$\frac{\partial \ell n \sum_{c,z} q_c^{i,z}}{\partial a_{\underline{x}}} = \frac{\sum_{z,D,m} q_c^{i,z}}{\sum_{z,c} q_c^{i,z}} \qquad \text{for each } \underline{x} \tag{4.106}$$

$$\frac{\partial \ell n \sum_{c,z} q_c^{i,z}}{\partial b_{\underline{x}}} = \frac{\sum_{z,D,m} -m q_c^{i,z}}{\sum_{z,c} q_c^{i,z}} \qquad \text{for each } \underline{x} \tag{4.107}$$

$$\frac{\partial \ell n \sum_{c,z} q_c^{i,z}}{\partial \theta_k} = \frac{\sum_{z,c} \frac{Q_{kD}^z}{P_D^z} q_c^{i,z}}{\sum_{z,c} q_c^{i,z}} \qquad \text{for each } \theta_k \tag{4.108}$$

where, as in Section 4.6, Q^z_{kD} is the partial derivative of P^z_D with respect to θ_k. What is important to notice in Equations 4.106 to 4.108 is that the ratio $q_c^{i,z} / \sum_{c,z} q_c^{i,z}$ appears in all of them and can be interpreted as the normalized posterior density of (c_i, z_i). This posterior density will be denoted by $\tilde{n}_c^{i,z}$ and can be thought of as a fractional a-posteriori count assigned to each category (c,z) for the i'th earthquake. When this notation is introduced into Equations 4.106 to 4.108 and summation is performed over all earthquakes, one obtains the equations

$$\sum_i \frac{\partial \ell n \sum_{c,z} q_c^{i,z}}{\partial a_{\underline{x}}} = \sum_i \sum_{z,D,m} \tilde{n}_D^{i,z} = \tilde{n}_{\underline{x}} \tag{4.109}$$

$$\sum_i \frac{\partial \ell n \sum_{c,z} q_c^{i,z}}{\partial b_{\underline{x}}} = \sum_i \sum_{z,D,m} m\, \tilde{n}_D^{i,z} = \tilde{m}_{\underline{x}} \tag{4.110}$$

$$\sum_i \frac{\partial \ell n \sum_{c,z} q_c^{i,z}}{\partial \theta_k} = \sum_i \sum_{\underline{x}} \tilde{n}_D^{i,z} = \tilde{n}_D^z \tag{4.111}$$

where $\tilde{n}_{\underline{x}}$, $\tilde{m}_{\underline{x}}$ and $\tilde{n}_D^c$ are a-posteriori values for the total reported count at location $\underline{x}$, the total reported magnitude at location $\underline{x}$, and the total reported count in detection category (D,z), respectively. Final expressions for the maximum likelihood equations are then:

$$\frac{\partial \ell n \ell^m}{\partial a_{\underline{x}}} = \tilde{n}_{\underline{x}} - \sum_{\underline{x}} T^*_{\underline{x}m} \exp\{a_{\underline{x}} - b_{\underline{x}} m\} = 0 \qquad \text{for each } \underline{x} \tag{4.112}$$

$$\frac{\partial \ell n \ell^m}{\partial b_{\underline{x}}} = -\tilde{m}_{\underline{x}} - \sum_m T^*_{\underline{x}m}\, m \exp\{a_{\underline{x}} - b_{\underline{x}} m\} = 0 \qquad \text{for each } \underline{x} \tag{4.113}$$

$$\frac{\partial \ell n \ell^m}{\partial \theta_{\underline{x}}} = \sum_{D_k} \left[\sum_{z_k} \tilde{n}_D^z \frac{Q_{k,D}^z}{P_D^z} - n_D^* Q_{k,D} \exp\{a_{\underline{x}} - b_{\underline{x}} m\} \right] = 0$$

for each θ_k (4.114)

Apart from the fact that the counts $\tilde{n}_{\underline{x}}$, $\tilde{m}_{\underline{x}}$, and $\tilde{n}_D^z$ are functions of the unknown parameters $a_{\underline{x}}$, $b_{\underline{x}}$ and θ, the above expressions are identical to the maximum likelihood equations derived for the case when no errors on the data are considered. This feature suggests a simple iteration scheme to obtain maximum likelihood estimates of the parameters: A reasonable initial solution for the a-posteriori counts is found by using a-priori information only, e.g. $q_c^{i,z(0)} \equiv p_c^{i,z}$ in Equations 4.106 to 4.108. For given counts $\tilde{n}_{\underline{x}}$, $\tilde{m}_{\underline{x}}$ and $\tilde{n}_D^z$, the parameters $a_{\underline{x}}$, $b_{\underline{x}}$ and θ_k can be estimated using techniques given earlier and including a-priori information as described in Section 4.7. A-posteriori counts can then be updated using Equation 4.104 and iteration should proceed until convergence. It is clear intuitively that the likelihood in each of these steps must monotonically increase, since the counts are redistributed in accordance with the seismicity which is estimated. Hence, a solution is always guaranteed. However, it is not evident whether only a single maximum of the likelihood exists and whether the likelihood is stationary at the maximum point. For instance, if location $\underline{x}$ has a-priori large recurrence rates relative to the other locations, then the a-posteriori recurrence rate at $\underline{x}$ will be even larger and the spatial distribution of recurrence rates more variable. Because the likelihood function may have more than one local maximum, the solution may depend on the initial values used in the algorithm. In application of the method to Model C, it is shown how this effect can be counteracted by imposing smoothness on the spatial variation of recurrence rates. In this case, the solution is expected to be more stable.

4.8.3 Modification to the Maximum Likelihood Estimation for Earthquakes Falling Outside the Range of Analysis

So far, it has been assumed that the prior distribution of the variables $\underline{x}$ and m, when subject to error, is entirely within the domain of interest. However, in the analysis of the earthquake data, interest typically focuses on a given magnitude range $[m_0, m_1]$ and only earthquakes with $m_0 \leq m \leq m_1$ are analyzed. The problem then arises of dealing with earthquakes for which the prior distribution falls in part outside this range of analysis. The same problem evidently occurs for earthquakes with uncertain location and near the boundary of the region of interest. Whereas in the latter case, the easiest solution is to extend the domain of interest, this is not very practical for the size measure m, since the assumptions of the model (such as exponentiality of the recurrence law and spatial homogeneity of incompleteness) may hold only over a limited range of size measures. The following approximate solution is therefore used:

1. The recurrence rate at magnitudes lower than m_0 is assumed to be

$$\lambda^z_{D\underline{x}m} = \lambda^z_{D\underline{x}m_0} \exp [\, b_{\underline{x}} (m_0 - m)] \qquad \text{for } m < m_0 \qquad (4.115)$$

where $\lambda^z_{D\underline{x}m}$ is the rate of reported earthquakes for detection category (z,D), location $\underline{x}$ and size m

2. The recurrence at magnitude larger than m_1 is assumed equal to zero

$$\lambda^z_{D\underline{x}m} = 0 \qquad \text{for } m > m_1 \qquad (4.116)$$

In the analysis, all the data that possibly fall inside the range $[m_0,m_1]$ are considered, but only the fraction of $\tilde{n}_i$ inside this range is used. The approximation lies in the fact that $\lambda^z_{D\underline{xm}}$ for $m<m_0$ or $m>m_1$ is not estimated from the data and therefore does not enter into the likelihood formulation. Equivalently, one could say that $\lambda^z_{D\underline{xm}}$ does enter into the likelihood formulation, but is associated with unknown periods of observation and satisfies Eqs. 4.115 and Eqs. 4.116. In this last interpretation, the modified solution is an exact one, in sofar as Eqs. 4.115 and 4.116 are satisfied.

A final remark is necessary on the treatment of earthquakes originally reported in a magnitude scale other than m. Suppose for instance that the analysis is in terms of Modified Mercalli (MM) intensity I_0. In this case, earthquakes reported in the MM scale and with uncertain I_0 should be redistributed according to the recurrence rate of <u>reported</u> events in MM. Suppose on the other hand that an earthquake is reported in an alternative scale, such as bodywave magnitude m_b. Then one should distinguish between two types of earthquake size uncertainty: 1. uncertainty on the reported value of m_b and 2. uncertainty on the estimated value of I_0. To account for uncertainty on m_b, the redistribution should incorporate the probability of detection of earthquakes with m_b reported. However, uncertainty on the estimated value of I_0 must be treated differently. In this case, it is known that I_0 is <u>not</u> reported and, in principle, the recurrence rate varies with $1-P_D(I_0)$. A simple example is useful to clarify the procedure: Suppose we know that, for events with $I_0>IV$ that have occurred after 1950, I_0 is reported in the catalog with probability one. Then one must accept the consequence

that, if only m_b is reported, the corresponding value of I_0 must be lower than IV, irrespective of the value of m_b. As will be seen in the applications, there are in fact several such events in the Chiburis catalog. However, it does not appear plausible that these events are not detected by human observers. Rather, it appears that I_0 was not reported in the catalog because instrumental magnitude is a more accurate size measure. For this reason, it is assumed that if only an instrumental size measure is reported, the distribution of the unknown I_0 value is simply exponential, and is not corrected for mode of detection.

4.9 GOODNESS-OF-FIT AND UNCERTAINTY OF THE ESTIMATORS

In the previous sections, attention has been focused on the formulation of statistical models and on the estimation of their parameters. The structure of those models is based in part on intuitive reasoning, in part on exploratory analysis of the data. Estimation of the parameters has been through maximum penalized likelihood. The present section discusses two additional issues which are important to the analysis: 1. evaluation of the goodness-of-fit of each model, 2. calculation of uncertainty on the estimated parameters. Examination of the goodness-of-fit of the models is of importance to validate the assumptions underlying the models, to detect possible deficiencies and to compare their relative performance.

Uncertainty on the estimated parameters is of concern in the prediction of future recurrence rates, which is for example necessary for seismic hazard analysis. Both problems are found to be extremely complex and this section is suggesting possible approaches, rather than giving

definite answers. Complexity is mainly the result of two characteristic of the data and the models:

1. The data is sparse and prohibits the use of asymptotic properties of usual goodness-of-fit statistics or asymptotic expressions for maximum likelihood estimators.
2. The estimated parameters can be strongly dependent due to the use of smoothing, constraints and other a-priori conditions (see Section 4.7). Consequently, the number of degrees of freedom, which are necessary to judge the usual goodness-of-fit statistics, are not well defined and the likelihood function, which is the key to calculating uncertainty of the estimators, has a complicated form.

Approximate procedures that bypass these problems are discussed next.

4.9.1 Goodness-of-Fit of the Models

In their most general form, the statistical models proposed in this thesis classify the earthquake data according to geographical location $\underline{x}$, size m, time of occurrence t, population density p, distance to the nearest instrument d and mode of detection z, hence into categories $(\underline{x},m,t,p,d,z)$.

In principle, an evaluation of goodness-of-fit must consider the expected and observed counts in each of these categories. Distinction should be made here between a *global* test and a *local* test of the model. In a global test, a summary statistic such as χ^2 is compared wth its theoretical distribution and, if found significantly large, the model is rejected. Apart from the fact that in the present case the distribution of the χ^2 statistic is not known (its distribution and the distribution of

any other summary statistic could of course be found by simulation), global testing does not reveal the nature of lack-of-fit, should lack-of-fit be found. A more fruitful approach is then to study the pattern of local violations of the model. In that perspective, various marginal classifications of the earthquake counts are of interest: The most important assumption made in all the models is perhaps that nonstationarity of the observed recurrence rates is due to incompleteness. To check this assumption, it is logical to compare expected and observed counts at each location $\underline{x}$ in different time periods. Another important assumption is that of exponentiality of the recurrence rate as a function of size m. The validity of this assumption can be assessed by comparing expected and observed counts as a function of m for different regions $\Omega(\underline{x})$. Similarly the appropriateness of the assumed model for the probability of detection can be checked using classifications of the data in detection categories (t,p,z) and (t,d,z).

A simple Poisson test is useful for this purpose. Given that the expected count in a certain category i equals $\bar{n}_i$, the probability of the count being less or equal than the observed count n_i is easily calculated. For instance,

$$P[N_i \leq n_i] = \sum_{k=0}^{n_i} \frac{(\bar{n}_i)^{n_i} e^{-\bar{n}_i}}{n_i!} = \alpha_i \qquad (4.117)$$

Very low and very high values of α_i indicate that the expected count is too high or too low respectively. It should be emphasized that no strict interpretation must be given to α_i, because the expected count used in the

test is data dependent. However, the true α_i is more "extreme" than the calculated α_i. These "significance levels" are used here only to compare model predictions with observations in an intelligible way, by flagging categories i associated with very low and very high values of of α_i. The fraction of cells that are flagged and the pattern of flagging is then of interest. Examples will be shown in the application of the models (see for instance Fig. 4.25a).

Typically many cells have very low or zero counts and the test of Equation 4.117 may flag as significant the occurrence of just one or only very few earthquakes. Various ways have been suggested to deal with problems of this type in the context of contingency tables, e.g. by Fienberg and Holland (1980). One that is found useful in the analysis of the earthquake counts consists of adding a small quantity δ to both n_i and n_i prior to the test. Compare for instance Fig. 4.25a with Fig. 4.25b where δ has been set equal to 1.

Traditionally, examination of the validity of the exponential recurrence relation or of the completeness model has been done directly on the basis of empirical plots. Since total earthquake counts for each size measure are large, one could also use the approximate assumption that n_i has Gaussian distribution $N(\bar{n}_i, \bar{n}_i)$ and examine the standardized residuals

$$\Delta_i = \frac{n_i - \bar{n}_i}{\sqrt{\bar{n}_i}} \qquad (4.118)$$

Again, one should be careful in the interpretation of the associated significance level, since $\bar{n}_i$ depends on n_i.

In application of the models, it is often found that deviations from the exponential recurrence relation are significant and indicate a faster than exponential decrease. Several possibilities could be considered. One is that the earthquakes of low size are incomplete even today. Such an assumption is however contrary to general belief based on the detection ability of the seismic network. Another possibility is that magnitude has non-exponential distribution for relatively low values; in this case, one might exclude from the analysis earthquakes in the lower mangitude range. As will be shown in the application sections, this may lead to unrealistic results because of the sparseness of the remaining data. An alternative and perhaps better technique is to allow for larger deviations from the assumed recurrence relation for small values of m. This can be done by using a weighted likelihood formulation, such that the contribution to the likelihood of events with small size is less than that of large-size events. Since the various terms in the log-likelihood without considering errors in the data (Eq. 4.45) are proportional either to the observed count or to the period of obervation, a simple way to do so is to replace these values with weighted ones depending on the size m. For instance

$$T_c^* = T_c \quad w_m \tag{4.119}$$

$$n_c^{z*} = n_c^z \quad w_m \tag{4.120}$$

Thus, if w_m is zero, earthquakes of size m are not considered in the analysis. The same technique is also used in the case when the size measure is uncertain, by applying weights to the a-posteriori counts, i.e.

$$\tilde{n}_c^{z*} = \tilde{n}_c^{z} \, w_m \qquad (4.121)$$

In this case, a direct interpretation in terms of the total likelihood is less evident. Notice however that, if w_m is set to zero, Eq. 4.121 is compatible with the treatment of earthquakes for which the prior distribution partially falls outside the analyzed magnitude interval $[m_0, m_1]$: although the recurrence rate of earthquakes with size measure below m_0 is assumed to follow the exponential relation when calculating a-posteriori counts, the a-posteriori counts below m_0 are not used in the analysis. The probability of detection for size m, which also enters into the redistribution, is determined by the smoothness imposed on P_D if w_m is zero.

4.9.2 Uncertainty on Recurrence Rates

It is convenient to separate uncertainty on the seismicity parameters due to two different sources:

1. Model uncertainty, by which we mean uncertainty on the appropriate treatment of a given data set. This includes uncertainty on the input parameters and on the analysis options used in the various models (i.e. the degree of smoothing imposed on the estimates, the choice of the model, constraints, etc.). A convenient way to characterize model uncertainty is to specify a discrete set of alternative input conditions or models and to assign a probability to each alternative. These probabilities are then applied to the resulting parameter estimates and seismic hazard curves.

2. Statistical uncertainty on the parameter vectors $a_{\underline{x}}$ and $b_{\underline{x}}$, given

the input conditions. In Bayesian analysis, this uncertainty is quantified by the posterior distribution of $a_{\underline{x}}$ and $b_{\underline{x}}$, given the input conditions. The posterior density of these parameters is proportional to the likelihood function, penalized and weighted in various ways and possibly modified by prior distributions, e.g. on $b_{\underline{x}}$.

A major obstacle to the calculation of the joint distribution of $a_{\underline{x}}$ and $b_{\underline{x}}$ is the high dimensionality of these vectors. Convenient procedures for the numerical characterization of parameter uncertainty in complex inferential problems are based on 1. simulating a large number n of data sets to represent the variability of the statistical sample, 2. analyzing each simulated set j to produce estimates $(\hat{a}^j_{\underline{x}}, \hat{b}^j_{\underline{x}})$, j=1,...,n, and 3. estimating properties of the joint distribution of $a_{\underline{x}}$ and $b_{\underline{x}}$ by considering $(\hat{a}^1_{\underline{x}}, \hat{b}^1_{\underline{x}}),\ldots,(\hat{a}^n_{\underline{x}}, \hat{b}^n_{\underline{x}})$ as a random sample from that distribution. For example, the variance of $a_{\underline{xy}} = a_{\underline{x}} a_{\underline{y}}$ may be estimated as

$$s^2_{a_{\underline{xy}}} = \frac{1}{n-1} \sum_{j=1}^{n} \left[a^j_{\underline{xy}} - \bar{a}_{\underline{xy}}\right]^2 \qquad (4.122)$$

where $\bar{a}_{\underline{xy}}$ is the sample mean of $\hat{a}_{\underline{x}} \hat{a}_{\underline{y}}$. Similarly for other variances and covariances. For the purpose of seismic hazard analysis, calculation of distribution characteristics of $a_{\underline{x}}$ and $b_{\underline{x}}$ is not necessary: one may simply calculate the hazard curve at the site that corresponds to the parameters $(\hat{a}^j_{\underline{x}}, \hat{b}^j_{\underline{x}})$ for each j and then treat the set of n hazard curves as a statistical sample.

Methods for the generation of artificial data sets are broadly referred to as resampling techniques. The best known such methods are

bootstrapping and jackknifing (Efron, 1979, 1982), each with several variants. One possibility in our case is to generate artificial samples assuming that the true earthquake process is Poisson with parameters $(\hat{a}_{\underline{x}}, \hat{b}_{\underline{x}}, \hat{P}_D)$ obtained from the historical data. A limitation of this procedure is that, if the method of estimating $a_{\underline{x}}$, $b_{\underline{x}}$ and P_D is biased, then sampling is from a biased model and the results may not be representative of actual uncertainty. For example, in the case of earthquake rates one should be careful not to sample from very "erratic solutions", in which the "spikes" may be caused by the tendency of ML to concentrate seismicity in a few cells when uncertainty on location is considered (model C). On the other hand, one should not sample from an excessively smooth solution, or else smoothing again the counts in the process of estimating $(\hat{a}^j_{\underline{x}}, \hat{b}^j_{\underline{x}})$ will produce flat and nearly identical solutions.

More work is required to address the issue of estimating uncertainty on the parameter estimates. Although the generation of artificial data sets is relatively straightforward, this is a computationally demanding task and one gains little insight into the influence of different modelling options on the uncertainty. In addition, it is not clear how uncertainty on the earthquake attributes $(t,\underline{x},m)$ can be accounted for in such a method. A method which avoids the latter problem is to generate artificial samples directly from the data (selecting each of the earthquakes with equal probability and with replacement until a sample of the required size is obtained). Such a method (empirical bootstrapping) has the disadvantage that empty categories always remain empty and would probably favor less smoothed estimates. A comparison of the empirical and parametric bootstrapping methods will be shown in Sect. 4.13 for Model D.

4.10 APPLICATION OF MODEL A

4.10.1 Introduction

In this and the following three sections, various examples of application of the models A to D to actual data will be shown. The purpose of these applications is primarily to illustrate the methods, to check their validity and to show the sensitivity of the results to the input parameters. Therefore, the estimated recurrence rates should not be used directly for seismic hazard analysis. Such an analysis would certainly require additional expert opinion about reasonable spatial configurations of seismicity, the composition of the catalog and the quality of the seismic network. In addition, input parameters have been selected to demonstrate the effect of certain assumptions, even if their actual values are sometimes debatable.

Except for Model C the discussion of each application is separated into six subsections: First, a brief review of the assumptions used in each model is given, with reference to the earlier theoretical sections. Next, the earthquake data and the discretization of the explanatory variables is briefly discussed. The third subsection describes the prior information used in the analysis and the fourth subsection summarizes the sensitivity cases that are considered. In the fifth part, the results of these analyses are discussed, followed by conclusions about the merits and deficiencies of the model. Because Model C partially overlaps with Model D, a more concise and qualitative discussion of the results is presented for this model in Section 12.

4.10.2 Review of Assumptions and Methods

Model A developed from considerations of incompleteness of the

catalog, while less attention was given to the spatial modelling of the seismicity rates. Following assumptions are made in this model:

- Earthquake occurrences follow a Poisson process
- True seismicity is stationary in time, is spatially homogeneous over specified regions Ω_k and follows an exponential relation as a function of size m (Eq. 4.2).
- The probability of detection can be separated into three independent effects (Eq. 4.26):
  * the transmittal loss of reports β_{tm} depends on the time of occurrence of the earthquake t and its size m
  * the detection of earthquakes by human observers α_{pm} depends on the population density in a region around the epicenter and the size of the earthquake
  * the detection of earthquakes by seismic instruments γ_{dm} depends on the distance to the nearest instrument and the size of the earthquake
- The slope parameters b_k in Eq. 4.3 satisfy one of the following three conditions:
  * The slopes b_k are independent
  * The slopes b_k are identical
  * The slopes b_k are i.i.d. random variables with normal distribution $N(m_B, \sigma_B^2)$ and unknown mean value m_B and variance σ_B^2

The corresponding maximum likelihood equations and the methods used to solve them have been discussed in Sections 4.6 and 4.7.

4.10.3 Earthquake Data and Discretization of Explanatory Variables

Earthquake data are obtained from the Chiburis catalog within the region of study indicated in Fig. 4.1b. Since most earthquake sizes are reported as epicentral intensities I_0 in a Modified Mercalli (MM) scale, I_0 is used as the common size measure (in this and in the following sections, the symbol m is however maintained when referring to I_0 as an explanatory variable). Uncertainty on the size of the historical events is not considered in this model. When two different values of I_0 are reported, the smaller one is used. This corresponds to an intuitive correction for uncertainty, since smaller values of I_0 are more likely to occur. When I_0 is not reported, the instrumental size measure is converted to I_0 using the relationship proposed by Chiburis (1981),

$$I_0 = (M - 1)/0.6 \tag{4.123}$$

Only integer values of I_0 are considered and in Eq. 4.123, I_0 is rounded off to the nearest integer. As a rough correction to the problem of clustering, earthquakes indicated in the catalog as aftershocks are removed, since at the time of application the identification of clusters as discussed in Chapter 3 had not been developed yet.

To model the spatial variation of seismicity, the seismogenic provinces shown in Fig. 4.1b are used. These sources are one of many alternative configurations proposed for New England (WGC, 1983). The temporal variation of seismicity rates within each province has been illustrated in Figs. 4.2. Incompleteness for small I_0 and early periods of the catalog is evident.

To model the probability of detection P_D, the population density near the epicenter p, the distance to the nearest instrument d and the time of

occurrence t need to be discretized. Two cases are considered to summarize the spatial configuration of population density. In the first case (Case 1), p corresponds to the population density at the epicenter, as discretized in Figs. 4.4. In the second case (Case 2), p is taken to be the maximum category found in a square region around the epicenter, the size of which depends on I_0. In this case, p accounts, at least to some degree, for the fact that more severe earthquakes can be detected by people at larger distances from the epicenter. The extent of the epicentral region, in units of quarter-degree cells, is given in Table 4.5 as a function of I_0. The epicentral region is also larger than in Case 1 for small I_0, to account for possible inaccuracy of the population maps or the reported epicentral coordinates. The net effect of using the maximum population category over an extended epicentral region is a shift towards higher values of p and a smoothing of the original population maps. Note that, because the degree of smoothing depends on the size measure m (I_0), the periods of observation T_C in Eq. 4.33 also depend on m. For instance, Fig. 4.11 compares the fraction of the area occupied in each province for different p for Case 1 and for the maximum smoothing level used in Case 2. Those fractions vary in time and the results shown are time averages. Fig. 4.11b indicates that category p=0 practically disappears for all provinces. The effect of smoothing is largest for Province 5, due to the fact that a substantial part of this province extends over the Atlantic Ocean. Figs. 4.12 show the variation in time of each population category for the different smoothing levels and should be compared with Fig. 4.5. Here, the fractional area is an average over all provinces. Notice that, for the maximum smoothing level (Fig. 4.12d), the

entire area of study has been settled with population categories 4 and 5 since 1860 and that category 4 has disappeared since 1950.

The distance to the nearest seismic instrument is discretized as shown in Table 4.2. A representative set of the spatial distribution of d for different time periods is shown in Figs. 4.7. Discrete time intervals t are defined as in Table 4.6. Basically, these time periods separate the different modes of reporting as discussed in Section 4.4, and include also some additional intervals to better model the temporal variation of P_D.

4.10.4 Prior Information

Apart from the fact that different options are allowed to relate the slope parameters b_k in different provinces, prior information is needed to constrain the estimates of b_{tm}, α_{pm} and γ_{dm}. The following constraints have been used:

- $\beta_{tm} = 1$ for all m and t = 5 (since 1950)
- $\beta_{tm} = \alpha_{pm} = \gamma_{dm} = 1$ for all t,p and d and m = 7 (I_0 = VIII)

(4.124)

These constraints have been discussed earlier in Section 4.2.1 and appear to be reasonable ones.

Smoothness of the estimates is imposed by including a term in the log-likelihood, which penalizes deviations from a locally linear variation of the parameters α, β, and γ with their subscript indices. After a number of preliminary runs, the penalty coefficients were choosen as follows:

- $P_{\alpha}^{p} = P_{\alpha}^{m} = 20$
- $P_{\beta}^{t} = P_{\beta}^{m} = 200$ (4.125)
- $P_{\gamma}^{d} = P_{\gamma}^{m} = 20$

4.10.5 Analysis Cases

Estimates using Model A were obtained using variations on the following:

1. the definition of the population density around the epicenter
2. whether or not, for small size measures, the exponential recurrence relation is satisfied
3. the condition of similarity among the parameters b_k for different provinces
4. the prior information on β_{tm}, α_{pm}, γ_{dm}

For the purpose of discussing these results, the following labelling is useful:

- Case 1 refers to the use of epicentral population density and considers all earthquakes (I_0=I to VIII, or m=0 to 7)
- Case 2 refers to the use of I_0-dependent population density and also uses all earthquakes
- Case 3 is identical to Case 2, except that earthquakes with I_0 equal to I are excluded from the analysis (i.e. 7 size categories are considered, m=0,6)

For each of the above cases, the three alternative assumptions on the similarity of b_k values are used. Sensitivity to the prior information on the completeness parameters has been considered only in Case 1.

4.10.6 Discussion of Results

Probability of Detection

Estimates of the incompleteness of parameters β_{tm}, γ_{dm} and α_{pm} are plotted in Figs. 4.13a-c, for Case 1 and using independent values of b_k. Fig. 4.13a shows that time has a direct influence on incompleteness, which is separate from that induced through variation of population and instruments. This independent effect of time can be attributed to more likely loss of records from earlier periods as well as to the evolution in time of instrument sensitivity, people awareness, and mode of recording. Fig. 4.13b contains similar plots for the probability of detection by instruments. The values of γ_{dm} in that figure might at first seem too small, especially for I_0 in excess of IV or V. However, these estimates are consistent with the number of historical earthquakes reported by people, but not detected by instruments (i.e. without an assigned magnitude); see Table 4.7. All the earthquakes of intensity VI and VII that do not have an assigned magnitude (see Table 4.8) occurred prior to 1955, indicating that instrument characteristics and network management may have improved significantly over the last 25 years. If this is the case, then the current probability of detection by instruments would be higher and the probability of detection in the first few decades of the century would be lower than displayed in Fig 4.13b. The values in the figures may in any case be interpreted as time-average probabilities. The reason why time effects may be only partially removed from the probability of instrument detection is that the effect of time may be different for people and instruments, contrary to the assumption of this model. One should also use caution in extrapolating the results of Fig. 4.13b beyond

the geographical limits considered in this study, because of likely regional variations of instrument types and network management.

Estimates of the probability of detection and recording by people are plotted in Fig. 4.13c. These probabilities should be geographically more stable than those for instruments. The sharp increase of people sensitivity between intensities III and IV is in good correspondence with the definition of these intensities in the Modified Mercalli scale. Fig. 4.13d shows the estimates of α_{pm} for case 2, where p is a function of I_0. An immediate consequence of the I_0 dependent smoothing of population is that α_{pm} is less dependent on I_0 for intermediate intensities. In this case, the constraint that α equals one for all p for I_0 = VIII may have been inappropriate: for I_0 = VIII, the epicentral region in Table 4.3 is so large that small values of p are very unlikely (see Figure 4.13d), and hence the constraint would have been unnecessary. Removing this constraint would probably lead to estimates of α that are even less variable as a function of size measure m.

The effect of the definition of p on the global probability of detection is more easily judged on the basis of the equivalent period of completeness $T^*_{\underline{x}m}$ (Eq. 4.46), integrated over the area of each province k, i.e. of T^*_{km}. Remember that $T^*_{\underline{x}m}$ refers to the total timespan of the catalog (from 1625 to 1980) appropriately scaled at each location $\underline{x}$ by the average of the probability of detection over time. The latter average value may be thought of as an incompleteness factor and is shown in Fig. 4.14 for each province k and earthquake size m. Fig. 4.14 compares estimates of T^*_{km} for Cases 1 and 2, and for Case 1 when the dependence of the recurrence rate on m is not assumed to be exponential (in this case, estimates of the recurrence rate for size m correspond to the earthquake

count reported for size m divided by the associated equivalent period of completeness) The estimates for Cases 1 and 2 are quite similar. The largest difference occurs for Province 5, due to the fact that this province is composed of a densely populated region on land and a region of the Atlantic Ocean with no population. Under the assumption of Case 1, the probability of detecting earthquakes in the ocean is based on the (zero) population at the epicenter, whereas in Case 2, the proximity of settlements along the shore is taken into account, particularly for earthquakes for high intensity.

A more detailed picture of the spatial variation of the probability of detection is shown in Fig. 4.15 where maps of $T^*_{\underline{xm}}$ are given for each I_0 in Case 2 using independent b_k. The figure actually gives values of $10 \times T^*_{xm}/T_{max,m}$ where $T_{max,m}$ is the maximum period of completeness with the following values (in years)

I_0	I	II	III	IV	V	VI	VII	VIII
$T^*_{max,m}$	2.2	12.0	37.8	90.3	129.0	160.9	237.8	356.0

(4.126)

For instance, a value of 4 in the figure for I_0 = V means that

$$4\ (12.9) < T^*_{\underline{xm}} < 5\ (12.9) \tag{4.127}$$

Because P_D is constrained to one for I_0 = VIII, the corresponding map is uniform in space and the equivalent period of completeness corresponds to the timespan of the catalog.

The effect on the detection probabilities of the similarity condition assumed for b_k is small. Estimates of the completeness parameters and equivalent periods of completeness also remain nearly the same when, in Case 2, earthquakes with I_0 = I are neglected (Case 3). As

shown in Fig. 4.14, the extreme case of non-parametric variation with m of the recurrence rates produces more substantial differences. The main effect is to increase P_D and hence to decrease the recurrence rates for small earthquake sizes. This indicates lower than exponential rates for small m.

The preceding results are based on the constraints and prior information described under Section 4.10.4. Increasing the value of the penalty coefficients in Eq. 4.125 reduces the curvature of the functions α, β and γ with respect to their subscripts. Differences in the estimated recurrence rates are however modest. Lowering the value of the penalty coefficients tends to produce rather erratic estimates. This is no surprise since the number of parameters is large and the earthquake counts in the various categories is often small.

Recurrence Parameters

Estimates of the recurrence parameters a_k and b_k are shown in Table 4.9 for Cases 1, 2 and 3 and for different assumptions on the similarity of the b_k parameters. In Case 1, the assumption that all b_k are identical appears not to be realistic, given the large differences in individual estimates when b_k are treated as independent quantities and the fact that the same estimates are not found to be identical under the assumption that the slopes b_k are i.i.d. random variables. A better assumption in this case is that the slopes b_k are identical for two groups of provinces, (1,2,3,4,5) and (6,7). For Case 1, the fitted exponential relations are plotted in Figs. 4.16: crosses indicate the historical earthquake counts in each province and boxes indicate non-parametric corrections for incompleteness. In practically all cases, the earthquake

counts for low m are systematically overpredicted by the exponential fits.

Recurrence parameter estimates for Case 2 are very similar to those for Case 1. It is interesting that, for Case 3 where $I_0 = I$ is not considered, the assumption that b_k are i.i.d. random variables leads to identical estimates. Since the assumption of exponentiality appears to be a crucial one, the case when all b_k are identical has been analyzed for three size measure ranges: I_0 = I-VIII, I_0 = III-VIII and I_0 = V-VIII, using the I_0-dependent definition of p. The results are shown in Fig. 4.17 by pooling all provinces together. The progressively higher value of b appears to indicate faster-than-exponential decay of the rate for increasing I_0.

4.10.7 Conclusions

Contrary to more traditional techniques where only data within periods of the catalog judged to be complete are used, the present model uses all the data. To do so, a physical process leading to incompleteness is proposed, which is explicitly related to the spatial and temporal variation of population density and seismic instrument location. This technique allows one to estimate a more refined spatial description of incompleteness, is more objective and should lead to more reliable recurrence parameter estimates.

Careful consideration should be given to the information used in the present model to estimate the probability of detection. This information includes:

- The constraints and smoothing of the completeness parametes, which are necessary to stabilize the solution. Individual estimates of

the parameters may change substantially if this prior information is changed (although the estimated recurrence parameters are rather stable). The variation of the completeness parameters with earthquake size m (with intensity I_0) is suspected to have large statistical uncertainty, since it interacts with the estimate of the slope parameters of the recurrence relation.

- The assumption that the variation in time of P_D is the same for detection by both people and seismic instruments is not a very reasonable one and produces rather low estimates of the detection probability by seismic instruments. These estimates should be interpreted as time averages rather than time-specific values.
- The completeness parameters depend on the configuration of seismogenic provinces.
- Exponentiality of the recurrence rate as a function of m is debatable for low sizes and tends to produce too low estimates of the probability of detection for these earthquake sizes.

4.11 APPLICATION OF MODEL B

4.11.1 Review of Assumptions and Methods

Model B developed from various limitations noticed on model A. The main changes with respect to Model A are as follows,

1. In order to arrive at estimates of the probability of detection that do not depend on the seismogenic provinces, model B uses a spatial grid to represent the spatial variation of the recurrence rates.
2. Because the variation with m of the completeness and recurrence parameters are to some extent interchangeable, model B assumes

that the completeness parameters are independent of m. To do so, the population density and the distance to the nearest seismic instrument are redefined to indirectly capture the influence of m on the probability of detection.

3. The probability of detection by people and instruments is allowed to vary independently in time. For seismic instruments, it is further assumed that the effect of time does not depend on ground motion intensity at the site of the nearest instrument.

Maximum likelihood equations for the estimation of the probability of detection are discussed in Section 4.6.5 (Eqs. 4.55). Estimation of the recurrence parameters $a_{\underline{x}}$ and $b_{\underline{x}}$ at location $\underline{x}$ is done using maximum penalized likelihood, as explained in Section 4.7.3.4 *Tk. 4.68).

4.11.2 Earthquake Data and Discretization of Explanatory Variables

The Chiburis catalog is used for application to the region of Fig. 4.1a. Again, I_0 is used as a uniform size measure and magnitude is converted to I_0 through Eq. 4.123). Contrary to Model A, a correction for clustering is made by elimination of the dependent events identified in the base case analysis discussed in Chapter 3. Consequently, earthquake counts are somewhat lower than in Case A. Also, earthquakes with I_0 less than II and events prior to 1625 are excluded. All events in the region of study have I_0 less or equal than VIII.

To account for the influence of earthquake size on the probability of detection, population category p is redefined by using a weighted average of the population density around the epicenter, where the weights depend on the site intensity at each location (Eq. 4.29). Nominal values q of the population density in Eq. 4.19 are shown in Table 4.10. The

attenuation function used for the calculation of p is the modified Gupta-Nuttli regression for the Central United States,

$$I = 3.2 + I_0 - 1.17 \ln(R) - 0.0011R \qquad (4.128)$$

where R is epicentral distance in kilometers. Values of r equal to 3, 5 and ∞ are used in the present analysis. For finite r, p corresponds to reduction of the spatial distribution of population density and site intensity to a single scalar. Fig. 4.18 illustrates contributions to p from combination of different site intensities I and population densities q. In the figure, p is arbitrarily scaled to one for the highest value of I and q. Note that the variation of p with I is larger for r=5 than for r=3. Eight discrete population categories p are defined on the basis of the logarithm of the continuous variable as shown in Table 4.11. A logarithmic transformation is used to increase the resolution at low population density and for small values of m. In the limiting case when $r=\infty$, p corresponds to I_0 and seven size categories are used, which correspond to unit intensities on the Modified Mercalli scale.

Time categories to model the variation of detection by people as a function of time are the same as in Model A. (Table 4.6) The distance to the nearest seismic instrument used in Model A is also redefined in the present analysis to account for the size of the earthquake. This is done on the basis of the distance to the nearest instrument and the epicentral intensity as shown in Table 4.12. The new classification corresponds approximately to site intensities -1,0,1,2 and $I_0 \geq 3$ at the location of the instruments (intensities -1 and 0 refer here to an extrapolation of the Modified Mercalli scale, using the attenuation function in Eq. 4.128). The time categories for the variation in time of the detection capability

of seismic instruments has been refined with respect to model A and are shown in Table 4.13. Spatial variation of the recurrence rates is modelled by dividing the geographical region into 56 unit-degree cells. Thus, there are 56 pairs of parameters $a_{\underline{x}}$ and $b_{\underline{x}}$.

For the interpretation of the results in this section it is useful to reconsider the earthquake data. Fig. 4.19 shows plots of the empirical recurrence rate over the entire region of study as a function of time and for different I_0. The difference with Fig. 4.2a is that the entire region is used and clustering of the earthquakes has been treated differently. Fig. 4.20 shows the spatial distribution of earthquakes for each time category. Variation with time of the spatial distribution of earthquakes is evident, especially in the early periods. Later in this section, it will be shown how well the model explains this variation through incompleteness. Table 4.14 shows the earthquake counts for each population-time and instrument-time category, depending on the mode of detection for r=5. The interpretation of these counts is not easy, because counts are associated with different observational areas and are therefore not directly comparable. However, direct comparison of time-totals and of counts in different detection modes is possible. For example, the column-sums and the empirical rates in the last row of Table 4.14a show two important facts: 1. the rate of earthquake detection by people has steadily increased until about 1950 but has since declined, presumably due to a shift of attention towards instrumental determinations of earthquakes size. Analoguous statistics for instruments in Tables 4.14c and 4.14d indicate that, except for events in the first time category, the rate of reporting has increased and the rate of non-reporting has decreased in recent times. The effect of increasing the

number of seismic instruments during the last decade (see Fig. 4.6) is especially evident.

Fig. 4.21 summarizes the spatial distribution of earthquake counts. The cumulative count and average total magnitude in each cell are sufficient statistics for the estimation of $a_{\underline{x}}$ and $b_{\underline{x}}$ if all I_0 larger or equal than 2 are used. Because in several of the analyses that follow the smaller size measures are excluded or down weighted in the maximum likelihood, Fig. 4.21c presents a breakdown of those counts according to each I_0. A singular case that will be discussed later is that of the cell with coordinates (70.5° W, 39.5° N), for which the only event reported in the catalog has intensity 7. By contrast, the only 3 events that are known to have occurred in the cell immediately to the west of this location have all intensity 3.

4.11.3 Prior Information

To obtain reasonable estimates of the completeness parameters α_{pt}, γ_d and β_{t*} in Eq. 4.28, it is found necessary to constrain the probability of detection by people α_{pt}. In this analysis it is assumed that, for the largest value of p (p=8 when r=3 or 5, p=7 when r=∞), earthquakes are reported at all times. Earthquakes associated with the next lower value of p (7 and 6 respectively) are assumed complete since 1910. Thus, for r=3 or 5,

$$\alpha_{tp} = 1 \text{ for } (p=8, \text{ all } t) \text{ and for } (p=7,\ t=4,5) \qquad (4.129)$$

Smoothness of the α estimates is again imposed by penalizing the likelihood (Section 4.7.2.2). In this case, interpolated values are calculated after transforming α to a logit-scale (Eq. 4.60). Reasonable estimates of α have been found using

$$P_\alpha = 100 \tag{4.130}$$

Smoothness of the spatial variation of the recurrence parameters is only imposed on the slope parameters $b_{\underline{x}}$. A global average of $b_{\underline{x}}$ over the entire region is used as an interpolator, in Eq. 4.80, i.e.

$$\hat{b} = \frac{1}{56} \sum_{\underline{x}} b_{\underline{x}} \tag{4.131}$$

where 56 is the number of one-degree cells used in the analysis. The corresponding penalty coefficient is set to 10,

$$P_b = 10 \tag{4.132}$$

Sensitivity of the results to these assumptions will be illustrated in the following applications.

4.11.4 Analysis Cases

Based on several preliminary runs of the model, it was decided that the assumption of exponential recurrence relation does not hold for the lower size measures and may bias the incompleteness results and recurrence rates. To correct for this problem, a weighted likelihood solution has been used as a base case and variations of the parameters are relative to this case. The value of r for the definition of population density in the base case is choosen equal to 5. The input parameters for all cases presented here are summarized in Table 4.15 and correspond to the following sensitivity analysis:

- Case 1 uses a value of r equal to 3, thus reducing the influence of epicentral intensity on the definition of p. For small I_0 the space-time variation of p is reasonably close to that of the original population maps shown in Figs. 4.4.

- Case 2 replaces p with the epicentral intensity I_0 (seven categories are used in this case). Consequently, p does not vary in time or space.
- Case 3 uses a smaller penalty coefficient $b_{\underline{x}}$ and thus allows for more local estimates.
- Case 4 uses a local interpolator of $b_{\underline{x}}$ from neighboring cells only, instead of the global average in Eq. 4.131.
- Cases 5 to 8 apply various weighting factors to the likelihood contributions of different I_0.

4.11.5 Discussion of Results

Probability of Detection

Parameter estimates for the probability of detection obtained in the base case are shown in Fig. 4.22. An interesting fact to be noticed is the strong dependence of the detection probability by people on p and the relatively small influence of t. One might conclude that, for r=5, p is a fundamental explanatory variable and that, after such a variable has been included in the analysis, time has only a marginal additional effect. On the other hand, it should be pointed out that the observational periods associated with low value of p in recent times and with high values of p in early periods are small and hence that the statistical uncertainty on these estimates can be large. Another interesting feature in Fig. 4.22a is the general decay of the detection probability since 1950 (the increase of α for categories 5 and higher are due mainly to the constraints and smoothness condition and are not suggested by the data). Figs. 4.22b and 4.22c give the effect of time and site intensity on the probability of detection by instruments. As one would expect, the parameters β and γ are

both monotonically varying with t and d. Of the two, γ is the more variable one. Notice that the produce $\beta_t * \gamma_d$ is smaller than one for all categories. It would be possible to set β and γ equal to 1 for the most recent time period and for the highest value of d if one believes that all earthquakes with site intensity $I \geq 3$ (at the nearest seismic instrument) that occurred in the region since 1970 have been detected by instruments. The influence on the recurrence rates or on the equivalent periods of completeness is however small.

Figure 4.23 shows estimates of α when r is set to infinity. One may note that the variation with time of the estimates is much larger in this case. The effect of this variation on the estimated recurrence rates is a rather substantial one, as will be illustrated later.

The spatial variation of incompleteness is illustrated in Fig. 4.24 by showing the estimated equivalent periods of completeness T^*_{xm} for each size measure and location. The values are shown here for one-degree cells, to be consistent with the spatial discretization of the recurrence rates. A qualitative comparison with earlier results obtained with Model A indicates that the present results are comparable. The most distinct difference between the two models is that in Model A the probability of detection is constrained to one for I_o=8 and hence the equivalent period of completeness is spatially homogeneous and equals the time span of the catalog (356 years). In Model B, where earthquake size enters only implicitly in the probability of detection, this condition is not imposed. Consequently, the incompleteness factor is less than one even for I_o=8.

One way to check the goodness-of-fit of the model with respect to the incompleteness model is to compare the expected and observed counts in each unit-degree cell for the various time periods. As explained in

Section 4.9, it is difficult to obtain exact significance tests for such a comparison, since the expected counts are data dependent, but an exploratory analysis of 'significant' deviation is of interest. Figs. 4.25a and 4.25b show the result of the significance test in Eq. 4.124 for the base case. In the latter figure, a count of one has been added to both expected and observed counts to eliminate flagging of cells with very few or zero events. Interpretation of the symbols is as follows:

= indicates that the observed count is much less than expected ($\alpha<0.02$)

\- indicates that the observed count is 'significantly' less than predicted ($0.02 < \alpha < 0.10$) (4.133)

\+ indicates that the observed count is 'significantly' larger than predicted ($0.90 < \alpha < 0.98$)

* indicates that the observed count is much larger than expected ($\alpha<0.98$)

The pattern emerging from these tests is that the recurrence rate is over-predicted for the most recent time periods in Massachusetts, for parts of New York State and in Southern New Hampshire. Correspondingly, the recurrence rate for this region is underpredicted in the earlier time periods. By contrast, high predicted rates in early periods and low predicted values in recent periods are found for the remainder of the region. Figure 4.25 refers to the base case, where small size measures are only partially considered in the analysis. As shown in Fig. 4.26 a better fit is obtained for Cases 5 and 6, where more weight is given to small size measures (see Table 4.15). This is especially true for Case 6, where all events with $I_o>2$ are weighted equally and many of the '-' and '=' flags disappear, due to the fact that small intensity counts are better fitted. The overall picture remains however the same. One might

propose various alternative explanations for this lack of fit. One is that factors other than population density, seismic instrumentation and time are necessary to explain the variation in time and space of incompleteness. Another is nonstationarity of the earthquake activity at the time scale of a few decades. In any case, it is clear that if one were to estimate recurrence rates based on seismic data obtained after 1900 only, the results would be substantially different.

Recurrence Rates

Estimates of the seismicity parameters a_x and b_x for the base case are shown in Fig. 4.27. The average value of b over the entire region is 1.38. The estimation methods assigns this value to cells with zero counts, since b_x in these cells is undefined and the global average is used as an interpolator. Although differences between neighboring values of b are generally modest, in a few cases differences are relatively large (e.g. b=1.05 near the south-west corner). These differences are due to the fact that the interpolator based on a global average does not effectively remove local spatial variations, unless a very high penalty is used or the earthquake counts are small. Estimates in the cells located at (70.5 W, 39.5 N) and (71.5 W, 39.5 N) are rather particular. As pointed out before, these cells have low counts, but whereas the former contains events with high I_o, the latter contains events with low I_o. Because the value of b in these cells is practically forced to the global average, the expected recurrence rates are very different.

Figure 4.28 shows the empirical earthquake count for the entire region as a function of I_o, the earthquake counts corrected for incompleteness but without assuming an exponential relation for the recurrence rates, and the expected count based on the exponential relation

integrated over the region. Because b is allowed to vary with $\underline{x}$, the latter count is not exactly an exponential function of I_o. The goodness-of-fit is better illustrated by plotting the standardized residuals as proposed in Eq. 4.118. Fig. 4.29 shows such residuals for the base case as well as for Case 6 where all I_o are given equal weight. For the base case also the reduced residuals, when multiplied with the weights used in the analysis, are shown (the sum of those residuals must add to zero, see Eq. 4.70). Examination of the solid line indicates that the exponential fit is good for $I_o<5$ but that the model underpredicts the rate of events with intensity $I_o=4$ and substantially overpredicts the rate for $I_o=2$ and 3. This is what one would expect if the slope of the exponential relation of the true recurrence rates increases with higher intensities.

Estimated recurrence parameters for the various sensitivity cases are summarized in Figs. 4.30 and 4.31. The expected rate at $I_o=2$, 4 and 6 is also tabulated for each case in Table 4.16 to facilitate later comparison with results obtained in Model D. Before discussing each of these results, it is instructive to compare the expected recurrence rate of the various models integrated over the entire region. This is shown in Fig. 4.32 for two sets of analyses: The first set uses weights for different I_o identical to the base case, and the integrated recurrence law is similar. The change from r=5 to r=3 (Base case versus Case 1) and the use of a local instead of global interpolator (Case 4 versus Case 1) for smoothing of b leads to practical identical results. The recurrence law for Case 2, where population density is not used in the model, has a relatively steep slope. This is so, because for $I_o=8$ the probability of detection is set to one and hence counts at large I_o receive more weight

in the likelihood. Case 3 allows more spatial variation of b_x and the associated integrated recurrence relation is more strongly non-exponential.

In the second set of cases, the weights assigned to the various intensities, I_o are varied. This has an important effect on the exponential relation, which is gradually flattened as the smaller intensities are assigned more weight. When all intensities are equally weighted (Case 6), the expected rate at high intensities is rather inaccurate.

The fact that the integrated recurrence law is similar in any two cases does not imply that the local estimates are also similar. Such local variations are discussed next for each case. Case 1 produces values of a and b that are close to those of the base case, except that a is more variable in space. This is to be expected since for r=3, the probability of detection is more dependent on actual population density and is therefore more variable in space.

Case 2 is one for which population has no influence on p. As a result, incompleteness does not change in space (except for the effect of seismic instruments after 1910). This is reflected in the recurrence rates by an increase in highly populated cells and a decrease in sparsely populated areas. The overall increase of the recurrence rates is due to setting P_D=1 for I_o=8, which increases the slope parameter b and, as a consequence increases the recurrence parameters a.

In Case 3, the parameters b_x are allowed to vary more freely in space. It then becomes more apparent that b_x values tend to be lower in the South-West corner than in the central part or North-East of the region. Considering the statistical uncertainty on the slope parameter

b_x (see Section 4.2.2), the spatial variation in the present case or even in the base case is rather extreme. Note also that in this case the peculiar earthquake counts in the cells located at 39.5° N and 70.5-71.5° W is interpreted as a difference in b_x more than in a_x.

The fact that b_x has a spatial trend makes it appropriate to use a local averaging rule rather than a global rule. This is done in Case 4, which emphasises the linear trend.

Increasing the weights for lower intensities (Cases 5 and 6) generally decreases the value of b_x, in some cases quite significantly. The same effect has been noted before globally; however, the effect on the local recurrence rates needs to be clarified. Consider again the cells around 71° W, with latitude 39.5° N. Contrary to the base case, the value of a_x for the cell at 70.5° W is lower than that for the cell at 71.5° W. This inversion is explained by the fact that low and high intensity events are now given equal weight. Hence, with respect to the base case, the parameter a_x tends to decrease if strong earthquakes are known to have occurred. The exponential recurrence relationships fitted in the two cells under Base-Case and Case-6 conditions (Fig. 4.33) give a dramatic illustration of this effect.

Neglecting events of intensity 2 when fitting exponential recurrence relationships (Case 7) produces results similar to those of Case 5, but typically with lower b values due mainly to the large influence of earthquakes of intensity 3.

The last case is a rather extreme one: because only historical events of intensity 5 or greater are used, the estimates of a_x and b_x are based on low counts and subject to large statistical error. Clearly, a trade-off needs to be made between using earthquakes of small intensity to

reduce uncertainty on the estimates and the resulting bias because of the non-exponential recurrence relation. In addition, one should consider reducing uncertainty on the estimates by allowing for a smoother variation of the estimates.

4.11.6 Conclusions

The major novelty of model B with respect to model A is that of using a non-parametric representation of the spatial variation of the recurrence parameters. This assumption serves two purposes:

- To obtain estimates of the probability of detection that are less dependent on seismic source geometry
- To obtain preliminary estimates of the recurrence parameters which may serve as a basis for more strict assumptions on their spatial variation

With respect to the physical process leading to incompleteness, two important changes are made:

1. The temporal variation of detection by instruments and by people is allowed to differ
2. The influence of the epicentral size measure is accounted for in the model by using site intensity at the nearest seismic instrument and an integrated value of the population density in a region surrounding the epicenter as explanatory variables

The benefit of explicitly incorporating the probable causes of incompleteness is that objective (within the assumptions of the model) and spatially detailed estimates of incompleteness can be obtained. As a consequence, the technique allows one to use with some confidence earthquakes of small intensity. These earthquakes are of importance,

since they often delineate spatial variation of seismicity (after proper correction for incompleteness) better than earthquakes of higher intensity, which are sparse. A disadvantage of the proposed technique is that estimation of the parameters is computationally demanding. Also the possibility of examining the goodness-of-fit or evaluating uncertainty on the estimates is limited because of computational constraints. As a result, the selection of an appropriate model and the choice of input parameters is a difficult and partially judgemental process. Further improvement of the model is certainly necessary in that respect.

Another point of concern is that the recurrence rates of earthquakes with small intensity do not appear to follow the postulated exponential variation of recurrence rates. Although this is corrected for by using a weighted likelihood formulation to eliminate bias of the estimates at high intensities, one evidently loses some of the benefits of the model.

The proposed model to correlate incompleteness explicitly to population density and location of seismic instruments should also be considered preliminary in a few respects:

- the redefinition of instrument and population categories to incorporate the effect of earthquake size is rather simplistic: Instrumental measures of earthquake sizes are usually reported in the catalog only when several seismic instruments have been triggered. Differences in the quality of the seismic instruments at different locations are not accounted for. The representation of population density by a single category is undoubtly a simplification. With a better knowledge of the original sources of earthquake reports used in the catalog, one could perhaps consider

alternative or additional explanatory variables such as the location of major cities, missionary stations, communication capability, etcetera.

- Whether or not earthquakes are detected by people or by instruments is an important piece of information, since it allows one to establish the absolute values of the probability of detection, rather than relative ones. In this analysis, the detection mode is based on the presence of an empirical size measure and that of an instrumental size measure. This information appears not very reliable, especially in recent periods, where interest has focused on reporting of instrumental size measures only.
- Regional differences in detection capabilities are unlikely to occur in the small region studied here. In application of the model to a larger region, evidence of such differences has been however found. In particular, it appeared that the level of reporting for Canadian and U.S. earthquakes differs.

In summary, it is thought that the merit of Models A and B depend on the purpose of analysis. If interest is only in the evaluation of seismic hazard and, therefore, recurrence rates at high intensity are most important, a reasonable alternative is to simplify the model by considering only earthquakes with I_o sufficiently large such that 1) the assumption of exponentiality holds, 2) incompleteness can be assumed reasonably constant within prespecified regions without further assumptions on the effect of population or instruments. On the other hand, if interest is in detecting non-stationarity of seismicity over longer periods, in the detailed spatial variation of seismicity to identify seismogenic provinces, in the quality of the seismic network or

in incompleteness of the catalog itself, the statistical technique presented here is thought to be superior.

4.12 APPLICATION OF MODEL C

4.12.1 Introduction

Because of problems in evaluating the goodness-of-fit and the assessment of uncertainty on the estimates, various simplifications to Models A and B are considered in Models C and D. A major modification common to both models is that the analysis is restricted to earthquakes with larger size measure only (i.e., $I_o>3$). For these earthquakes the spatial variation of incompleteness is less important and can be assumed constant within prespecified portions of the region of study. In this case, the population density and location of seismic instruments need not be considered. In each region, incompleteness is a function of time of occurrence and earthquake size.

This approach is not the same as that of Stepp, who proposes to further restrict the analysis to periods over which the catalog is assumed complete. The Stepp approach is the best one can do, if recurrence rates are assumed spatially homogeneous within given seismic sources, incompleteness is different in different sources and no prior information (e.g. smoothness or monotonicity) is available on the probability of detection. In all other cases, estimates of the recurrence rates can be improved by considering also data outside the periods of completeness. For instance, it is clear that, even if earthquakes in a certain time period are incomplete, the relative observed earthquake count at two locations is indicative of spatial variation, assuming that incompleteness

at both locations is the same. The present approach is therefore most relevant when seismogenic provinces smaller than the completeness region or a non-parametric spatial representation of recurrence rates is used.

Basic assumptions used in Models C and D are summarized below:

- Earthquake occurrences follow a Poisson process
- True seismicity is stationary and constant within the cells of a spatial grid or over prespecified seismogenic provinces Ω_k
- Incompleteness is spatially constant within prespecified completeness regions S and varies only with time t and size m

Model C differs from Model D in two important aspects:

- Uncertainty on location is accounted for in the model
- Smoothing of the recurrence parameters is done directly on the earthquake counts

Minor variations on the methods described in Section 4.7 are also used in Model C to impose smoothness on P_D and b_x. However, these are particular to the application and need not be discussed here. Since Models C and D overlap to a large extent, only a few selected results illustrating the effect of the uncertainty on epicentral location are shown here and discussed on a qualitative basis. The smoothing of the earthquake counts has been extensively commented upon in Section 4.7.3.5.

4.12.2 Qualitative Discussion of Selected Results

An exploratory analysis of the catalog used in the application of Model C has been presented earlier in Section 4.8.1 emphasizing the importance of uncertainty on epicentral location and, to a lesser extent, on earthquake size. The results discussed here consider only events with $I_o \geq 4$, for which it is reasonable to assume that incompleteness is

spatially constant inside the entire region of study, which extends from 11°20'E to 13°50'E and 45°N to 46°35'N (see Fig. 4.9). Because seismicity in some parts of this region is relatively strong and spatially variable, unit cells of width 10' along latitude and 5' along latitude are used for the estimation of a_x and b_x. Smoothing of the earthquake counts is imposed separately inside each of the seismic sources in Fig. 4.10f. These sources are also used to define upper-bound intensities as follows: $I_{o,max}$=11 for Source 2 and $I_{o,max}$=9 for Sources 1 and 3. To obtain reasonable estimates of the slope parameters, it was found necessary to include an independent prior distribution of b, in addition to a moderate spatial smoothing. The independent prior estimate is chosen equal to 1.1 in all cases.

Several variants of the incompleteness model were considered in the analysis: After a preliminary analysis, it was decided that the catalog is reasonably complete since 1874 for all intensities above 4. The probability of detection is assumed equal to one also since 1000 for I_o=10 and 11, and since 1700 for I_o=9 and 10. Results given next were obtained analyzing all data within given time envelopes. It is worth mentioning that estimation of incompleteness inside these envelopes creates some problems when the periods differ with magnitude: This is so because, at the boundaries, interpolated values are not well defined and estimates tend to be systematically too large if they are based on a simple average of estimates in neighboring categories inside the envelope. In the present application, this bias has been eliminated by determining the probability of detection using all the data and then keeping this probability fixed when using a time envelope to estimate the recurrence parameters.

The results shown next illustrate:

- the influence of the prior distribution for uncertainty on location
- the variation in time of the spatial pattern of earthquakes
- the influence of the smoothing of the recurrence rates on the redistribution of the earthquake counts due to location uncertainty
- the effect of uncertainty on the size measure I_o

To do so without going into details of the various input parameters, only a short qualitative description of the assumptions made in the various cases is given. The results are also shown in a qualitative form (Fig. 4.34): For each case, contour plots of the recurrence rate at I_o=4 and I_o =6 are shown. The actual values are of no importance, since equal contouring intervals are used in each of the plots and interest is in global variations.

Sensitivity of the results to the accuracy of the epicentral locations is illustrated in Figs. 4.34a,b and c. In all three cases, only earthquakes with I_o larger than 4 and inside the completeness periods are used. Uncertainty on I_o is assumed uniform between the minimum and maximum values reported in the catalog. Moderate smoothing is applied to the earthquake counts to produce a non-erratic variation of the recurrence rate parameters. The slope parameters b_x are further constrained by an independent prior value. Uncertainty on epicentral location differs as follows: Case 1 (Fig. 4.34a) assumes that all earthquakes are accurately located, Case 2 (Fig. 4.34b) uses a prior distribution of earthquake location that varies linearly from 1 at the center to 0 at the radii in

Table 4.3, Case 3 (Fig. 4.34c) uses radii that are 50% larger than in Case 2. The change in recurrence rate estimates at I_o=4 going from Case 1 to Case 2 is very large. This is not surprising, because also in recent times the number of events with inaccurately determined location is large (i.e., Fig. 4.10.c). Especially earthquakes in the Eastern part of the region are relocated. Case 3 shows the effect of increasing the radius of uncertainty. This effect is less pronounced, although still substantial. It is interesting that the contour plots for I_o=6 show much less contrast between the various cases. This is so because the b_x value is small in the North-Eastern part of the region, which therefore dominates the spatial picture at high values of I_o. Apparently, the recurrence rate in this North-Eastern region is also relatively stable with respect to uncertainty on location.

The effect of extending the time periods to include incomplete parts of the catalog is more influential. Case 4 (Fig. 4.34d) corresponds to analyzing all earthquakes since 1700, 1500 and 1000 for I_o less than, equal to and larger than 8, respectively. Cases 5 and 6 (Figs. 4.34e and f) correspond to an analysis of the entire catalog with and without considering uncertainty on location, respectively. If one compares these results with Case 2 (where only events inside the completeness periods are used), it is clear that there is a gradual spread of seismicity towards the Southern and Central Eastern parts of the region, while seismicity in the Northern region decreases significantly. It would lead us too far to comment on individual differences. The trend is consistent with earlier observations during exploratory analysis of this catalog. The overall decrease in seismicity at high intensities is due to the fact that

relatively few earthquakes of large intensity are reported in early periods of the catalog and hence the b parameter increases. Case 6 where no uncertainty on location is considered, is shown here only to illustrate the amount of spatial relocation of events, especially for early periods.

The variation of seismicity in time is dramatically illustrated in Cases 7 to 9 (Figs. 4.34g,h and i) where the analysis is performed using only a portion of the catalog, 1000 to 1699, 1700 to 1873, 1874 to the present respectively, while fixing the incompleteness parameters. Case 9 differs from Case 2, because also for large I_o only the last 110 years are used as periods of observation. Because most of the large earthquakes in this region have occurred in recent times, this further increases the recurrence rate estimates.

Smoothing of the recurrence parameters a_x (here achieved by smoothing the counts) is also influential on the estimates. Cases 10 and 11 (Figs. 4.34j and 4.34l) illustrate this effect. These cases are variants of Case 2, which is also shown for ease of comparison (Fig. 4.34k). Case 10 is rather extreme and does not impose any smoothness on a. In this case, all earthquakes at $\underline{x}$ may be completed relocated if all reported locations $\underline{x}$ are subject to measurement error. Such a solution is not a very stable one and uncertainty on the estimates is likely to be very large. Smoothing of the estimates, as is done in Case 11, however stabilizes the solution. The choice of an appropriate smoothing level is not evident and some judgement is required. For instance, Case 11 which uses more smoothing than Case 2 is thought to be excessively smooth and obscures information in the actual data. On the other hand, it should be noted

that the entire area of Source 2 has been proposed as being homogeneous based on geophysical information.

The redistribution of the counts due to uncertainty on the size measure versus that due to location uncertainty is illustrated in Fig. 4.35 for the case when all earthquake data are analyzed (Case 5) and on an aggregated basis for Sources 2 and 3. Presented in this figure are 1) the actual earthquake counts based on reported location and average I_o, 2) the redistributed count when only uncertainty on I_o is considered, 3) the redistributed count when uncertainty on I_o and epicentral location is analyzed. The effect of uncertainty of I_o is most visible for large I_o, where the method redistributes the counts up to the upper bound value of I_o in each source. Globally, the redistribution tends to be such that the exponential relation is better satisfied. The effect of uncertainty on location is larger than it would appear from this figure, because summing the counts over each source does not show the redistribution of counts within each source. In total, the effect is one of relocating earthquakes of Source 3 to 2.

In summary, seismicity in the geographical region used in this analysis has a rather peculiar behavior. The fact that even after considering uncertainty on epicentral location, temporal variations of the spatial distribution remain, supports the hypothesis of earthquake migration. Given the small size of the region, it appears unlikely that the observed effect might be explained through spatial variation of incompleteness only. With respect to the method, the present application illustrates the importance of considering also incomplete periods for the interpretation of the seismic data. On the other hand, for the purpose of

seismic hazard, use of only the last time intervals appears most sensible. In this particular case, early events in this catalog are likely to introduce bias and not to improve the recurrence rate estimates for future seismicity. Consideration of uncertainty on location, even in recent periods, becomes more important if one attempts to model spatial variations on a smaller scale or considers recurrence rates at low I_o. Uncertainty on the size measure is thought to be less important, but should be considered at least for earthquakes with large I_o.

4.13 APPLICATION OF MODEL D

4.13.1 Review of Assumptions and Methods

The basic assumptions used in Model D are identical to those of Model C: true seismicity is assumed to follow a stationary Poisson process, the spatial variation of which is modelled on a spatial grid. Incompleteness is assumed spatially homogeneous within prespecified regions S_1 above a given size measure, but varies with time t and size m. Contrary to Model C, uncertainty on epicentral location is however not accounted for and a penalized maximum likelihood formulation is used to smooth the spatial variation of recurrence parameters a_x and b_x. Uncertainty on the size measure m is allowed for.

Some details of the solution techniques are briefly discussed next. The incompleteness parameters α_{tm} (Eq. 4.30) are smoothed by penalizing deviations with respect to a local average based on neighboring values. Because using a local average tends to increase the estimates of α_{tm} at the boundaries (i.e., for categories t=1 or m=0), the penalty coefficient P_α is decreased with a factor 1/2 and 1/4 for α_{tm} along a boundary or on a

corner respectively. Smoothness of the recurrence parameters is imposed in a similar fashion and uses penalty coefficients P_a and P_b for recurrence parameters $a_{\underline{x}}$ and $b_{\underline{x}}$ respectively. In this case no correction is made at the boundary. If the region is large, the boundary effects are small, and if the region is small, smoothing towards an average value is reasonable. The problem of slow convergence of the maximum likelihood estimates for a region with many cells $\underline{x}$ and a linear trend of the estimates, has been explained in Section 4.7.3.4. The second method, which uses a penalty term explicit in the parameters, is used and no problems of convergence were encountered in its application. Also discussed in Section 4.7.3.4 and applied here is the correction at each iteration to balance the total expected and observed counts and magnitudes over the region.

Uncertainty on earthquake size has been treated by iteratively calculating the posterior distribution of size for each earthquake and by accordingly redistributing its unit count over the various categories (Eqs. 4.109, 4.110 and 4.111). Events reported in the chosen size measure, here I_o, are treated differently from those reported in an alternative scale. In the former case, when I_o is reported, the posterior distribution depends on the probability of detection. In the second case, when only an instrumental size measure is reported, I_o is unknown and no correction for P_D is applied (for a more detailed discussion, see Section 4.8.4).

In addition to spatial smoothing of the $b_{\underline{x}}$ parameters, a penalty term $P\hat{b}$ is included in the log-likelihood that penalizes deviations of $b_{\underline{x}}$ from a prior value of b, which is independent of location $\underline{x}$. Estimates of $b_{\underline{x}}$

are also constrained to the interval [0.5,2.0]. Because of spatial smoothing and the independent prior used in the following results, all estimates of $b_{\underline{x}}$ fall inside this interval without activating the constraint. The rate parameter $a_{\underline{x}}$ is assumed larger than -7.0, which simply prevents estimates from going to $-\infty$ when the count at $\underline{x}$ is zero $\alpha\nu\phi$ no spatial smoothing is used for $a_{\underline{x}}$.

In application of Model D to the data in the Chiburis catalog, it was found that with appropriate smoothing of P_D, monotonicity with time t and size m is satisfied, except for some minor violations in a few categories. To avoid smoothing P_D too much to correct for this problem, a simple heuristic change is made to the method that leads to monotonic estimates:

- Violations of monotonicity for neighboring cells α_{tm} and m fixed are checked first. If such a violation occurs, e.g., $\alpha_{tm} < \alpha_{t-1,m}$, the estimate in these categories is replaced with one found from pooling the counts and observational areas in those two cells together and by penalizing deviations from an averaged interpolated value. To impose monotonicity on the entire set $\{\alpha_{tm}\}$, this estimate is further restricted to be larger than or equal to $\alpha_{t-2,m}$ when necessary
- Next, the same procedure is applied for fixed t, with the additional constraint that the new estimate should be larger than estimates for the same t, but lower m, i.e., $\alpha_{tm} > \alpha_{t,m-1}$
- Finally, interpolated values α_{tm} are calculated as usual based on the modified estimates and new penalized maximum likelihood estimates α_{tm} are found.

Obviously, such a method does not necessarily converge to final estimates of α_{tm} that are monotonic. Therefore, convergence is only checked on α_{tm} prior to imposing monotonicity, and, when convergence is reached, the modified monotonic estimates are used. In this particular application, where monotonicity is nearly satisfied to start with, the technique is considered acceptable. In a more general case, alternative techniques that are likelihood based should be considered

To quantify uncertainty on the estimates, a bootstrapping technique is used in Model D. Both a parametric and an empirical version of bootstrapping are applied. In the former, the estimated incompleteness and recurrence parameters are assumed to be the true ones. In that case, an artificial sample can be generated by simulating the earthquake count for each category ($\underline{x}$,t,m). These counts have Poisson distribution with expected value determined by the recurrence rate, the probability of detection and the period of observation in each category. Note that the total size of each simulated sample is not constant but rather a random variable with Poisson distribution. Uncertainty on earthquake size is neglected in this method. In application of the empirical bootstrapping method on the other hand, each sample is generated by random selection of earthquakes from the actual catalog without replacement, until the original sample size is reached. The uncertainty on the size m of selected earthquakes is treated as usual in this case. Technically, the second method is the simpler one, although possibly computationally more demanding. The relative benefits and disadvantages of both methods will be discussed later in this section.

4.13.2 Earthquake Data and Distribution of Explanatory Variables

The earthquake catalog used is that of Chiburis (1981). As in Model A, earthquakes tagged as aftershocks in this catalog have been removed prior to the analysis. To do so, the original identification found in the Chiburis catalog is used.

Only earthquakes with true $I_o>4$ are used in this analysis. Thus, category m=0 corresponds to $I_o=4$, category 5 corresponds to the largest intensity found in the catalog, $I_o=8$. The accuracy level of the size measure for different earthquakes is consistent with that assumed in Section 1.5, where a deterministic correction is proposed. If I_o is reported but $\Delta_{I_o} = I_{o,max}-I_{o,min}$ is not zero, the prior distribution of I_o is assumed to be normal with mean value $(I_{o,min} + I_{o,max})/2$ and $\sigma_{I_o}=0.5$ and 1 for $\Delta_{I_o}=1$ and 2, respectively. The normal distribution is truncated at $\pm 3\ \sigma_{I_o}$ and discretized to a mass density function p'_m for different categories m (including m<0). The posterior mass density function p''_m is then assumed proportional to:

$$p''_m \propto p'_m\ P_D \exp\{-b_{\underline{x}}m\} \qquad (4.134)$$

All parameters in Equation 4.134 are earthquake dependent. For instance, P_D refers to the probability of detection at the time of occurrence of the earthquake and $\underline{x}$ refers to its epicental location. For m<0, the probability P_D is assumed equal to P_D for m=0. Finally, p''_m is normalized and then truncated for m<0.

Treatment of earthquakes with I_o not reported is similar, except that the prior distribution is assumed to be normal with mean value

$$E'[I_o] = (M - 1)/0.6 \qquad (4.135)$$

where M is the reported instrumental size measure and E' refers to the prior expected value of I_o. The standard deviation σ_{I_o} is assumed equal to 0.6. One should note that Eq. 4.135 is an estimate of the prior value of I_o, i.e., independent of the marginal distribution of I_o, and should be interpreted as $E[M|I_o]^{-1}$. Another difference of the treatment of uncertainty on m is that for I_o not reported, the factor P_D in Eq. 4.134 is not included.

Based on a preliminary analysis of the data, it is decided that two completeness regions are sufficient to capture the spatial variation of incompleteness. These regions are shown in Fig. 4.36. Basically, the coastal region of the U.S., which has been settled evenly in the early periods of the catalog, is separated from the remainder of the region. The simplicity of this configuration follows from the sparseness of the earthquake counts in much of the region. For instance, it may appear strange that locations in the Atlantic Ocean are not treated as a separate completeness region. The recurrence rate in this area for the intensity interval considered here is however so low that 1) incompleteness cannot be determined separately, and 2) adding the region to areas over land does not introduce any change in the estimates α_{tm}. Without smoothing, α_{tm} is determined as the ratio of observed to expected true counts and both are almost zero for locations over the ocean. An alternative and perhaps better solution is to assign for this region values for P_D based on earlier analyses accounting for population density and seismic

instruments. Such a solution is for instance necessary if a seismogenic province with spatially constant recurrence rate covers part of the ocean. In this case, the true recurrence rate is no longer close to zero, and P_D could also be estimated and would be small. From a practical point of view, the present choice of only two regions corresponds to assuming that recurrence rates are small in this part of the region. Including a separate region to account for early settlements around Quebec and Montreal has been also considered. In this case, it was found that estimates of P_D are very similar to those in the surrounding region.

The temporal variation of seismicity in both regions is illustrated in Figs. 4.37 and 4.38. In these figures, I_o corresponds to the expected value of the prior distribution. The most characteristic feature of the usual plot of cumulative recurrence rate versus period of observation (Figs. 4.37a and 4.38a) is that I_o=6 appears incomplete more recently than I_o=4 or 5. To aid in the interpretation of these figures, some alternative representations of the temporal variation of seismicity are also shown. Figs. 4.37b and 4.38b correspond to the recurrence rate estimated over different time periods. The first period starts from the present and is chosen such that the recurrence rate is maximum. The periods that follow are determined similarly after shifting the origin of time. This procedure evidently enforces a monotonic decrease of the recurrence rate and corresponds to non-parametric maximum likelihood estimation of a monotone density (Groeneboom et al., 1983). Although these figures exemplify the overall temporal variation of seismicity, clearly many of the small jumps in the rate density are non-significant. The picture is greatly simplified if, for each I_o, one merges subsequent time periods for which differences are small. This is done in Figs. 4.37c

and 4.38c using a statistical test: Starting from the most recent period, the significance of the difference in recurrence rate with the next time period is calculated under the assumption that the recurrence rate is the same. Earthquake counts in both periods are assumed independent Poisson and a test similar to Eq. 3.6 is used. Because the time periods correspond to maximum values of the recurrence rate, this is evidently an approximation. When the difference in rate is found non-significant, the two time intervals are merged, and the next time interval is compared with the merged one. If the difference is significant, the first time interval is fixed and the procedure is repeated to merge subsequent time intervals. The length of the first time interval can be thought of as an estimate of the period of stationarity, or of completeness. Figs. 4.37c and 4.38c show the result of this procedure for a rather moderate amount of merging (the significance level used is 0.2 but should not be strictly interpreted as such). Figs. 4.37d and 4.38d show the result of reapplying the same method to the already simplified results using a smaller significance level. The reason for doing so is that it is clearly better to merge first the most obviously close time intervals. It is unlikely that the first time periods in the last plots are periods of complete reporting, because the abrupt and large changes in the recurrence rate and hence in the probability of detection do not appear realistic. Estimates for moderate merging are more consistent with the original Stepp plots, but also exemplify the problem with such estimates: exponential variation of the recurrence rate with I_o is clearly violated for Source 1 and I_o=4, and the periods of 'completeness' do not increase monotonically with I_o. In view of these problems, the time categories used in Models A and B have

been maintained and only moderate assumptions on completeness have been made, as explained in the next subsection.

4.13.3 Prior Information

Several states of prior information have been used in application of the model. A reference case has been defined first, for which the input is described here. Deviations from this input will be indicated in the discussion of results and summarized in the following subsection.

For the probability of detection, it is assumed that all earthquakes independent of I_o, have been reported since 1950; hence

$$\alpha_{tm} = 1 \qquad \text{for } t = 5 \tag{4.136}$$

It is further assumed that, for I_o=7 and 8, the catalog is complete since 1860 and 1625 respectively, so that

$$\alpha_{tm} = 1 \qquad \text{for } t=4,5 \text{ and } m=4$$
$$\text{for } t=1,\ldots,5 \text{ and } m=5 \tag{4.137}$$

The penalty coefficient P_α which regulates the smoothness of the variation of α_{tm} with its subscripts is taken as

$$P_\alpha = 20 \tag{4.138}$$

This corresponds to a moderately smooth change of the estimates. As explained before, monotonicity of the estimates is also imposed. In the reference case, the location vector $\underline{x}$ is discretized according to unit-degree cells. Because the size of these cells is reasonably large, no smoothing is imposed on $a_{\underline{x}}$, i.e.,

$$P_a = 0 \tag{4.139}$$

Smoothing of the estimates b_x is however required to obtain reliable estimates. The penalty coefficient used to pull the estimates of b_x towards a locally linear trend is chosen as

$$P_b = 50 \tag{4.140}$$

Also, an independent prior on b_x is used, with mean value

$$\hat{b} = 1.3 \tag{4.141}$$

and associated penalty coefficient

$$\hat{P_b} = 10 \tag{4.142}$$

$\hat{P_b}$ refers here to a cell with unit-degree equatorial width and is scaled according to the area of the cell it is applied to. The present value in Eq. 4.142 introduces only a moderate amount of prior information, which can be seen by calculating the corresponding standard deviation used in the prior distribution: for a unit cell in the present analysis,

$$\hat{\sigma_b} = (0.71\ \hat{P_b})^{-1/2} = 0.38 \tag{4.143}$$

where 0.71 corresponds to the area of a cell at 45 degrees latitude.

4.13.4 Analysis Cases

Input parameters to the analysis have been varied primarily to demonstrate sensitivity of the results to assumptions on P_D and to the smoothing of the recurrence parameters a_x, and b_x. Also considered is the influence of uncertainty on earthquake size and the width of the spatial discretization. Two versions (one empirical, the other parametric) of bootstrapping are used to evaluate uncertainty on the estimates for the reference case. Input parameters for other cases are summarized below:

- Case 1 : no uncertainty on I_o. In this case, a deterministic correction $-0.5b\sigma 2_{I_o}$ is applied to the expected prior value of I_o and then the nearest integer is used.
- Case 2: P_α=5, allowing for more data-dependent estimates of α_{tm}
- Case 3: Only part of the earthquakes in completeness region 2 is used in the analysis. Specifically, the following time intervals are used:
  * for I_o=4 and 5, only earthquakes since 1860
  * for I_o=6, only earthquakes since 1780
  * for I_o=7 and 8, all earthquakes since 1625
- Case 4: α_{tm} is not constrained to 1 for m=0(I_o=4) and t=5 (since 1950).
- Cases 5 to 9: correspond to variations of the prior information on a_x and b_x. Deviations with respect to the reference case are as follows
  * P_b = 12.5 for Case 5
  * P_b = 200. for Case 6
  * $\hat{p_b}$ = 2.5 for Case 7
  * $\hat{P_b}$ = 40. for Case 8
  * P_a = 10. for Case 9
- Case 10: uses a weighted likelihood formulation to improve the fit of the exponential recurrence relation to data with large I_o. To do so, earthquakes with I_o=4 are weighted as

  $w(I_o=4) = 0.2$

- Cases 11 to 13: the earthquake data is analyzed using half-degree cells. In these cases, the smoothness of a_x is controlled as follows

  * P_a = 0.0 for Case 11
  * P_a = 2.0 for Case 12
  * P_a = 10.0 for Case 13

Results from the bootstrapping techniques are only for base case input.

Some comment on the presentation of the results is appropriate. Numerical results for all cases are shown in Table 4.17. For each case, the expected earthquake count over a 100-year time interval and for a unit-degree equatorial cell is shown for I_o=2, 4 and 6 at each location. Results obtained with model B have been presented in the same format in Table 4.16 for easy comparison. For cases with significant differences of a_x and b_x, contour plots of the recurrence rate at I_o=4 (per 100 years and unit-degree equatorial area) and of b_x are shown over the region of study. The contouring interval of b_x in these plots is 0.1 (i.e., a label 13 corresponds to b=1.3). The contouring interval for the recurrence rate equals 5. One may note that the algorithm used in producing these plots is a very simple one and produces jagged contours. For the present purpose of comparing results, these figures are however useful.

4.13.5 Discussion of Results

The effect of explicitly including uncertainty on the historical earthquake magnitudes in the likelihood is not very large as far as recurrence rates are concerned. (Compare the contour plots for Case 1 and the Reference Case in Figs. 4.40a and 4.40b). The recurrence rate at I_o=4 increases somewhat in the North-West corner and the parameter b_x increases

in the East, when uncertainty is not considered. Changes in the probability of detection are also moderate (see Fig. 4.39), except for some isolated categories. For completeness region 1, neglecting uncertainty consistently produces lower estimates of P_D. As will be shown later, this effect is similar to that from reducing the smoothness of α_{tm}. The reason why the two operations may be equivalent is that, when uncertainty is considered, the earthquake counts can be redistributed; hence smoothing is facilitated and has more effect. When uncertainty is neglected, higher estimates of P_D are obtained for small I_o in completeness region 2. To understand the estimated values in detail it is necessary to consider the difference between the earthquake counts used in each case: a-posteriori earthquake counts for the Reference Case and the usual earthquake counts (after a deterministic correction) for Case 1. Actual counts in each category, and expected earthquake counts predicted by the model are shown in Table 4.8. In general, differences are small. For large I_o, the Reference Case typically produces larger counts, which may be expected because the a-posteriori counts reflect the increase of the probability of detection with increasing I_o (if I_o is reported), whereas the deterministic correction does not. The fact that, in the Reference Case, the counts for I_o=4 are lower explains the relative decrease of P_D in region 2. For completeness region 1, the decrease is probably counteracted by the smoothing effect. The systematically lower count at I_o=4 must be due to the discretization. For example, if many earthquakes have a value of I_o between 3.5 and 4.5 after the deterministic correction, they are classified as I_o=4 in Case 1. On the other hand, in the Reference Case, only a fraction of those low-intensity counts is used,

since part of the posterior distribution is outside the range of analysis. Although normally this would be balanced by fractional counts from earthquakes with $I_o<3.5$, these counts are fewer. This fact is not considered in the Reference Case, where it is assumed that P_D for I_o below the range of analysis equals P_D for the lowest intensity interval $I_o=4$.

Cases 2, 3 and 4 all consider variations of the incompleteness model. Of these cases, only the last one, where the probability of detection for $I_o=4$ is no longer fixed, produces substantial differences in the recurrence rates (Fig. 4.40c). The increase of the recurrence rates at $I_o=4$ is however offset by a corresponding increase of the slope parameter b, and the effect at $I_o=6$ is less important (Table 4.17). It should be mentioned that considerably larger estimates of b_x would have been obtained without the constraining effect of the independent prior on b: relative to the Reference Case the prior has more effects on the estimated slope parameters. The probability of detection for Cases 2, 3 and 4 are shown in Fig. 4.39. Case 4 evidently predicts very low values of α for $I_o=4$ and illustrates the importance of constraining P_D in recent time periods. Allowing for a less smooth variation of α_{tm} (Case 2) typically produces lower estimates of α. This is due to the fact that the local averaging rule used in this model to calculate interpolated values tends to increase the estimates. The effect of using a time envelope in Case 3 is not very large, except for the estimate of α_{tm} at t=2 and m=3($I_o=6$). Again this is due to the averaging rule, which for this (t,m) category calculates interpolated values close to 1., because the probability of detection at lower intensities is unknown and cannot be used in the interpolation. The significance of these deviations is better appreciated

if one considers the minimum and maximum estimates obtained in 50 empirical bootstrapping samples. It is clear that the statistical variability on the estimates is high and deviations due to the prior information are relatively moderate. Assumptions on the constraints (as in Case 4) appear however important. One should not draw the conclusion that, given the large uncertainty on these estimates, one might as well not use the data in the incomplete time periods. If one is correct in assuming that incompleteness is spatially constant within the given incompleteness regions, then earthquake counts are important to estimate the spatial variation of seismicity. It is true, however, that the actual level of seismicity is primarily determined by the counts in periods when α_{tm} is constrained to 1.

Cases 5 to 9 show the influence of varying the smoothness of b_x and a_x. Case 5, the Reference Case and Case 6 illustrate the effect of increasing the value of P_b. For the lowest value of P_b (Case 5), the spatial trend of increasing b_x from the South-West to the North-East as well as the local maximum of b_x in eastern Massachusetts is very clear. Increasing P_b gradually removes these features, first the local maximum, then the overall linear trend, which appears to be quite strong. In Case 6, b_x is practically constant and equals the prior mean value 1.3. Although there is a slight change in the a_x estimates which counteracts the increase and decrease of b_x, the global effect at high I_o is to increase the recurrence rates for areas in the central part of the region and to decrease the rates in the North-East corner (see Table 4.17).

Cases 7 and 8 vary the influence of the independent prior $\hat{b}$. For low values of $P_{\hat{b}}$ (Case 7) the linear trend of b is more pronounced and this case is similar to Case 5. Higher values of $P_{\hat{b}}$ produce results similar to Case 6. The effect of P_b and $P_{\hat{b}}$ thus appears interchangeable in the present case. Both parameters are however necessary. For instance, setting $P_b=0$ and only applying penalties to deviations from $\hat{b}$, would produce a more erratic fluctuation of the b estimates (e.g., compare the present results with those of Model B). On the other hand, if $P_{\hat{b}}$ is set to zero, then the global linear trend is too extreme, unless very high values of P_b are used, so that boundary effects become important and again b becomes independent of $\underline{x}$.

Smoothing of $a_{\underline{x}}$ is of less interest for the cell size used in the reference case but is illustrated more extensively when smaller cells will be used in the analysis. Case 9 applies moderate smoothing of $a_{\underline{x}}$ to the Reference Case. The peak values in the Massachusetts area are especially influenced. It is worth mentioning that, contrary to the other cases, the incompleteness parameters changed substantially. In the second completeness region, the probability of detection increases whereas in the first region the same probability decreases. This is to be expected, since spatial smoothing of the recurrence rates tends to decrease the higher recurrence rates in the second region and correspondingly increases the recurrence rates in the first region. In general, it is clear that assumptions on the spatial continuity of the recurrence parameters <u>across</u> the boundaries of the completeness regions can be influential on the relative values of the probability of detection in both regions, when such continuity is not suggested by the data.

Comparison of observed and expected global counts for each value of I_o (Table 4.18) shows that, in the Reference Case, counts for I_o=4 are overpredicted, whereas counts for I_o=5 are underpredicted. One would expect this trend if the slope of the exponential relation increased with higher values of I_o. Although the trend is less pronounced than in earlier analyses (Models A and B), one might correct for this problem by using a weighted likelihood formulation. Case 10 shows results when earthquakes with I_o=4 are down weighted by a factor 0.2. Relative to the reference case, b_x is spatially more constant. This is so, because the data are less informative on the actual value of b and as a result the influence of the prior value $\hat{b}$ is larger. Values of a_x are higher in this case, because the relative low counts at I_o=4 are weighted less in the analysis. The combination of these two effects produces higher expected recurrence rates over the entire I_o range (see Table 4.17).

Cases 11 to 13 use smaller cell sizes to model the variation of a_x and b_x. Smoothing of a_x is gradually increased in the three cases. Because the spatial smoothing of b is left unchanged, the estimates of b are found to be slightly more variable than in the reference case. It is interesting to notice that the spatial smoothing of a eventually produces results (Case 13) very similar to those in the reference case. Considerable more detail is found in the spatial variation of a_x for lower values of P_a. Theoretically, one might test whether such spatial variation is statistically significant by applying a bootstrapping technique to evaluate uncertainty on the estimated parameters. Although such an analysis is not performed here, it appears from the results shown next for the Reference Case that the significance is low (See for instance

the minimum and maximum values of a obtained in 50 samples). Of course, such a result does not incorporate any prior knowledge, one might have, based on geophysical or seismological considerations.

To assess the uncertainty on the estimated parameters an empirical and parametric bootstrapping technique has been applied to the Reference Case. In each case, 50 samples have been generated and analyzed using the same parameters as in the Reference Case. In empirical bootstrapping, the sample size has been fixed to be the same as the original sample size. Alternatively, one could have used for each sample a size generated according to a Poisson distribution with expected value equal to the original sample size. The latter is a better approach if interest is not only in the relative recurrence rates over space, time, and magnitude but also in their absolute values. The additional uncertainty due to a variable sample size is however small, at least when N is large. In applying parametric bootstrapping, no independent prior has been used on b (i.e. $P_{\hat{b}}=0$). The reason why this is necessary requires some explanation: if the penalty terms on the spatial variation of the recurrence rates are included also in the estimation procedure when applied to the artificially generated samples, then the estimates obtained from these samples are smoother (and biased) with respect to the true parameters used to generate the samples. It is important to note that this does not imply that the estimation procedure always produces biased estimates: in an ideal application of MPL, the penalty coefficients are not fixed a-priori, but are determined on basis of the obtained data sets. Such a procedure is however not a very practical one in the case of bootstrapping, where many such samples are generated. If however the penalty terms are interpreted as a-priori distributions of the recurrence parameters, then

it is not clear how the parametric bootstrapping method should be modified. Although neglecting the prior information in the estimation procedure when applied to the artificially generated sample seems intuitively valid, more work is necessary to resolve this issue. Because of this problem, the empirical bootstrapping technique is thought more appropriate this time for evaluating uncertainty on the estimates.

Summarizing the results from bootstrapping is in general not a simple task, because many are simultaneously estimated and there is correlation among the estimates. Fig. 4.41 shows selected results for the first 20 samples in empirical bootstrapping. Fig. 4.41a shows estimates of α_{tm} as a function of t and for different m, in both completeness regions. The sample average and minimum and maximum values have been shown earlier in Fig. 4.39. Variation of the parameters is obviously large and has been commented upon earlier. As one would expect, the parameters are also strongly dependent due to the imposed smoothness condition. Note for instance how the different lines predominantly shift up or down, with relative few crossings. Fig. 4.41b shows estimates of a_x and b_x as functions of longitude, for different latitudes. One may note that, whereas the a_x are relatively independent (see the large number of crossings), estimates of b_x are more dependent and tend to produce parallel lines. Another point of interest illustrated by this figure is that, at locations with zero count, the empirical bootstrapping technique produces always the same estimate a_x=-7. because no smoothing is applied to the estimates and the count at those locations is zero in all samples. This not true for the parametric bootstrapping where the expected recurrence rate is used to generate counts in these cells. In practice,

however, the estimated recurrence rate is so low that an enormous sample size is required to actually generate any of these counts.

To assess uncertainty on a single parameter or on any function of the parameters, one may calculate statistics from the generated sample. Fig. 4.42 shows for instance the sample average, standard deviation and minimum and maximum estimates for $a_{\underline{x}}$ and $b_{\underline{x}}$. Table 4.19 shows the numerical value of the average and standard deviation for both a and b. The contouring intervals for the standard deviation of a and b in the figures are 0.5 and 0.025, respectively. Consider first the results of parametric bootstrapping (Fig. 4.42a). The standard deviation of $a_{\underline{x}}$ increases with the estimated average, while the coefficient of variation decreases, as one would expect. The standard deviation of $b_{\underline{x}}$ is reasonably constant, but increases at the boundaries. Again this is to be expected intuitively, since for values at the boundary, spatial smoothing is less effective. It is also interesting to compare results obtained from empirical and parametric bootstrapping. Whereas results for $a_{\underline{x}}$ are nearly identical, the values of $b\underline{x}$ are quite different. This is of course due to the fact that, in parametric bootstrapping, the independent value of b is not used, whereas the same value is used in empirical bootstrapping.

Ultimately, the interest of statistically analyzing earthquake catalogs is to evaluate the seismic hazard at a given site. Uncertainty on seismic hazard estimates can be separated as follows:

1. Uncertainty due to model assumptions such as the value of the smoothing parameters or other prior information based on judgement rather than on data.
2. Uncertainty due to the limited size of the sample.

3. Uncertainty due to other parameters in a seismic hazard analysis such as the attenuation law and the upper-bound magnitude, which are not discussed in this thesis.

As a simple illustration of the magnitude of uncertainty on seismic hazard, the recurrence rate of earthquakes with site intensity larger than I is calculated for the Boston area (45°20'N and 71°10'W). The modified Guppta-Nuttli attenuation function (Eq.4.128) is used without considering attenuation uncertainty. Seismic hazard curves are calculated for all cases considered in Models B and D. Fig. 4.43 shows six curves that envelope all results.

'DR' and 'BR' refer to the reference cases of Models D and B respectively. 'PR', 'PR+' and 'PR-' correspond to the sample average and the sample average ±2 standard deviations from parametric bootstrapping (the uncertainty band obtained from empirical bootstrapping is narrower, because of inclusion of independent information on b). Finally, 'B6' corresponds to sensitivity case 6 in Model B, for which all earthquakes with $I_o \geq 2$ are used equally in the likelihood formulation. Although not shown in this figure, it is worth mentioning that all sensitivity cases considered in Model D fall inside the uncertainty band from parametric bootstrapping. This suggest that the sample size is at least as important as the model assumptions. It is also interesting that results obtained from 'BR' do not significantly differ from 'DR'. Presumably, the seismic hazard results are reasonably stable, because they are dominated by the historical events at large intensities which in both cases are fitted reasonably well. Case 'B6' on the other hand deviates considerably from the other results and indicates the importance of the assumption of

exponentiality of earthquakes if small intensity are included in the analysis.

To combine the various results into a single seismic hazard prediction, the credibility of the various curves should be established. It should also be emphasized that uncertainty on the upper-bound magnitude and on attenuation are not considered here. In addition, the results are based on the estimation of recurrence rates inside one-degree and half-degree cells. Whereas more local estimates are not thought to alter the estimates very much, the assumption of spatial homogenity inside large seismogenic provinces might do so.

4.13.6 Conclusions

Model D uses a statistical model for earthquake occurrences that differs from Models A and B in two basic aspects: 1) The spatial variation of incompleteness is judgementally defined, 2) Uncertainty on historical earthquake sizes is accounted for. Major differences with a more traditional analysis of the data are that: 1) Incompleteness is corrected for by estimation of the probability of detection, 2) A non-parametric representation is used to model spatial variation of the recurrence rates.

The application of the model to the Chiburis catalog indicates that explicitly considering uncertainty on the earthquake sizes does not substantially alter the results. Of course, such a conclusion is data dependent and should not be generalized to other catalogs or geographical regions. On a theoretical basis, the deterministic correction proposed in Section 2 is considered sufficiently accurate if interest is on obtaining best estimates of the recurrence rates or of the seismic hazard. Note

however that, in using this correction, the effect of the probability of detection is not considered and that the increase of uncertainty on the estimated values cannot be assessed. Thus, if an exploratory analysis shows that uncertainty on the earthquake sizes is large for a substantial portion of the historical catalog, then the likelihood based approach is recommended.

Comparison with results obtained earlier with Model B shows that, for large intensities, the results of the two models are comparable. This indicates that, for the purpose of seismic hazard calculation, the simplification in modelling the spatial variation of incompleteness is justified. If also earthquakes of small intensity are considered, for example to delineate regions of different seismic activity, Model B is considered more appropriate.

Evaluation of total uncertainty on the results of interest (e.g., on seismic hazard) is a complicated task, because uncertainty from many sources needs to be combined. The use of empirical and parametric bootstrapping to evaluate uncertainty due to limited sample size has been illustrated. In parametric bootstrapping, the problem arises of including uncertainty on earthquake sizes and of specifying judgementally determined input parameters for the artificially generated samples. In this regard, the empirical bootstrapping is easier to use and thought to be more appropriate.

CHAPTER 5

SUMMARY AND CONCLUSION

Several new methods are proposed in this thesis to address three major problems in the statistical analysis of earthquake catalogs: the conversion of different magnitude measures to a single scale, the identification of earthquake clusters, and the estimation of incompleteness and recurrence rates. Techniques that are currently used to account for these problems and their limitations have been identified earlier in Chapter 1. In the present chapter, the innovations introduced in this study are summarized and main conclusions are stated. Topics that should be subject to further research are also indicated.

5.1 MAGNITUDE CONVERSION

Earthquakes are typically reported in different magnitude scales, however, many operations are greatly simplified if earthquake size is expressed in a single scale. Chapter 2 considers this "magnitude conversion" problem and proposes a method which has following distinguishing features:

- the regression of m against other size measures may be nonlinear and the residuals need not have constant variance. Outliers present in the data set can be identified or removed.
- measurement errors in the reported size measures can be accounted for by correcting the individual regression estimates.
- when more than one size measure is reported for an earthquake, the different regression estimates are combined into a single, more accurate estimate.

- the conversion formula is such that the ordering of the earthquakes by size is invariant with respect to the choice of magnitude scale and estimates of the parameters a and b of the exponential recurrence law are not biased. The unbiasedness property of m is valid if the historical catalog is complete; some bias may result in the case of incomplete reporting.

The corrections for measurement error, the combination of different size measures into a single scale and the correction for bias are theoretically based. The marginal distribution of the size measures is assumed exponential, whereas the distributions of the regression residuals and of the measurement errors are assumed independent Gaussian. These are common assumptions. In the derivation of the correction for measurement error and bias, the regression line is further assumed to be linear. Althought this may not be the case for the actual regression line, the proposed correction remains accurate if the regression line is well approximated locally by a straight line.

Further research on the sensitivity of the corrected lines to the modelling assumptions would be useful. Another point of interest is to study the effect of incompleteness on the regression lines more rigorously, for instance by using results on the degree of incomplete reporting as a function of the earthquake size obtained in the estimation of incompleteness and recurrence rates.

5.2 IDENTIFICATION OF CLUSTERS

Plots of historical earthquake events as points in space and time typically reveal various non-Poisson characteristics of the earthquake process. The most common phenomenon is clustering of the events, as

discussed in Chapter 2. A statistical method has been developed for the identification of such clusters, which has following features:

- In classifying earthquakes as main or dependent events, the spatial-temporal extent of the cluster region is not fixed a-priori or assumed equal for main events of the same size. Rather, the region of the clusters is estimated separately for each main event by performing statistical tests;
- Contrary to many methods in the literature, the procedure works well with spatially non-homogeneous catalogs and with catalogs that display incompleteness-induced nonstationarity . Both features are very pronounced in most earthquake catalogs.

To study the performance of the method, the procedure has been applied to two simulated catalogs and to the Chiburis (1981) catalog. For the latter catalog, the classification of events produced by the proposed method has been compared with a judgemental classification by seismologists.

The automatic procedure is found in all cases to perform quite well. Sensitivity of the results to the input parameters has been extensively studied in the case of the New England catalog. The conclusion is that the identification of clusters is robust with respect to rather substantial variations in such parameters.

The final result of the clustering procedure is the separation of the historical earthquakes into a set of independent ("Poisson") counts and a set of dependent events. These results are documented in Fig. 3.14. and are commented upon in Section 3.5. Displaying the independent events in time and space sometimes reveals non-Poission patterns other than clustering; for example, in the New England catalog,

one may notice bursts of seismic activity. These bursts (e.g., around 1860) have been noticed also by Chiburis (1981), who attributes them to increased reporting of earthquake events. In the most recent time interval, an on-and-off phenomena seems however to provide a more reasonable explanation.

The modelling of the clustered events themselves is not addressed in this thesis but is a topic of interest for future study. A difficulty in such modelling is posed by the distortion of the shape and size of the historical clusters due to incompleteness of the catalog.

5.3 ESTIMATION OF INCOMPLETENESS AND RECURRENCE RATES

Chapter 4 discusses the estimation of incompleteness and recurrence rates under the assumption that the main events follow a stationary Poisson process. Thus, nonstationarity is attributed entirely to incompleteness. Several new concepts are developed in this chapter and illustrated through four different models:

For incompleteness

- Incompleteness of the catalog is allowed to vary not only with earthquake magnitude and time but also with geographical location. This can be done by relating the probability of earthquake detection and recording to the spatial distributions of population and instruments at the time of the event, or else by specifying regions with different incompleteness characteristics.
- The notion of period of incompleteness is replaced with that of equivalent period of completeness, T_E. The latter is the period of time by which the total number of recorded events must be

divided to obtain an unbiased estimate of the recurrence rate. For a given magnitude and at a fixed geographical location, the equivalent period of completeness is obtained as the integral over time of the probability of detection for that magnitude and location. The probability of detection itself is estimated from the data, simultaneous with the recurrence rates.

- Because data and estimates of incompleteness for early periods of the catalog may be subject to large uncertainty, the analysis can be restricted to use only part of the data (within a different time interval for each magnitude). This is similar to traditional recurrence rate analysis, except that no assumption of completeness is made within the time intervals used in the analysis.

For recurrence rates

- Homogeneous earthquake sources must not be identified. Rather, seismicity parameters a and b are allowed to vary continuously on the geographical plane.
- In some cases, it is possible to identify regions with similar seismotectonic characteristics. The methods proposed allow one to use such information but does not require seismicity to be homogeneous within each region. Rather, the user can control the smoothness of the spatial variation of the recurrence relationship, separately for the a and b parameters. The standard model with homogeneous earthquake sources is obtained as a limiting special case, when total smoothness is imposed.
- Recurrence rates and incompleteness of the catalog are estimated jointly.

- Uncertainty on epicentral location or magnitude of the earthquake events can be accounted for.

For model validation

- Local significance tests that compare expected counts from the model with actual counts can be used to detect nonstationarity or non-exponentiality of the recurrence rate. Such tests can be further extended to judge the appropriateness of an assumed degree of smoothness of a and b within a given region.

For uncertainty on the estimates

- Bootstrapping techniques (empirical or parametric) are effective tools to assess uncertainty on the estimates. These methods can be used also to find uncertainty on desired quantities such as seismic hazard of a given site.

Different combinations of these new concepts have been used in Chapter 4 to formulate alternative models (A to D). From application of these models to the analysis of actual catalogs and from other considerations, the following conclusions are drawn.

- Explicitly accounting for the actual distribution of population and instruments is of importance to estimate small scale spatial variations of incompleteness. For large regions however, such a model may need to be extended to account for regional differences in the effect of population and instruments on incompleteness. It also appears that the reporting of events by people and by instruments are not independent events and the model, which now assumes independence, should be modified accordingly.
- The assumption that incompleteness is homogeneous inside given regions is a good alternative for the purpose of identifying

seismic hazard, especially when incompleteness needs to be estimated over large geographical regions. Models with homogeneously incomplete regions are easier to understand and verify. In addition, through the selection of the regions they allow one to incorporate information other than changes in population and instrument location, for example regional differences in the compilation of the catalog.

- Maximum penalized likelihood estimation of the recurrence parameters a or b is preferred to kernel estimation. The former method is a more flexible one (although it is also computationally more demanding) and allows one to combine a nonparametric specification on the spatial variation of the recurrence rates with parametric assumptions on the distribution of magnitude. Spatial variations of incompleteness are also more easily accounted for.

The methods proposed in Chapter 4 relax several questionable assumptions of traditional methods of seismicity analysis. Some assumptions, such as that of exponential recurrence rates, stationarity of the earthquake process and Poisson distribution in space and time of the main events are maintained. An interesting future development would be to further relax these assumptions and allow deviations from the stationary-Poisson-exponential model, whenever these deviations are clearly indicated by the data. Another point of interest is the spatial modelling of the recurrence rates. The degree of smoothness of the parameters can be interactively determined by comparing observed with predicted counts, for instance by using local significance tests. A possibly better technique to determine the optimal degree of smoothing is

cross-validation. The advantage of cross-validation is that the degree of smoothing is determined automatically, without the need for external intervention.

References

Agresti, A. (1984), Analysis of Ordinal Categorical Data, New York: John Wiley.

Aitchison, J. and Dunsmore, I.R. (1975), Statistical Prediction Analysis, London: Cambridge University Press.

Aki, K. (1965), "Maximum Likelihood Estimate of b in the Formula LogN=a-bM and its Confidence Limits," Bull. Earthquake Res. Inst., vol. 43, pp. 237-239.

Aki, K. (1956), "Some Problems in Statistical Seismology," Zisin, Vol. 8, No. 4, pp. 205-228 (English translation by A.S. Furimoto, Hawaii, 1963).

Andrews, D.F. (1983), "Likelihood, Shape, and Adaptive Inference." In G.E.P. Box, J. Leonard, and C.F. Wu (Eds.), Scientific Inference, Data Analysis, and Robustness. New York: Academic Press.

Barlow, R.E., Bartholomew, D.J., Brenner, J.M. and Brunk, H.D. (1972), Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression. New York: John Wiley & Sons.

Barosh. (1981), Personal Communication.

Barstow, N.L., Brill, K.G., Nuttli, O.W., and Pomeroy, P.W. (1981), "An Approach to Seismic Zonation for Siting Nuclear Electric Power Generating Facilities in the Eastern United States," NUREG/CR-1577, U.S. Nuclear Regulatory Commission, Washington, D.C.

Basu, S. (1977), Statistical Analysis of Seismic Data and Seismic Risk Analysis of Indian Peninsula", PhD. Thesis, Dept. of Civil Engineering, Indian Institute of Technology, Konpur, India.

Bath, M. (1956), "A Note on the Measure of Seismicity," Bull. Seism. Soc. Am., Vol. 46, pp. 217-218.

Bender, B. (1983), "Maximum Likelihood Estimation of b Values for Magnitude Grouped Data," Bull. Seism. Soc. Am., Vol. 73, No. 3, pp. 831-851.

Bernreuter, D.L., Savy, J.B., Mensing, R.W., and Chung, D.H. (1984), "Seismic Hazard Characterization of the Eastern United States: Methodology and Interim Results for Ten Sites", NUREG/CR-3756, USNRC, Washington, D.C.

Bishop, Y. (1975), Discrete Multivariate Analysis. MIT Press, Cambridge.

Box, G.E.P. and Cox, D.R. (1964), An Analysis of Transformations." J. Royal Statist. Soc., Series B, Vol. 26, pp. 211-252.

Brillinger, D.R. (1976), "Measuring the Association of Point Processes: A Case History," The American Mathematical Monthly, Vol. 83, pp. 16-22.

Chiburis, E.F. (1981), "Seismicity, Recurrence Rates, and Regionalization of the Northeastern United States and Adjacent Southeastern Canada," NUREG/CR-2309, USNRC, Washington, D.C.

Chung, D.H. and Bernreuter, D.L. (1980), "Regional Relationships Among Earthquake Magnitude Scales," NUREG/CR-1457.

Cleveland, W.S and McGill, R. (1984), "The Many Faces of a Scatterplot," J. Am. Stat. Assoc., Vol. 79, No. 388, Theory and Methods Section, pp. 807-822.

Cleveland, W.S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots," J. Am. Stat. Assoc., Vol. 74, No. 368, Theory and Methods Section, pp. 829-836.

Cox, D.R. and Hinckley, D.V. (1974), Theoretical Statistics, New York: Chapman and Hill.

Devroye, L. and Györfi, L. (1985), Nonparametric Density Estimation, The L_1 View, New York: John Wiley & Sons.

Draper, N.R., and Smith H. (1981), Applied Regression Analysis, New York: O. Wiley & Sons.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," The Annals of Statistics, Vol. 7, No. 1, pp. 1-26.

Efron, B. (1982), The Jackknife, the Bootstrap and Other Resampling Plans CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia.

Ellis, S.P. (1984), "An Analysis of the Haicheng Aftershock Sequence, a Nonstationary Point Process", Technical Report, No. U.S.G.S. 1, Statistics Center, M.I.T., Cambridge.

ENEA, (1984)W Catalog of Norther Italian Earthquake Data in the Friuli Region, personal communication.

Energy, Mines, and Resources Canada. (1974), The National Atlas of Canada, Ottawa.

EPRI (1985), "Seismic Hazard Methodology for Nuclear Facilities in the Eatern United States", EPRI Research Project P101-29, Palo Alto, California.

Epstein, B. and Lomnitz, C. (1966), "A Model for the Occurrence of Large Earthquakes," Nature, Vol. 211, pp. 954-956.

Fienberg, S.E. (1977), The Analysis of Cross-Classified Categorical Data, MIT Press, Cambridge.

Fienberg, S.E. and Holland, P.W. (1973), "Silmultaneous Estimation of Multinomial Cell Probabilities." J. Amer. Stat. Assoc., Vol. 68, pp. 683-691.

Fienberg, S.E. and Holland, P.W. (1969), "Methods for Eliminating Zero Counts in Contingency Tables." In Patil, G.P. (ed.), Random Counts in Scientific Work, Vol. 1, University Park, Pennsylvania: The Pennsylvania State University Press.

Fraser, D.A.S. (1976), "Necessary Analysis and Adaptive Inference." J. Amer. Statist. Assoc., Vol. 71, pp. 99-113.

Friis, H.R. (1960), "A Series of Population Maps of the Colonies and the United States 1625-1790." Mimeographed Publication No. 3, American Geographical Society, New York.

Ganse, R.A., Amemiya, Y. and Fuller, W.A. (1983), "Prediction When Both Variables Are Subject to Error, With Application to Earthquake Magnitudes," J. Am. Stat. Assoc., Vol. 78, No. 384, Applications Section, pp. 761-765.

Gardner, J.K. and Knopoff, L. (1974), "Is the Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian?," Bulletin of the Seismological Society of America, Vol. 64, No. 5, pp. 1363-1367.

Good, I.J. and Gaskins, R. (1971), "Nonparametric Roughness Penalties for Probability Densities." Biometrika, vol. 58, pp. 255-277.

Good, I.J. and Gaskins, R. (1980), "Density Estimation and Bump-hunting by Penalized Likelihood Method Exemplified by Scattering and Meteorite Data." J. Amer. Stat. Assoc., vol. 75, pp. 42-74.

Groeneboom, P. and Pyke R. (1983), "Asymptotic Normality of Statistics Based on the Convex Minorants of Empirical Distribution Functions," The Annals of Probability, Vol. 11, No. 2, pp. 328-345.

Haberman, S.J. (1973), "Log-Linear Models for Frequency Data: Sufficient Statistics and Likelihood Equations", The Annals of Statistics, Vol. 1, No. 4, pp. 617-632.

Halliday, R.J., Shannon, W.E., Lombardo, R., and Compton, B. (1981), "Canadian Seismograph Operations - 1980." Seismological Series, Earth Physics Branch, No. 86, Ottawa.

Halliday, R.J., Shannon, W.E., Lombardo, R., and Compton, B. (1977), "Canadian Seismograph Operations - 1975." Seismological Series, Earth Physics Branch, No. 75, Ottawa.

Hunter, W.G. and Lamboy, W.F. (1981), "A Bayesian Analysis of the Linear Calibration Problem," Technometrics, Vol. 23, No. 4, pp. 323-349.

Iacurto, O., Paciello, A., Musmeci, F., and Basili, M. (1981), "On the Completeness Analysis of Historical Catalogues," Proceedings, Annual Meeting of Progetto Finalizzato Geodinamica, Udine, Italy, May 12-14.

Johnson, N.L. and Kotz, S. (1970), Continuous Univariate Distributions-I, Distributions in Statistics, New York: J. Wiley & Sons.

Kagan, Y.Y. (1981), "Spatial Distribution of Earthquakes: the Three-Point Moment Function," Geophys. J.R. Astr. Soc., Vol. 67, pp. 697-717.

Kagan, Y.Y. and Knopoff, L. (1976), "Statistical Search for Non-random Features of the Seismicity of Strong Earthquakes," Physics of the Earth and Planetary Interiors, Vol. 12, pp. 291-318.

Kagan, Y.Y. and Knopoff, L. (1978), "Statistical Study of the Occurrence of Shallow Earthquakes," Geophys. J.R. Astr. Soc., Vol. 55, pp. 67-86.

Kagan, Y.Y. and Knopoff, L. (1980), "Spatial Distribution of Earthquakes: the Two-Point Correlation Function," Geophys. J.R. Astr. Soc., Vol. 62, pp. 303-320.

Kagan, Y.Y. and Knopoff, L. (1981), "Stochastic Synthesis of Earthquake Catalogs," Journal of Geophysical Research, Vol. 86, No. 4, pp. 2853-2862.

Kelly, E.J. and Lacoss, R.T. (1969), "Statistical Estimation of Seismicity and Dection Probability," Semiannual Technical Summary, Seismic Discrimination, Lincoln Laboratory, M.I.T.

Kendall, M.G. and Stuart, A. (1967), The Advanced Theory of Statistics. Vol. 2, Second Edition, New York: Hafner Publishing.

Ketellapper, R.H. (1983), "On Estimating Parameters in a Simple Linear Errors-in-Variables Model", Technometrics, Vol. 25, No. 1, pp. 43-47.

Knopoff, L., Kagan, Y.Y. and Knopoff, R. (1982), "b-Values for Foreshocks and Aftershocks in Real and Simulated Earthquake Sequences," Bulletin of the Seismological Society of America, Vol. 72, No. 5, pp. 1663-1676.

Kulldorff, G. (1961), Estimation from Grouped and Partially Grouped Samples, New York: John Wiley & Sons.

Laird, N.M. (1978), "Empirical Bayes Methods for Two-way Contigency Tables. Biometrika, Vol. 65, pp. 581-590.

Lee, W.H.K. and Brillinger, D.R. (1979), "On Chinese Earthquake History -An Attempt to Model an Incomplete Data Set by Point Process Analysis," Pageoph., Vol. 117, pp. 1229-1257.

Lehmann, E.E. (1959), Testing Statistical Hypotheses, p. 140, New York: John Wiley and Sons.

Leonard, T. (1973), "A Bayesian Method for Histograms." Biometrika, Vol. 60, pp. 297-308.

Leonard, T. (1975), "Bayesian Estimation Models for Two-way Tables." J. Royal Stat. Soc., Series B, Vol. 37, pp. 23-37.

Leonard, T. (1978), "Density Estimation, Stochastic Processes, and Prior Information." J. Royal Stat. Soc., Series B, Vol. 40, pp. 113-146.

Levin, T. and Maritz, J.S. (1982), "An Analysis of the Linear-Calibration Controversy From the Perspective of Compound Estimation, Technometrics, Vol. 24, No. 3, pp. 235-242.

Lindley, D.V. (1947), Regression Lines and the Linear Functional Relationships." J. Royal Statist. Soc., (Suppl.), Vol. 9, pp. 218-244.

Little, R.J.A. (1982), "Models for Non-Response in Sample Surveys," Journal of the American Statistical Association, Vol. 77, pp. 237-250.

Lomnitz-Adler, J. and Lomnitz, C. (1979), "A Modified Form of the Gutenberg-Richter Magnitude-Frequency Relation," Bull. Seism. Soc. Am., Vol. 69, pp. 1209-1214.

Lord, C.L. and Lord, E.H. (1953), Historical Atlas of the United States. Revised Edition, New York: Holt.

Mandansky, A. (1959), "The Fitting of Straight Lines When Both Variables Are Subject to Error," J. Am. Stat. Assoc., Vol. 54, pp. 173-205.

Matérn, B. (1960), Spatial Variation, Comm. Swed. Forestry Res. Inst., 49, pp. 1-144.

McGuire, R.K. (1977), "Effects of Uncertainty in Seismicity on Estimates of Seismic Hazard on the East Coast of the United States," Bull. Seism. Soc. Am., Vol. 67, pp. 827-838.

Merz, H.A. and Cornell, C.A. (1973), "Aftershocks in Engineering Seismic Risk Analysis," Research Report R73-25, Dept. of Civil Engineering, M.I.T., Cambridge, Massachusetts.

Merz, H.A. and Cornell, C.A. (1973b), "Seismic Risk Analysis Based on a Quadratic Magnitude-Frequency Law," Bull. Seism. Soc. Am., Vol. 63, No. 6, pp. 1999-2006.

Montgomery, D.C. and Peck, E.A. (1982), Introduction to Linear Regression Analysis. New York: John Wiley & Sons.

Nordheim, E.V. (1984), "Inference from Nonrandomly Missing Categorical Data: An Example From a Genetic Study on Turner's Syndrome," Journal of the American Statistical Association, Vol. 79, No. 388, pp. 772-780.

Nuttli, O.W. (1974), "Magnitude Recurrence Relation for Central Mississippi Valley Earthquakes." Bull. Seism. Soc. Am., Vol. 64, No. 4, pp. 1189-1207.

Page, R. (1968), "Aftershocks and Microaftershocks of the Great Alaska Earthquake of 1964", Bull. Seism. Soc. Am., Vol. 58, pp. 1131-1168.

Parzen, E. (1962), "On the Estimation Of a Probability Density Function and the Mode." Annals of Mathematical Statistics, Vol. 33, pp. 1065-1076.

Plante, A. (1970), "Counter-Examples and Likelihood," Proceedings, Symposium on the Foundations of Statistical Inference, University of Waterloo, Ontario, Canada.

Poppe, B.B. (1979), "Historical Survey of U.S. Seismograph Stations." U.S. Geological Survey, Professional Paper 1096, U.S. Dept. of Interior, Washington.

Prozorov, A.G. and Dziewonski, A.M. (1982), "A Method of Studying Variations in the Clustering Property of Earthquakes: Application to the Analysis of Global Seismicity," Journal of Geophysical Research, Vol. 87, No. B4, pp. 2829-2839.

Pregibon, D. (1977), "Typical Survey Data: Estimation and Imputation," Survey Methodology, Vol. 2, pp. 70-102.

Press, S.J. (1968), "Estimating from Misclassified Data," Journal of the American Statistical Association, Vol. 63, pp. 123-133.

Reasenberg, P. (1984), "Second-Order Moment of Central California Seismicity, 1969-1982," U.S. Geological Survey, Menlo Park.

Reilly, P.M. and Patino-Leal, H.V. (1981), "A Bayesian Study of the Error-in-Variables Model," Technometrics, Vol. 23, No.3, pp. 221-231.

Richter, C.F. (1958), Elementary Seismology, San Francisco: W.H. Freeman and Co.

Rosenblatt, M. (1956), "Remarks on Some Nonparametric Estimates of a Density Function." Annals of Mathematical Statistics, Vol. 27, pp. 832-837.

Savage, W.U. (1975), "Earthquake Probability Models: Recurrence Curves, Aftershocks, and Clusters", Ph.D. Thesis, University of Nevada, Reno.

Shlien, S. and Toksöz, M.N. (1975), "A Branching Poisson-Markov Model of Earthquake Occurrences," Geophys. J.R. Astr. Soc., Vol. 42, pp. 49-59.

Shlien, S. and Toksöz, M.N. (1970), "A Clustering Model for Earthquake Occurrences," Bulletin of the Seismological Society of America, Vol. 60, No. 6, pp. 1765-1787.

Shlien, S. and Toksöz, M.N. (1970b), "Frequency-Magnitude Statistics of Earthquake Occurrence," Earthquake Notes (Eastern Section of the Seismological Society of America), Vol. 41, pp. 5-18.

Simonoff, J.S. (1983), "A Penalty Function Approach to Smoothing Large Sparse Contingency Tables." The Annals of Statistics, vol. 11, pp. 208-218.

Simpson, D.W. and Richards, P.G. (Eds.) (1981), Earthquake Prediction: An International Review, American Geophysical Union, Washington, D.C.

Stepp, J.C. (1972), "Analysis of Completeness of the Earthquake Sample in the Puget Sound Area and Its Effect on Statistical Estimates of Earthquake Hazard," Proceedings, International Conference on Microzonation, vol. 2, pp. 897-910.

Stepp, J.C., Rinehart, W.A., Algermissen, S.T. (1965), "Earthquakes in the United States 1963-64 and an Evaluation of the Detection Capability of the United States Seismograph Stations." ARPA Order No. 620, U.S. Dept. of Commerce, Environmental Science Services Administration, Coast and Geodetic Survey, Washington, D.C.

Stevens, A.E. (1980), "History of Some Canadian and Adjacent American Seismograph Stations." Bull. Seism. Soc. Am., vol. 70, pp. 1381-1393.

Street, R.L. and Turcotte, F.T. (1977), "A Study of Northeastern North American Spectral Moments, Magnitudes, and Intensities." Bull. Seism. Soc. Am., Vol. 67, No. 3, pp. 599-614.

Tapia, R.A. and Thompson, J.R. (1978), Nonparametric Probability Density Estimation. Baltimore: The John Hopkins Univesity Press.

Utsu, T. (1969,1970,1971), "Aftershocks and Earthquake Statistics, Parts I-III," Journal of the Faculty of Science, Hokkaido University, Ser. VII, Vol. III, No. 3 (1969), No. 4 (1970), and No. 5 (1971).

Utsu, T. (1966), "A Statistical Significance Test of the Difference in b-Value Between Two Earthquake Groups", J. Phys. Earth, 14, pp. 37-40.

Utsu, T. (1961), "A Statistical Study of the Occurrence of Aftershocks," Geophysical Magazine, Tokyo, Vol. 30, pp. 521-605.

Veneziano, D. and Van Dyck, J. (1985a), "Seismic Hazard Analysis for the Friuli Region, Part I: Magnitude Conversion and Earthquake Clustering," Report to ENEA, Rome, Italy.

Veneziano, D. and Van Dyck, J. (1985b), "Seismic Hazard for the Friuli Region: Part II: Catalog Incompleteness, Seismic Sources, and Recurrence Rates", Report to ENEA, Rome, Italy.

Vere-Jones, D. (1978), "Space Time Correlations for Microearthquakes -a Pilot Study," Proceedings, Conf. on Spatial Patterns and Processes, Advances in Applied Probability, Vol. 10 (Special Suppl.), pp. 73-87.

Vere-Jones, D. (1970), "Stochastic Models of Earthquake Occurrence," Journal of the Royal Statistical Society, Vol. 32, No. 1, pp. 1-62.

Vere-Jones, D., Turnovsky, S., Eiby, G.A., and Davis, R.B. (1964,1965), "A Statistical Survey of Earthquakes in the Main Seismic Regions of New Zealand, Parts I and II," New Zealand Journal of Geology and Geophysics, Vol. 7 (1964), pp. 722-744, Vol. 9 (1965), pp. 251-284.

Weichert, D.H. (1980), "Estimation of the Earthquake Recurrence Parameters for Unequal Observation Periods for Different Magnitudes," Bull. Seism. Soc. Am., Vol. 70, pp. 1337-1346.

Weston Geophysical Corporation. (1983), Appendix C of "Maine Yankee Seismic Hazard Analysis." YAEC-1356, Framingham, Massachusetts.

Weston Geophysical Corporation. (1982), "Supplementary Seismic Probabilistic Study, Yankee Atomic Electric Company, Rowe, Massachusetts." YAE-1331, Appendix D1.

Appendix

ROBUST LOCALLY WEIGHTED REGRESSION

Robust locally weighted regression is a non-parametric regression method, originally proposed by Cleveland (1979). For application to magnitude conversion, a local estimate of uncertainty about the regression is necessary to account for heteroscedasticity. Because such an estimate is not derived in Cleveland's paper, the method is reviewed here in more detail. Robust locally weighted regression is a technique designed to analyze data for which the regression of y on x is a smoothly varying function:

$$y_i = g(x_i) + \varepsilon_i \qquad (A.1)$$

Subscript i indicates the i'th point in the sample ordered for increasing x. The total number of points in the sample is n.

In locally weighted regression, estimates of y_i are obtained by fitting locally at x_i a straight line:

$$\hat{y} = \beta_i^0 + \beta_i^1 x \qquad (A.2)$$

One might consider fitting a polynomial of any order, but in practice, a straight line is often sufficient. Denote by $w_{i,k}$ the weight given to the k'th datapoint when estimating the linear regression at x_1. Total parameters β_i^0 and β_i^1 are found by minimizing the weighted sum of squares

$$SS_i = \sum_{k=1}^{n} w_{i,k}\left(\beta_i^0 + \beta_i^1 x_k - y_k\right)^2 \qquad (A.3)$$

The following matrix notation is useful:

$$\underline{\beta}_i = [\beta_i^0, \beta_i^1]^T \text{ , with dimension 2 x 1} \quad \text{(A.4)}$$

$$(\underline{X})_{k,} = [1 \; x_k] \text{ , with dimension n x 2} \quad \text{(A.5)}$$

$$(\underline{Y})_{k,} = [y_k] \text{ , with dimension n x 1} \quad \text{(A.6)}$$

$$(\underline{W}_i)_{j,k} = \begin{cases} w_{i,k} & j=k \\ 0 & j \neq k \end{cases} \text{ , with dimension n x n} \quad \text{(A.7)}$$

The sum-of-squares in Equation A.3 is the same as in weighted least squares (Draper and Smith, 1981). The associated estimators of $\underline{\beta}_i$ and γ_i are:

$$\hat{\underline{\beta}}_i = (\underline{X}^T \underline{W}_i \underline{X}_i)^{-1} \underline{X}^T \underline{W}_i \underline{Y} \quad \text{(A.8)}$$

and

$$\hat{y}_i = [1 \; x_i] \hat{\underline{\beta}}_i \quad \text{(A.9)}$$

From Equations A.8 and A.9 it follows that $\hat{y}_i$ is a linear combination of the observed y values and can be written as:

$$\hat{y}_i = \sum_{k=1}^{n} r_{i,k} \, y_k \quad \text{(A.10)}$$

In general,

$$\hat{\underline{Y}} = \underline{R} \, \underline{y} \quad \text{(A.11)}$$

where

$$(\hat{\underline{y}})_i = \hat{y}_i$$

$$(\underline{R})_{i,k} = r_{i,k}$$

If the regression is homoscedastic with residual variance σ^2, then an unbiased estimate of σ^2 is:

$$\sigma^2 = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 \mid \text{trace } (\underline{C}) \tag{A.12}$$

where $\hat{\varepsilon}_i$ is the residual

$$\hat{\varepsilon}_i = \hat{y}_i - y_i,$$

and $\underline{C}$ is the covariance matrix of $[\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n]$, given by

$$\underline{C} = (\underline{I} - \underline{R})(\underline{I} - \underline{R})^T \tag{A.13}$$

$\underline{I}$ is the n x n identity matrix.

If the regression is heteroscedastic, a local estimate of σ^2 is needed. One such estimate is obtained by assuming that the local residuals $\hat{\varepsilon}_{i,k}$ used in Equation A.3,

$$\hat{\varepsilon}_{i,k} = y_k - \beta_i^0 - \beta_i^1 x_k, \tag{A.14}$$

are homoscedastic with local variance σ_i^2 and have mean value equal to zero. If $\underline{\hat{\varepsilon}_i}$ is the column-vector with elements $\hat{\varepsilon}_{i,k}$, then

$$\underline{\hat{\varepsilon}_i} = \underline{Y} - \underline{X}\,\underline{\hat{\beta}_i} \tag{A.15}$$

Using Equation A.8, this is expanded to:

$$\underline{\hat{\varepsilon}_i} = [\underline{I} - \underline{M}]\,\underline{Y} \tag{A.16}$$

where:

$$\underline{M} = \underline{X}(\underline{X}^T\underline{W}_i\underline{X})^{-1}\,\underline{X}^T\underline{W}_i \tag{A.17}$$

Notice that the matrix $\underline{M}$ is symmetric and idempotent, so that $\underline{M}\,\underline{M} = \underline{M}$. Therefore, the covariance matrix of $\hat{\underline{\varepsilon}}_i$ in Equation A.16

$$T = \sigma_i^2(\underline{I} - \underline{M})(\underline{I} - \underline{M})^T$$

$$= \sigma_i^2(\underline{I} - \underline{M}) \qquad (A.18)$$

The local sum of squares in Equation A.3 can also be written as:

$$SS_i = \text{trace}(\underline{W}_i\ \hat{\underline{\varepsilon}}_i\ \hat{\underline{\varepsilon}}_i^T), \qquad (A.19)$$

and the expected value of SS_i is, using A.18,

$$E[SS_i] = \sigma_i^2[\text{trace}(\underline{W}_i) - \text{trace}(\underline{W}_i\underline{M})] \qquad (A.20)$$

Using the property that trace $(\underline{S}\ \underline{T})$ = trace$(\underline{T}\ \underline{S})$, it follows from Equation A.17, that

$$E[SS_i] = \sigma_i^2[\text{trace}(\underline{W}_i) - \text{trace}(\underline{A}_1^{-1}\underline{A}_2)] \qquad (A.21)$$

where:

A_1 and A_2 are defined as for s=1 and 2 respectively.

$$\underline{A}_s = \underline{X}^T\ \underline{W}_i^s\ \underline{X}$$

So far, no attention has been given to the presence of outliers and the method as described above is referred to as locally weighted regression. In a robustified version of the method, this procedure is applied first to find initial estimates, say $\hat{y}_i^{(0)}$. Subsequent iterations use modified weights. In the j'th iteration, the weights are:

$$w_{i,k}^{(j)} = w_{i,k}\ \delta\left(\hat{\varepsilon}_i^{(j-1)}\right) \qquad (A.22)$$

where δ is a function that decreases with increasing absolute value of the of the residual. For example, Cleveland uses for δ a bisquare function with $\hat{\varepsilon}_i^{(j-1)}$ scaled to the median value of all residuals at iteration j-1.

To account for heteroscedasticity, it is preferable to normalize first $\hat{\varepsilon}_i$ with respect to $\hat{\sigma}_i$,

$$\hat{\varepsilon}'_i = \hat{\varepsilon}_i / \hat{\sigma}_i \tag{A.23}$$

Then, a bisquare function is used with argument $\hat{\varepsilon}'_i$:

$$\delta(\hat{\varepsilon}'_i) = \begin{cases} \left[1 - \left(\dfrac{\hat{\varepsilon}'_i}{bm_{\hat{\varepsilon}'_i}}\right)^2\right]^2 & \text{for } \varepsilon \leq 6m_\varepsilon \\ 0 & \text{for } \varepsilon > 6m_\varepsilon \end{cases} \tag{A.24}$$

where $m_{\hat{\varepsilon}'_i}$ is the median value of all normalized residuals $\hat{\varepsilon}'_i$ - for a given iteration. To complete the details of the procedure, the weights $w_{i,k}$ used in Equation A.22 need to be specified. Cleveland uses a trisquare function of the distance of point k to point i, normalized to the r'th nearest-neighbor distance for point i. Such a choice has the advantage of automatically modulating the width of the local window according to the density of the points. A disadvantage of this technique is that the window size is always very large for high values of the size measure, because the earthquake count is small. As a simple alternative, a fixed local window is proposed here, with weighting function:

$$w_{i,k} = \begin{cases} \exp\left\{-\dfrac{1}{2}\dfrac{(x_k - x_i)^2}{h^2}\right\} & |x_k - x_i| \leq 4h \\ 0 & |x_k - x_i| > 4h \end{cases} \tag{A.25}$$

given	ξ	η	x	y
ξ	b_ξ	$\beta_0 + \beta_1 \xi$ σ_e^2	ξ σ_u^2	$\beta_0 + \beta_1 \xi$ $\sigma_e^2 + \sigma_v^2$
η	$\frac{1}{\beta_1}(\eta - \beta_0 - \frac{b_\xi}{\beta_1}\sigma_e^2)$ $\frac{\sigma_e^2}{\beta_1^2}$	$\frac{b_\xi}{\beta_1}$	$\frac{1}{\beta_1}(\eta - \beta_0 - \frac{b_\xi}{\beta_1}\sigma_e^2)$ $\frac{\sigma_e^2 + \sigma_v^2}{\beta_1^2}$	η σ_v^2
x	$x - b_\xi \sigma_u^2$ σ_u^2	$\beta_0 + \beta_1 x - \beta_1 b_\xi \sigma_u^2$ $\sigma_e^2 + \beta_1^2 \sigma_u^2$	b_ξ	$\beta_0 + \beta_1 x - \beta_1 b_\xi \sigma_u^2$ $\sigma_e^2 + \sigma_v^2 + \beta_1^2 \sigma_u^2$
y	$\frac{1}{\beta_1}[y - \beta_0 - \frac{b_\xi}{\beta_1}(\sigma_e^2 + \sigma_v^2)]$ $\frac{\sigma_e^2 + \sigma_v^2}{\beta_1^2}$	$y - \frac{b_\xi}{\beta_1}\sigma_v^2$ σ_v^2	$\frac{1}{\beta_1}[y - \beta_0 - \frac{b_\xi}{\beta_1}(\sigma_e^2 + \sigma_v^2)]$ $\frac{\sigma_e^2 + \sigma_v^2}{\beta_1^2} + \sigma_u^2$	$\frac{b_\xi}{\beta_1}$

Note : diagonal elements correspond to the slope parameter of the marginal exponential distribution
off-diagonal elements refer to the mean and variance of the conditional Gaussian distributions

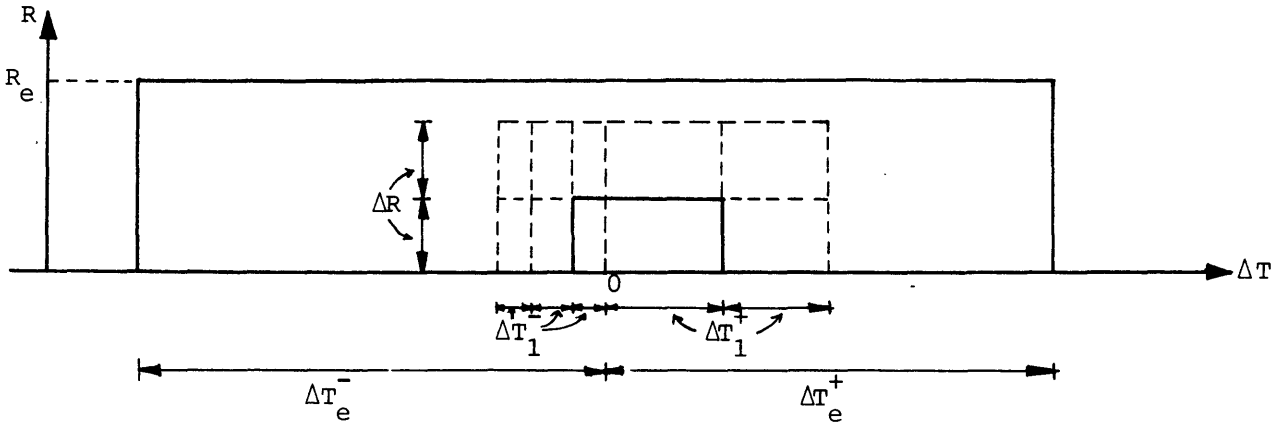
Table 2.1 - Summary of parameters for the marginally-exponential, conditionally-normal distribution

Magnitude	Radius (km)	Duration (days),
2.5	19.5	6
3.0	22.5	11.5
3.5	26.0	22
4.0	30.0	42
4.5	35.0	83
5.0	40.0	155
5.5	47.0	290
6.0	54.0	510
6.5	61.0	790
7.0	70.0	915
7.5	81.0	960
8.0	94.0	985

Table 3.1 - Dimensions of the space-time windows used by Gardner and Knopoff (1974) in the analysis of Southern California earthquake data

I_o	TIME OF OCCURRENCE					
	1534-1699	1700-1799	1800-1849	1850-1899	1900-1949	1950-1980
1	200	100	100	50	50	25
2	200	100	100	50	50	25
3	200	100	100	50	50	25
4	200	100	100	50	50	50
5	200	200	100	50	50	50
6	200	200	100	100	50	50
7	300	300	200	200	100	100
8	500	500	500	500	500	500
9	1000	1000	1000	1000	1000	1000
10	1000	1000	1000	1000	1000	1000

Table 3.2 - Intervals of randomization in years around the historical earthquake times used in the generation of the nonhomogeneous, nonstationary, quasi-Poisson catalog.



I_o	ΔR(degrees)	ΔT_1^-(days)	ΔT_1^+(days)	R_e(degrees)	n_T^- (1)	n_T^+ (2)	n_R (3)
1	0.20	5	10	1.00	4	4	2
2	0.20	30	40	1.00	4	4	2
3	0.20	50	100	1.00	4	4	2
4	0.22	60	200	1.00	4	4	2
5	0.28	70	300	1.00	4	4	2
6	0.30	80	400	1.00	4	4	2
7	0.32	90	500	1.00	4	4	2
8	0.35	100	500	1.00	4	4	2
9	0.38	110	500	1.00	4	4	2
10	0.40	120	500	1.00	4	4	2

(1) max number of backward extensions in time
(2) max number of foreward extensions in time
(3) max number of extensions in space

Time Interval	ΔT_e^-(years)	ΔT_e^+(years)
1534 - 1649	150	75
1650 - 1749	100	75
1750 - 1849	100	50
1850 - 1949	50	30
1950 - 1969	20	10
1970 - 1980	10	5

$q = 0.1$
$\alpha = 0.02$
$\alpha_{ext} = 0.02$
No. of iterations = 2
Method of earthquake classification: Method 2 of Sec. 3.3.5

Table 3.3 - Input parameters for the analysis of the stationary Poisson catalog.

I_o	EQKS.	MAIN	SECONDARY	CLUSTERS
1	245	236	9	0
2	659	644	15	1
3	761	742	19	7
4	648	628	20	25
5	339	336	3	11
6	133	132	1	3
7	57	57	0	1
8	12	12	0	2
9	4	4	0	2
10	2	2	0	0
TOTAL	2860	2793	67	52

Table 3.4 - Summary results for the stationary Poisson catalog.

I_o	EQKS.	MAIN	SECONDARY	CLUSTERS
1	245	200	45	3
2	659	583	76	14
3	761	698	63	36
4	648	615	33	29
5	339	326	13	22
6	133	133	0	5
7	57	57	0	2
8	12	12	0	0
9	4	4	0	0
10	2	2	0	0
TOTAL	2860	2630	230	111

Table 3.5 - Summary results for the simulated nonstationary catalog.

I_o	ΔR(degrees)	ΔT_l^-(days)	ΔT_l^+(days)	R_e(days)	n_T^-	n_T^+	n_R
1	0.20	5	10	1.00	3	4	2
2	0.20	30	40	1.00	3	4	2
3	0.20	50	100	1.00	3	4	3
4	0.22	60	200	1.00	3	6	3
5	0.28	70	300	1.00	3	6	3
6	0.30	80	400	1.00	3	6	3
7	0.32	90	500	1.00	3	8	3
8	0.35	100	500	1.00	3	6	2
9	0.38	110	500	1.00	3	4	2
10	0.40	120	500	1.00	3	4	2

Time Interval	ΔT_e^-(years)	ΔT_e^+(years)
1534 - 1649	150	75
1650 - 1749	100	75
1750 - 1849	100	50
1850 - 1949	50	30
1950 - 1969	20	10
1970 - 1980	10	5

$q = 0.1$
$\alpha = 0.02$
$\alpha_{ext} = 0.02$
no. of iterations = 2
Method of earthquake classification: Method 2 of Sec. 3.3.5

Table 3.6 - Input parameters for the analysis of the Weston Observatory Catalog (see Table 3.2 for explanation of symbols).

		This Study			Secondary Events		
I_o	EQKS.	MAIN	SECONDARY %	CLUSTERS	THIS STUDY ONLY	SEISMOL. ONLY	BOTH
1	245	155	90(37)	8	46	8	44
2	659	422	237(36)	26	95	11	142
3	761	532	229(30)	53	83	12	146
4	648	472	176(27)	86	62	14	114
5	339	296	43(13)	39	24	4	19
6	133	112	21(16)	21	7	1	14
7	57	51	6(11)	12	2	2	4
8	12	9	3(25)	4	1	0	2
9	4	4	0(0)	2	0	0	0
10	2	2	0(0)	2	0	0	0
TOTAL	2860	2055	805(28)	253	320	51	485

Table 3.7 - Summary results for the Weston Observatory catalog.

		SECONDARY	BREAKDOWN BY INTENSITY									
I_o	NO. CLUSTERS	EVENTS IN CLUSTERS	I_o=1	2	3	4	5	6	7	8	9	10
1	8	8	8									
2	25	46	23	23								
3	52	98	22	25	51							
4	86	168	7	62	59	40						
5	40	172	26	59	38	30	19					
6	22	91	1	26	19	33	9	3				
7	12	149	1	29	44	55	5	12	3			
8	4	40	2	12	13	6	5	2	0	0		
9	2	22	0	1	5	9	4	2	1	0	0	
10	2	11	0	0	0	3	1	2	2	3	0	0

Table 3.8 - Breakdown by intensity of secondary events in clusters.

I_O	NO. CLUSTERS	NO. TIME COMPRESSIONS	NO. TIME EXTENSIONS	NO. SPACE EXTENSIONS	n_T^- (1)	n_T^+ (2)	n_R (3)
1	8	0	1	1	0	0	0
2	25	1	1	2	1	1	1
3	52	5	4	4	1	2	1
4	86	10	8	6	2	3	1
5	40	14	8	4	2	4	2
6	22	5	6	3	1	3	1
7	12	2	5	0	2	7	0
8	4	0	1	1	1	3	0
9	2	0	0	0	0	0	0
10	2	0	1	0	0	1	0
TOTAL	253	37	35	21			

(1) Maximum number of extensions backward in time in a single cluster
(2) Maximum number of extensions foreward in time in a single cluster
(3) Maximum number of extensions in space in a single cluster

Table 3.9 - Cluster size statistics.

Sensitivity Case	Change in the parameters of Table 3.6
1	$\alpha = \alpha_{EXT} = 0.05$
2	$\alpha_{EXT} = 0.05$
3	ΔR doubled and n_R set to 1 for all I_O
4	ΔR halved and n_R doubled for all I_O
5	$R_e = 1.50$ for all I_O
6	ΔT_e^- and ΔT_e^+ doubled for each time interval
7	Total removal of secondary events inside cluster regions
8	Removal of secondary events by Method 1 in Sec. 3.3.4

Table 3.10 - Variants of Table 3.6 for sensitivity analysis.

I_o	BASE CASE	SENSITIVITY 1 α=0.05 α_{EXT}=0.05	2 α=0.02 α_{EXT}=0.05	3 ΔR doubled	4 ΔR halved	5 R_e=1.50	6 ΔT_e^-, ΔT_e^+ doubled	7 total eqk. removal	8 removal by catalog simulation
1	8	8	8	8	4	8	8	8	8
2	25	25	23	35	26	26	25	25	31
3	52	50	49	59	48	50	52	52	53
4	86	87	83	81	79	85	81	83	85
5	40	45	43	41	37	40	38	41	39
6	22	23	22	24	14	20	19	22	22
7	12	14	12	11	12	13	12	14	12
8	4	5	5	4	3	5	4	4	5
9	2	2	2	2	2	2	2	2	2
10	2	2	2	2	2	2	2	2	2
TOT-AL	253	261	249	267	227	251	243	253	257
%	12%	13%	12%	14%	10%	12%	12%	12%	12%

Table 3.11 - Number of clusters in base case and sensitivity cases.

I_o	BASE CASE	SENSITIVITY 1 α=0.05 α_{EXT}=0.05	2 α=0.02 α_{EXT}=0.05	3 ΔR doubled	4 ΔR halved	5 R_e=1.50	6 $\Delta T_e^-, \Delta T_e^+$ doubled	7 total eqk. removal	8 removal by catalog simulation
1	90	99	101	119	76	103	101	93	83
2	237	257	247	274	206	254	239	241	213
3	229	252	239	253	202	241	232	232	217
4	176	189	184	186	155	177	175	175	163
5	43	45	45	44	29	43	43	42	39
6	21	21	21	23	18	21	21	21	21
7	6	7	7	6	5	6	6	7	6
8	3	3	3	3	2	3	3	3	3
9	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0
TOT-AL	805	873	847	908	693	848	820	814	745
%	28%	31%	30%	32%	24%	30%	29%	28%	26%

Table 3.12 - Number of secondary events in base case and sensitivity cases.

Time period	1.	2.	3.	Aftershocks		% Aftershocks		present analysis	
From - To	All events	Events in clusters	Main events	4. WGC	present analysis	WGC	present analysis	No. of clusters	% of clusters
a. 1500 - 1800	204	149	73	114	131	56	64	18	25
b. 1800 - 1900	477	146	381	65	96	14	20	50	13
c. 1900 - 1940	488	185	340	103	148	21	30	37	11
d. 1940 - 1960	430	166	306	96	124	22	29	42	14
e. 1960 - 1974	401	95	334	55	67	14	17	40	12
f. 1974 - 1981	648	281	430	97	218	15	34	63	15
TOTAL	2648	1022	1864	530	784	20	28	238	13

Note : % of clusters is calculated relative to the number of main events

Table 3.13 - Number of events plotted in Figures 3.14a-3.14f

Category	Population density (inhabitants per square mile)
0	<2
1	2-5
2	6-17
3	18-44
4	45-81
5	≥90

Table 4.1 - Population categories

Category	Distance to closest station (kilometers)
0	≥305
1	195-304
2	110-194
3	55-109
4	0-54

Table 4.2 - Instrument categories

i_L	Maximum radius of uncertainty (km)
1	5
2	10
3	20
4	50
5	> or >> 50
6	instrumental estimate ($\equiv i_L=1$)

Table 4.3 - Maximum radius of uncertainty on epicentral location for various epicentral-accuracy classes

I_0	**average intensity**
dI_0	**difference of reported values of I_0**
source	see Figure 4.10f
UL	$\equiv i_L$, location uncertainty (Table 4.3)
time	1 : 1000 - 1249 2 : 1250 - 1499 3 : 1500 - 1699 4 : 1700 - 1873 5 : 1874 - 1984

	source		
I0	1	2	3
1	0	0	0
2	0	21	10
3	2	92	80
4	0	86	42
5	2	71	40
6	0	36	16
7	0	24	11
8	0	6	3
9	0	7	0
10	0	1	0
11	0	0	0

Source 1

	Time				
I0	1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	2	0
4	0	0	0	0	0
5	0	0	0	0	2
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	0	0	0	0	0
10	0	0	0	0	0
11	0	0	0	0	0

	UL					
I0	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	2	0	0
4	0	0	0	0	0	0
5	0	0	2	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0

	dIO					
I0	0	1	2	3	4	5
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	1	1	0	0
4	0	0	0	0	0	0
5	2	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0

	Time				
UL	1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	2
4	0	0	0	2	0
5	0	0	0	0	0
6	0	0	0	0	0

	dIO					
UL	0	1	2	3	4	5
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	2	0	0	0	0	0
4	0	0	1	1	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

	Time				
dIO	1	2	3	4	5
0	0	0	0	0	2
1	0	0	0	0	0
2	0	0	0	1	0
3	0	0	0	1	0
4	0	0	0	0	0
5	0	0	0	0	0

Table 4.4 - Earthquake counts in the Friuli region

Source 2

IO	Time 1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	5	16
3	0	0	1	32	59
4	0	3	5	11	67
5	0	3	3	18	47
6	1	0	1	2	32
7	0	4	1	7	12
8	1	0	0	3	2
9	0	0	2	3	2
10	0	0	0	0	1
11	0	0	0	0	0

IO	UL 1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	1	15	0	5	0
3	0	1	56	6	29	0
4	1	2	66	4	13	0
5	5	2	51	4	9	0
6	0	0	33	2	1	0
7	1	4	13	2	4	0
8	2	1	1	2	0	0
9	0	2	5	0	0	0
10	1	0	0	0	0	0
11	0	0	0	0	0	0

IO	dIO 0	1	2	3	4	5
1	0	0	0	0	0	0
2	13	8	0	0	0	0
3	63	11	0	18	0	0
4	64	11	7	1	2	1
5	55	13	1	0	0	2
6	28	8	0	0	0	0
7	20	2	2	0	0	0
8	5	1	0	0	0	0
9	3	4	0	0	0	0
10	0	1	0	0	0	0
11	0	0	0	0	0	0

UL	Time 1	2	3	4	5
1	0	0	0	3	7
2	0	0	0	8	5
3	0	2	5	23	210
4	2	3	2	9	4
5	0	5	6	38	12
6	0	0	0	0	0

UL	dIO 0	1	2	3	4	5
1	4	3	0	0	0	3
2	11	2	0	0	0	0
3	186	44	1	9	0	0
4	10	4	3	3	0	0
5	40	6	6	7	2	0
6	0	0	0	0	0	0

dIO	Time 1	2	3	4	5
0	0	3	2	50	196
1	2	2	7	14	34
2	0	3	4	3	0
3	0	0	0	14	5
4	0	2	0	0	0
5	0	0	0	0	3

Source 3

IO	Time 1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	2	8
3	0	1	2	58	19
4	0	2	8	12	20
5	1	1	12	12	14
6	3	6	2	2	3
7	6	4	0	1	0
8	1	1	1	0	0
9	0	0	0	0	0
10	0	0	0	0	0
11	0	0	0	0	0

IO	UL 1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	0	7	3	0	0
3	0	1	17	58	4	0
4	4	0	14	16	8	0
5	0	0	11	21	7	1
6	2	1	2	7	4	0
7	0	0	3	6	2	0
8	0	0	0	1	2	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0

IO	dIO 0	1	2	3	4	5
1	0	0	0	0	0	0
2	8	2	0	0	0	0
3	47	9	0	24	0	0
4	18	12	7	0	1	4
5	24	14	1	0	0	1
6	5	2	6	1	0	2
7	3	7	1	0	0	0
8	1	2	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0

UL	Time 1	2	3	4	5
1	0	0	0	0	6
2	0	0	0	1	1
3	1	3	2	6	42
4	7	10	19	67	9
5	3	2	4	13	5
6	0	0	0	0	1

UL	dIO 0	1	2	3	4	5
1	0	0	0	0	0	6
2	1	1	0	0	0	0
3	42	9	3	0	0	0
4	47	32	9	23	1	0
5	16	6	3	2	0	0
6	0	0	0	0	0	1

dIO	Time 1	2	3	4	5
0	3	4	6	43	50
1	4	6	15	16	7
2	4	4	3	4	0
3	0	0	1	24	0
4	0	1	0	0	0
5	0	0	0	0	7

Table 4.4 - (End)

Case	No. of cells on each side of the epicentral cell	Total no. of cells
Case A, all I_0	0	1
Case B, $I_0 \leq$ III	1	9
I_0 = IV	2	25
I_0 = V	3	49
$I_0 \geq$ VI	4	81

Table 4.5 - Number of quarter-degree cells around the epicenter used in the definition of population category

Category	Time interval
1	1625-1779
2	1780-1859
3	1860-1909
4	1910-1949
5	1950-1980

Table 4.6 - Time categories

Intensity, I_0	No. in catalog since 1910	No. not detected by instruments
V	75	34
VI	20	5
VII	4	2
VIII	1	0

Table 4.7 - Number of earthquakes of high intensity not detected by instruments

Date	Time	Coordinates		I_0	Location
1918 Aug 21	0515	44.2	70.5	VII	ME Bridgeton-Norway
1925 Oct 09	1355	43.7	71.1	VI	NE Ossipee
1927 Jun 01	1223	40.3	74.0	VI	NJ Toms River-Sandy Hook
1928 Feb 08	A.M.	45.3	69.0	VI	ME Milo
1952 Jan 30	0400	44.5	73.2	VI	VT Burlington
1954 Jan 07	0725	40.3	76.0	VI	PA Berks Co.
1954 Feb 21	2000	41.2	75.9	VII	Wilkes-Barre

Table 4.8 - Earthquakes of intensity VI and VII without assigned magnitude

	Assumption on slope parameters		Province 1	2	3	4	5	6	7
CASE A	Unrelated b-values	a	124.	356.	389.	215.	277.	55.	44.
		b	1.48	1.35	1.59	1.25	1.39	1.11	1.17
	b-values from iid random variables	a	110.	357.	305.	223.	269.	63.	55.
		b	1.43	1.35	1.50	1.25	1.38	1.14	1.23
	Identical b-values	a	71.	301.	160.	253.	208.	102.	67.
		b			1.29				
	Two groups with same b-value	a	75.	313.	168.	262.	216.	51.	33.
		b			1.33			1.10	
CASE B	Unrelated b-values	a	117.	343.	337.	200.	247.	54.	43.
		b	1.45	1.36	1.57	1.25	1.43	1.12	1.16
	b-values from iid random variables	a	107.	345.	264.	210.	233.	63.	56.
		b	1.41	1.36	1.47	1.26	1.40	1.16	1.23
	Identical b-values	a	77.	290.	155.	239.	170.	98.	69.
		b			1.30				
CASE C	Unrelated b-values	a	153.	314.	192.	295.	321.	67.	40.
		b	1.51	1.33	1.40	1.34	1.49	1.17	1.14
	b-values from iid random variables	a	83.	299.	143.	276.	184.	114.	75.
		b			1.32				
	Identical b-values	a	83.	299.	143.	276.	184.	114.	75.
		b			1.32				

Table 4.9 - Parameters a and b in the relationship $\ln \lambda = a - bI_0$ (λ is the recurrence rate per 100 years and 771.5 km^2)

Class	Population Density Actual(1)	Nominal(2)
0	<2	100
1	2-5	1,000
2	6-17	3,500
3	18-44	10,000
4	45-89	20,000
5	≥90	50,000

(1) Inhabitants per square mile
(2) Inhabitants per quarter degree cell (771 km^2)

Table 4.10 - Nominal values of population density

r	(1)	Equivalent population category 1	2	3	4	5	6	7	8
3	from	1.43	3.48	5.52	7.57	9.61	11.66	13.70	15.75
	to	3.48	5.52	7.57	9.61	11.66	13.70	15.75	17.79
5	from	-2.16	0.35	2.58	5.36	7.86	10.36	12.87	15.37
		0.35	2.58	5.36	7.86	10.36	12.87	15.37	17.87

(1) the intervals shown refer to the natural logarithm of the integrated population density (see Eq. 4.29). The nominal density q in this equation is taken from Table 4.10.

Table 4.11 - Definition of discrete population categories p in Model B

Distance to the nearest instrument	Epicentral intensity					
	2	3	4	5	6	>6
≥305 km	1	1	2	3	4	5
195-304	1	2	3	4	5	5
110-194	2	3	4	5	5	5
55-109	3	4	5	5	5	5
0-54	4	5	5	5	5	5

Table 4.12 - Definition of instrument category d in Model B

Category t^*	Time interval
1	1910-1929
2	1930-1949
3	1950-1969
4	1970-1980

Table 4.13 - Definition of time category t^* in Model B

	(a) Events detected by people					(b) Events detected by instruments but not by people		
population category(1)	time category t					time category t		
	1	2	3	4	5	1-3	4	5
1	3	0	0	3	1	0	0	4
2	6	1	4	7	3	0	2	4
3	10	14	68	50	16	0	6	92
4	21	47	41	56	16	0	17	80
5	5	27	66	72	55	0	11	26
6	5	17	23	21	35	0	10	4
7	0	1	6	4	6	0	1	0
8	0	0	0	0	0	0	0	0
Total	50	107	208	213	132	0	47	210
Rate per year	0.3	1.3	4.2	5.3	4.3	0	1.2	6.8

	(c) Events detected by instruments				(d) Events detected by people but not by instruments			
instrument category	time category t*				time category t*			
	1	2	3	4	1	2	3	4
1	0	0	0	0	2	3	1	0
2	0	1	3	1	6	15	4	0
3	0	12	22	18	17	30	6	0
4	0	20	25	84	26	35	8	1
5	3	27	44	89	35	28	35	1
Total	3	60	94	192	86	111	54	2
Rate per year	0.2	3.0	4.7	17.5	4.5	5.6	2.7	0.2

(1) population category p for r=5

Table 4.14 - Earthquake counts for different mode of detections (Model B)

		smoothness of a and b				weights on counts for each I_0 for (a,b)-estimation					
						I_0					
	r	Δ_a	Δ_b	P_a	P_b	2	3	4	5	6	7-8
Base Case	5	L	G	0	10	0.01	0.10	0.25	0.50	0.75	1.00
Case 1	3										
2	I_0										
3					1						
4			L								
5						0.10	0.25	0.50	0.75	1.00	1.00
6						1.00	1.00	1.00	1.00	1.00	1.00
7						0.00	1.00	1.00	1.00	1.00	1.00
8						0.00	0.00	0.00	1.00	1.00	1.00

Note : - only changes to the parameters are indicated for Cases 1-8

- Δ_a and Δ_b indicate how interpolated values are calculated
  "L" : local interpolation, using only neighboring estimates
  "G" : global interpolation, using all estimates

- parameters common to all cases are :

  $P_\alpha = 100$
  no smoothing of β and γ parameters
  $\alpha = 1$ for p=8 (p=7 for Case 2) and all t
  for p=7 (p=6 for Case 2) and t=4,5

Table 4.15 - Input data for base and sensitivity cases (Model B)

Base Case

$I_o = 2$								$I_o = 4$								$I_o = 6$							
259.	225.	641.	280.	74.	51.	59.	242.	15.1	14.2	33.9	27.0	6.1	2.9	4.0	12.3	0.88	0.90	1.79	2.60	0.50	0.16	0.28	0.62
14.	204.	483.	377.	65.	143.	155.	289.	0.8	13.4	28.2	25.8	3.4	8.2	10.6	15.6	0.05	0.89	1.65	1.77	0.18	0.47	0.73	0.84
0.	55.	81.	147.	127.	584.	560.	295.	0.0	3.6	4.4	9.5	6.7	34.1	23.3	15.6	0.00	0.24	0.24	0.61	0.35	1.99	0.97	0.82
38.	0.	69.	84.	126.	361.	468.	24.	2.0	0.0	4.6	4.5	5.8	17.3	38.4	1.4	0.11	0.00	0.31	0.24	0.27	0.83	3.15	0.08
13.	51.	64.	273.	377.	293.	180.	53.	0.8	4.3	3.7	17.6	20.3	14.9	10.9	3.2	0.05	0.36	0.21	1.14	1.10	0.76	0.67	0.20
61.	67.	120.	182.	6.	25.	13.	0.	5.4	5.8	11.3	18.6	0.4	1.5	0.8	0.0	0.48	0.51	1.07	1.90	0.02	0.09	0.05	0.00
0.	63.	89.	42.	27.	5.	118.	0.	0.0	7.7	5.4	2.6	1.5	0.3	9.3	0.0	0.00	0.94	0.33	0.16	0.09	0.02	0.73	0.00

Case 1

259.	231.	661.	297.	77.	48.	53.	242.	14.5	14.3	34.3	28.0	6.1	2.6	3.6	11.8	0.82	0.89	1.77	2.65	0.48	0.14	0.25	0.58
13.	232.	483.	382.	64.	140.	166.	305.	0.7	14.7	27.1	24.7	3.3	7.7	10.9	15.8	0.04	0.93	1.52	1.59	0.17	0.42	0.72	0.82
0.	59.	84.	156.	141.	601.	557.	274.	0.0	3.8	4.4	9.7	7.2	34.4	23.2	14.2	0.00	0.25	0.23	0.60	0.36	1.97	0.96	0.74
40.	0.	79.	83.	118.	340.	441.	23.	2.1	0.0	5.1	4.5	5.5	16.9	37.7	1.3	0.11	0.00	0.33	0.24	0.26	0.84	3.22	0.07
13.	50.	65.	232.	331.	284.	185.	51.	0.8	4.1	3.6	16.2	19.0	14.7	11.2	3.0	0.05	0.34	0.20	1.14	1.09	0.76	0.68	0.18
59.	66.	112.	170.	6.	23.	11.	0.	5.2	5.8	10.8	17.7	0.4	1.3	0.7	0.0	0.47	0.50	1.04	1.85	0.02	0.08	0.04	0.00
0.	61.	85.	36.	25.	3.	76.	0.	0.0	7.3	5.1	2.2	1.4	0.2	6.0	0.0	0.00	0.88	0.30	0.13	0.08	0.01	0.47	0.00

Case 2

250.	292.	835.	461.	93.	52.	52.	224.	12.2	15.1	36.9	34.2	6.3	2.4	3.0	9.7	0.60	0.78	1.63	2.54	0.42	0.11	0.18	0.42
15.	250.	558.	443.	81.	169.	218.	583.	0.7	15.3	27.8	24.9	3.5	7.9	12.0	16.6	0.03	0.81	1.38	1.40	0.15	0.37	0.66	0.72
0.	89.	111.	212.	189.	879.	748.	203.	0.0	4.3	4.9	11.2	8.0	42.0	26.5	9.7	0.00	0.23	0.22	0.59	0.34	2.01	0.94	0.46
51.	0.	102.	134.	243.	734.	662.	16.	2.3	0.0	5.6	5.9	8.6	27.1	44.5	0.8	0.10	0.00	0.31	0.26	0.31	1.00	2.99	0.04
17.	76.	97.	444.	675.	468.	213.	34.	0.8	5.1	4.5	23.5	29.2	19.1	10.8	1.8	0.04	0.34	0.21	1.24	1.26	0.78	0.55	0.09
100.	139.	232.	254.	7.	19.	7.	0.	7.0	9.2	16.6	21.3	0.3	0.9	0.3	0.0	0.49	0.60	1.18	1.78	0.02	0.05	0.02	0.00
1.	109.	124.	32.	16.	2.	39.	0.	0.1	10.3	6.1	1.7	0.8	0.1	2.9	0.0	0.00	0.97	0.30	0.09	0.04	0.01	0.22	0.00

Case 3

342.	259.	790.	244.	37.	132.	45.	424.	17.0	15.4	36.3	26.0	4.8	3.7	3.8	14.7	0.85	0.92	1.67	2.77	0.63	0.10	0.32	0.51
38.	224.	578.	413.	164.	213.	159.	412.	1.1	14.5	30.6	27.2	4.1	9.4	11.3	17.5	0.03	0.93	1.62	1.79	0.10	0.42	0.81	0.74
0.	55.	177.	165.	230.	692.	849.	525.	0.0	3.7	5.2	10.2	7.9	36.6	25.6	19.4	0.00	0.25	0.16	0.64	0.27	1.93	0.77	0.71
106.	0.	64.	166.	305.	487.	478.	111.	2.4	0.0	4.7	5.3	6.3	18.3	40.0	2.4	0.06	0.00	0.34	0.17	0.13	0.69	3.35	0.05
22.	22.	113.	305.	468.	418.	228.	93.	0.9	3.2	4.3	18.5	21.5	16.7	12.3	4.2	0.04	0.45	0.16	1.13	0.99	0.67	0.66	0.19
29.	41.	86.	135.	13.	70.	36.	0.	4.3	5.1	10.3	17.2	0.5	2.2	1.2	0.0	0.62	0.64	1.24	2.19	0.02	0.07	0.04	0.00
1.	25.	118.	71.	111.	10.	36.	0.	0.0	5.3	6.0	3.3	2.7	0.4	5.2	0.0	0.00	1.10	0.30	0.15	0.07	0.02	0.74	0.00

Case 4

231.	244.	648.	279.	73.	58.	80.	301.	15.8	14.8	34.3	26.9	6.1	2.9	4.6	13.3	0.89	0.90	1.81	2.59	0.51	0.15	0.26	0.59
15.	221.	496.	394.	70.	172.	197.	359.	0.8	14.0	29.0	26.6	3.6	8.7	11.5	16.2	0.05	0.89	1.69	1.78	0.18	0.44	0.57	0.73
0.	56.	85.	174.	163.	659.	651.	382.	0.0	3.6	4.4	10.0	7.2	34.8	24.0	20.6	0.00	0.23	0.23	0.57	0.32	1.84	0.89	1.11
31.	0.	66.	99.	159.	416.	522.	33.	2.0	0.0	4.6	4.7	6.0	17.6	36.3	1.5	0.12	0.00	0.32	0.23	0.23	0.75	3.00	0.07
7.	32.	47.	271.	410.	337.	211.	62.	0.6	3.6	3.4	17.5	20.4	15.5	11.6	3.5	0.06	0.41	0.24	1.13	1.02	0.71	0.64	0.20
28.	39.	87.	165.	6.	30.	15.	0.	4.1	5.0	10.2	18.3	0.4	1.6	0.8	0.0	0.60	0.63	1.20	2.03	0.02	0.08	0.05	0.00
0.	33.	47.	27.	24.	5.	128.	0.	0.0	5.9	4.4	2.2	1.5	0.3	9.9	0.0	0.00	1.06	0.42	0.17	0.09	0.02	0.76	0.00

Case 5

142.	180.	448.	189.	38.	43.	28.	172.	11.4	12.3	28.4	21.4	4.0	2.8	2.6	10.3	0.92	0.85	1.79	2.42	0.41	0.18	0.24	0.61
13.	132.	375.	235.	52.	86.	110.	175.	0.9	10.8	24.2	20.5	3.2	6.5	9.0	12.5	0.06	0.89	1.56	1.79	0.20	0.49	0.74	0.89
0.	44.	69.	110.	87.	386.	364.	194.	0.0	3.3	4.2	7.8	5.7	28.1	19.6	12.5	0.00	0.24	0.26	0.56	0.38	2.05	1.06	0.81
36.	0.	45.	67.	100.	246.	268.	27.	2.1	0.0	3.8	4.2	5.4	15.3	29.7	1.7	0.13	0.00	0.32	0.27	0.29	0.95	3.29	0.11
9.	28.	48.	175.	221.	180.	126.	27.	0.7	2.9	3.3	14.7	16.4	12.3	9.2	2.2	0.05	0.30	0.23	1.23	1.22	0.85	0.67	0.17
25.	30.	66.	108.	6.	23.	13.	0.	3.4	4.1	8.6	14.0	0.4	1.6	0.9	0.0	0.46	0.55	1.12	1.83	0.03	0.11	0.07	0.00
1.	28.	48.	22.	35.	22.	40.	0.	0.1	4.7	3.9	1.8	2.2	1.4	4.3	0.0	0.01	0.80	0.32	0.14	0.13	0.10	0.45	0.00

Case 6

55.	130.	267.	103.	19.	30.	109.	10.	7.7	10.7	23.3	17.4	2.8	2.5	15.4	0.8	1.09	0.88	2.03	2.93	0.43	0.21	2.16	0.06
10.	87.	227.	139.	32.	43.	88.	82.	1.0	8.9	20.6	16.0	2.8	4.9	7.4	10.0	0.09	0.91	1.87	1.85	0.25	0.55	0.62	1.23
0.	44.	54.	71.	49.	205.	163.	104.	0.0	3.1	3.9	7.0	4.8	22.7	15.7	9.2	0.00	0.22	0.27	0.69	0.47	2.52	1.51	0.82
28.	0.	34.	46.	58.	141.	122.	34.	2.3	0.0	3.3	3.8	5.0	13.6	21.4	2.4	0.18	0.00	0.32	0.31	0.42	1.31	3.76	0.17
3.	12.	31.	107.	102.	88.	97.	7.	0.4	2.0	2.8	11.9	13.8	9.8	7.5	0.9	0.04	0.33	0.26	1.31	1.87	1.08	0.58	0.12
8.	10.	43.	55.	5.	19.	13.	0.	1.7	2.4	6.7	10.3	0.5	1.6	1.2	0.0	0.38	0.58	1.04	1.91	0.05	0.13	0.11	0.00
3.	13.	19.	6.	50.	72.	9.	0.	0.3	3.1	2.7	0.8	3.3	3.8	1.5	0.0	0.02	0.76	0.39	0.10	0.22	0.20	0.26	0.00

Case 7

145.	138.	485.	220.	35.	41.	24.	200.	11.7	11.3	28.3	23.0	3.7	2.9	2.3	11.2	0.94	0.93	1.66	2.40	0.40	0.21	0.21	0.63
28.	125.	436.	211.	50.	43.	53.	190.	1.7	10.5	26.0	19.1	3.5	4.9	6.5	13.3	0.11	0.88	1.55	1.74	0.24	0.55	0.79	0.93
0.	32.	62.	140.	89.	303.	366.	143.	0.0	2.7	4.2	8.9	6.0	26.4	20.5	11.3	0.00	0.23	0.28	0.56	0.40	2.30	1.15	0.89
58.	0.	32.	47.	97.	220.	198.	59.	3.1	0.0	3.1	3.8	5.8	15.4	26.8	3.3	0.16	0.00	0.30	0.30	0.34	1.08	3.63	0.19
8.	18.	32.	117.	216.	105.	85.	17.	0.7	2.3	2.8	12.7	17.0	10.7	7.6	1.6	0.06	0.30	0.25	1.38	1.34	1.10	0.67	0.16
10.	14.	44.	97.	13.	18.	34.	0.	1.9	2.7	6.8	13.4	0.9	1.4	2.2	0.0	0.38	0.53	1.07	1.85	0.06	0.12	0.15	0.00
0.	10.	54.	14.	87.	0.	21.	0.	0.0	2.7	4.3	1.3	4.6	0.0	2.6	0.0	0.00	0.71	0.34	0.12	0.24	0.00	0.33	0.00

Case 8

380.	204.	892.	231.	139.	0.	100.	480.	19.7	12.4	44.4	23.2	9.2	0.0	6.0	21.2	1.02	0.75	2.21	2.32	0.60	0.00	0.36	0.94
0.	360.	289.	523.	0.	291.	229.	299.	0.0	20.2	20.2	32.4	0.0	13.4	13.6	15.2	0.00	1.13	1.41	2.01	0.00	0.61	0.81	0.77
0.	54.	81.	99.	188.	643.	440.	231.	0.0	3.4	4.1	6.8	6.8	34.7	18.7	12.2	0.00	0.21	0.21	0.47	0.41	1.87	0.79	0.65
0.	0.	116.	65.	0.	250.	882.	0.	0.0	0.0	6.8	3.4	0.0	12.4	55.8	0.0	0.00	0.00	0.40	0.17	0.00	0.62	3.53	0.00
0.	36.	68.	476.	480.	377.	397.	168.	0.0	2.7	3.5	24.7	22.5	16.3	19.4	7.7	0.00	0.21	0.18	1.28	1.06	0.71	0.95	0.38
208.	187.	348.	300.	0.	0.	0.	0.	12.4	11.1	22.0	24.6	0.0	0.0	0.0	0.0	0.74	0.66	1.39	2.02	0.00	0.00	0.00	0.00
0.	331.	315.	116.	0.	0.	163.	0.	0.0	20.9	13.9	5.8	0.0	0.0	11.2	0.0	0.00	1.33	0.61	0.29	0.00	0.00	0.77	0.00

Note: expected count for 100 years and per unit-degree equatorial area

Table 4.16 - Estimated earthquake counts in Model B

Base Case

$I_o = 2$	$I_o = 4$	$I_o = 6$
107. 123. 266. 251. 34. 22. 14. 74. 1. 97. 236. 180. 36. 102. 105. 139. 1. 18. 32. 108. 79. 342. 212. 100. 4. 1. 35. 90. 139. 253. 371. 2. 7. 22. 40. 130. 241. 177. 55. 13. 29. 41. 55. 90. 1. 8. 1. 1. 1. 39. 16. 9. 1. 1. 11. 1.	7.4 8.5 19.4 22.8 2.9 1.6 1.0 5.4 0.1 7.1 17.2 14.2 2.7 6.7 7.5 9.0 0.1 1.4 2.4 7.4 4.8 20.0 11.6 6.6 0.3 0.1 2.8 6.0 7.5 13.4 25.4 0.1 0.8 2.0 3.4 10.2 14.4 10.3 3.8 0.9 3.1 4.4 6.2 9.0 0.1 0.6 0.1 0.1 0.1 4.7 1.8 0.8 0.1 0.1 0.9 0.1	0.50 0.58 1.41 2.07 0.24 0.12 0.08 0.39 0.01 0.51 1.25 1.12 0.19 0.44 0.54 0.58 0.01 0.11 0.17 0.51 0.29 1.17 0.64 0.43 0.03 0.01 0.22 0.39 0.40 0.71 1.74 0.01 0.06 0.19 0.29 0.61 0.86 0.60 0.26 0.07 0.34 0.47 0.70 0.90 0.01 0.04 0.01 0.01 0.01 0.57 0.20 0.08 0.01 0.01 0.07 0.01

Case 1

111. 142. 310. 316. 38. 30. 15. 97. 1. 104. 266. 195. 45. 120. 131. 167. 1. 28. 44. 128. 93. 371. 251. 127. 13. 1. 41. 120. 173. 288. 444. 1. 11. 43. 47. 138. 264. 214. 62. 15. 34. 50. 54. 96. 1. 15. 1. 1. 1. 40. 15. 10. 1. 1. 12. 1.	7.6 9.7 22.5 27.5 3.1 2.1 1.1 6.5 0.1 7.4 19.0 15.4 3.2 7.6 8.6 9.9 0.1 2.0 3.1 8.5 5.4 20.4 12.5 7.6 1.0 0.1 3.0 7.4 8.8 13.7 27.0 0.1 1.0 3.8 3.9 10.4 14.5 11.3 3.8 1.0 3.5 5.0 6.0 9.1 0.1 1.0 0.1 0.1 0.1 4.7 1.6 0.9 0.1 0.1 1.0 0.1	0.52 0.67 1.64 2.40 0.26 0.15 0.08 0.44 0.01 0.53 1.36 1.22 0.22 0.48 0.57 0.59 0.01 0.15 0.22 0.56 0.32 1.12 0.62 0.45 0.08 0.01 0.22 0.46 0.45 0.65 1.64 0.01 0.08 0.34 0.32 0.79 0.80 0.60 0.24 0.07 0.37 0.50 0.66 0.86 0.01 0.07 0.01 0.01 0.01 0.55 0.17 0.09 0.01 0.01 0.07 0.01

Case 2

116. 134. 289. 266. 37. 24. 15. 81. 2. 106. 256. 195. 41. 111. 115. 152. 1. 20. 35. 119. 86. 378. 234. 108. 6. 1. 38. 98. 152. 280. 410. 2. 8. 24. 43. 143. 266. 195. 61. 14. 32. 47. 61. 99. 1. 9. 1. 1. 1. 44. 18. 10. 1. 1. 12. 1.	8.1 9.3 21.4 25.2 3.2 1.8 1.1 6.0 0.1 7.9 19.0 15.7 3.0 7.4 8.4 10.0 0.1 1.6 2.6 8.3 5.3 22.5 13.1 7.3 0.4 0.1 3.1 6.6 8.4 15.1 28.7 0.1 0.7 2.3 3.8 11.5 16.2 11.6 4.2 1.0 3.5 5.0 7.0 10.1 0.1 0.6 0.1 0.1 0.1 5.3 2.0 0.9 0.1 0.1 1.0 0.1	0.57 0.65 1.59 2.38 0.28 0.14 0.09 0.45 0.01 0.59 1.41 1.27 0.22 0.50 0.61 0.66 0.01 0.12 0.19 0.58 0.33 1.34 0.74 0.49 0.03 0.01 0.25 0.44 0.46 0.81 2.00 0.01 0.06 0.22 0.33 0.93 0.99 0.69 0.30 0.08 0.38 0.53 0.81 1.04 0.01 0.04 0.01 0.01 0.01 0.65 0.23 0.09 0.01 0.01 0.08 0.01

Case 3

115. 135. 292. 251. 23. 22. 14. 66. 2. 108. 251. 192. 26. 87. 106. 90. 1. 18. 34. 97. 47. 335. 205. 70. 4. 1. 27. 94. 122. 253. 363. 2. 7. 23. 41. 130. 241. 177. 55. 13. 30. 42. 55. 90. 1. 8. 1. 1. 1. 39. 17. 9. 1. 1. 11. 1.	7.6 8.7 20.0 22.3 1.9 1.6 1.0 4.7 0.1 7.3 17.6 14.6 1.8 5.8 7.6 6.1 0.1 1.4 2.4 6.7 2.9 20.0 11.5 4.8 0.3 0.1 2.1 6.1 6.7 13.4 25.4 0.1 0.6 2.1 3.5 10.2 14.4 10.3 3.8 1.0 3.2 4.4 6.2 9.0 0.1 0.6 0.1 0.1 0.1 4.7 1.8 0.9 0.1 0.1 0.9 0.1	0.50 0.56 1.37 1.98 0.16 0.12 0.08 0.33 0.01 0.60 1.23 1.10 0.14 0.39 0.54 0.41 0.01 0.10 0.17 0.46 0.18 1.19 0.65 0.33 0.03 0.01 0.18 0.39 0.37 0.71 1.78 0.01 0.06 0.19 0.29 0.61 0.86 0.60 0.26 0.07 0.34 0.46 0.70 0.90 0.01 0.04 0.01 0.01 0.01 0.57 0.20 0.08 0.01 0.01 0.07 0.01

Case 4

157. 186. 418. 394. 52. 32. 21. 114. 2. 143. 363. 283. 59. 162. 168. 218. 1. 25. 47. 168. 128. 850. 510. 168. 6. 1. 53. 143. 239. 649. 968. 3. 10. 31. 63. 303. 687. 418. 120. 20. 44. 88. 123. 203. 1. 12. 1. 1. 1. 87. 36. 14. 1. 1. 15. 1.	9.9 11.5 26.4 31.0 4.0 2.2 1.4 7.4 0.1 9.6 23.5 19.4 3.8 9.3 10.4 12.5 0.1 1.9 3.2 10.2 5.7 39.1 22.5 9.3 0.4 0.1 3.8 8.2 10.5 26.4 52.2 0.1 0.8 2.7 4.7 19.6 27.5 20.0 7.1 1.3 4.2 6.0 11.4 16.7 0.1 0.7 0.1 0.1 0.1 6.7 3.3 1.2 0.1 0.1 1.2 0.1	0.63 0.72 1.67 2.45 0.30 0.15 0.09 0.47 0.01 0.65 1.61 1.33 0.24 0.53 0.65 0.72 0.01 0.14 0.21 0.62 0.34 1.80 1.00 0.51 0.03 0.01 0.27 0.47 0.47 1.08 2.60 0.01 0.07 0.24 0.35 1.27 1.29 0.96 0.43 0.09 0.41 0.72 1.08 1.37 0.01 0.05 0.01 0.01 0.01 0.87 0.31 0.10 0.01 0.01 0.09 0.01

Case 5

112. 132. 292. 210. 29. 22. 12. 69. 1. 95. 253. 172. 40. 109. 93. 149. 1. 17. 32. 106. 86. 360. 259. 100. 4. 1. 33. 97. 168. 297. 329. 2. 8. 21. 43. 123. 283. 199. 52. 12. 28. 44. 47. 74. 1. 8. 1. 1. 1. 35. 17. 9. 1. 1. 10. 1.	7.4 8.5 19.6 21.9 2.8 1.6 1.0 5.3 0.1 7.1 17.4 14.1 2.8 6.8 7.4 9.1 0.1 1.4 2.4 7.4 4.8 20.2 11.9 6.6 0.3 0.1 2.7 6.0 7.7 13.7 24.9 0.1 0.8 2.0 3.5 10.1 14.7 10.5 3.7 0.9 3.1 4.5 6.0 8.5 0.1 0.6 0.1 0.1 0.1 4.6 1.8 0.8 0.1 0.1 0.9 0.1	0.48 0.55 1.32 2.28 0.27 0.12 0.08 0.41 0.01 0.52 1.19 1.16 0.19 0.42 0.58 0.55 0.01 0.11 0.17 0.52 0.27 1.13 0.55 0.43 0.03 0.01 0.22 0.37 0.38 0.63 1.89 0.01 0.06 0.20 0.28 0.63 0.76 0.56 0.27 0.07 0.34 0.46 0.75 0.99 0.01 0.04 0.01 0.01 0.01 0.60 0.20 0.08 0.01 0.01 0.08 0.01

Case 6

99. 114. 259. 289. 38. 22. 14. 78. 1. 94. 229. 188. 37. 96. 111. 132. 1. 18. 31. 103. 70. 297. 175. 99. 4. 1. 35. 82. 112. 205. 371. 2. 7. 23. 39. 126. 199. 151. 53. 14. 32. 44. 63. 99. 1. 8. 1. 1. 1. 45. 18. 9. 1. 1. 12. 1.	7.2 8.3 19.2 23.2 2.9 1.6 1.0 5.4 0.1 7.0 17.0 14.2 2.7 6.6 7.6 8.9 0.1 1.4 2.3 7.4 4.6 19.2 11.1 6.5 0.3 0.1 2.7 6.8 7.2 12.7 24.9 0.1 0.8 2.0 3.4 9.9 13.7 9.9 3.7 0.9 3.2 4.4 6.3 9.0 0.1 0.6 0.1 0.1 0.1 4.8 1.8 0.9 0.1 0.1 0.9 0.1	0.52 0.60 1.43 1.87 0.23 0.12 0.07 0.38 0.01 0.52 1.27 1.08 0.20 0.45 0.52 0.60 0.01 0.11 0.17 0.52 0.31 1.24 0.70 0.43 0.03 0.01 0.22 0.42 0.47 0.79 1.67 0.01 0.06 0.19 0.29 0.78 0.94 0.65 0.25 0.07 0.32 0.44 0.63 0.81 0.01 0.04 0.01 0.01 0.01 0.51 0.18 0.08 0.01 0.01 0.07 0.01

Case 7

109. 126. 266. 244. 34. 23. 14. 79. 1. 99. 234. 182. 39. 109. 113. 149. 1. 17. 31. 111. 86. 357. 225. 109. 4. 1. 31. 92. 152. 275. 390. 2. 5. 16. 32. 117. 239. 180. 56. 13. 18. 28. 39. 71. 1. 8. 1. 1. 1. 25. 11. 7. 1. 1. 11. 1.	7.4 8.5 19.4 22.5 2.9 1.6 1.0 5.4 0.1 7.1 17.0 14.1 2.7 6.8 7.6 9.1 0.1 1.4 2.3 7.4 4.8 19.8 11.4 6.7 0.3 0.1 2.7 6.0 7.6 13.1 25.2 0.1 0.6 2.0 3.2 9.8 13.9 10.1 3.7 0.9 2.8 3.9 5.6 8.4 0.1 0.6 0.1 0.1 0.1 4.1 1.6 0.8 0.1 0.1 0.9 0.1	0.49 0.57 1.41 2.09 0.24 0.11 0.07 0.37 0.01 0.50 1.24 1.09 0.19 0.42 0.51 0.55 0.01 0.11 0.17 0.50 0.27 1.07 0.58 0.40 0.03 0.01 0.23 0.38 0.38 0.63 1.62 0.01 0.07 0.24 0.32 0.82 0.81 0.57 0.24 0.07 0.41 0.55 0.80 0.99 0.01 0.04 0.01 0.01 0.01 0.67 0.23 0.09 0.01 0.01 0.07 0.01

Case 8

101. 117. 264. 284. 38. 21. 13. 72. 1. 94. 234. 182. 37. 94. 100. 126. 1. 18. 31. 100. 69. 292. 175. 90. 4. 1. 36. 84. 115. 203. 329. 2. 8. 26. 45. 128. 208. 151. 50. 12. 39. 54. 70. 104. 1. 8. 1. 1. 1. 55. 22. 11. 1. 1. 12. 1.	7.2 8.4 19.2 23.0 2.9 1.6 1.0 5.3 0.1 7.0 17.0 14.1 2.7 6.6 7.4 8.8 0.1 1.4 2.3 7.3 4.6 19.2 11.1 6.4 0.3 0.1 2.6 5.6 7.3 12.6 24.4 0.1 0.6 2.1 3.5 9.9 13.7 9.9 3.6 0.9 3.3 4.6 6.4 9.1 0.1 0.6 0.1 0.1 0.1 5.0 1.9 0.9 0.1 0.1 0.9 0.1	0.51 0.60 1.40 2.00 0.23 0.12 0.08 0.40 0.01 0.52 1.24 1.09 0.20 0.46 0.55 0.62 0.01 0.11 0.17 0.53 0.31 1.27 0.70 0.46 0.02 0.01 0.21 0.41 0.46 0.78 1.81 0.01 0.05 0.17 0.27 0.77 0.90 0.65 0.26 0.07 0.29 0.38 0.58 0.79 0.01 0.04 0.01 0.01 0.01 0.46 0.16 0.07 0.01 0.01 0.07 0.01

Case 9

107. 155. 275. 229. 95. 71. 70. 93. 58. 102. 218. 168. 113. 115. 111. 130. 30. 41. 77. 132. 151. 229. 173. 132. 21. 24. 51. 109. 184. 199. 199. 73. 19. 24. 45. 84. 142. 111. 48. 36. 20. 24. 34. 48. 37. 27. 16. 15. 17. 22. 24. 26. 22. 16. 11. 11.	7.2 10.4 20.8 21.7 7.2 4.6 4.7 6.5 3.6 7.3 16.5 14.8 7.4 6.9 7.7 8.4 2.0 2.8 5.3 8.9 8.6 14.2 9.7 8.7 1.5 1.8 3.8 7.2 10.1 11.0 15.7 4.7 1.6 2.2 4.0 7.2 9.3 6.9 3.2 2.5 2.2 2.8 4.1 5.3 2.7 1.8 1.1 1.0 1.8 2.8 2.7 2.5 1.8 1.1 0.8 0.8	0.48 0.70 1.58 2.04 0.55 0.30 0.32 0.46 0.23 0.52 1.25 1.17 0.49 0.43 0.54 0.54 0.13 0.20 0.36 0.60 0.49 0.89 0.55 0.57 0.11 0.13 0.29 0.47 0.56 0.60 1.24 0.30 0.14 0.21 0.36 0.62 0.62 0.43 0.22 0.17 0.24 0.29 0.60 0.59 0.20 0.12 0.08 0.07 0.18 0.36 0.30 0.24 0.14 0.08 0.06 0.06

Case 10

127. 102. 346. 292. 56. 13. 33. 120. 1. 135. 227. 234. 22. 101. 145. 136. 1. 33. 40. 113. 70. 348. 155. 109. 2. 1. 63. 112. 63. 225. 453. 1. 6. 32. 46. 205. 220. 173. 78. 30. 71. 76. 138. 157. 1. 5. 1. 1. 1. 95. 46. 23. 1. 1. 27. 1.	8.5 7.0 23.7 25.4 4.7 0.9 2.4 8.0 0.1 9.4 16.5 17.7 1.6 6.8 9.9 8.8 0.1 2.5 2.9 7.9 4.5 21.0 9.1 7.2 0.2 0.1 4.6 7.2 3.7 12.6 30.4 0.1 0.5 2.7 3.5 14.7 13.4 10.5 5.2 2.1 5.2 6.5 11.8 13.1 0.1 0.4 0.1 0.1 0.1 8.6 4.0 1.9 0.1 0.1 2.1 0.1	0.57 0.48 1.62 2.21 0.38 0.07 0.17 0.54 0.01 0.66 1.20 1.34 0.12 0.46 0.68 0.57 0.01 0.18 0.20 0.55 0.28 1.28 0.63 0.47 0.01 0.01 0.33 0.47 0.22 0.71 2.04 0.01 0.04 0.23 0.27 1.05 0.81 0.64 0.35 0.15 0.54 0.65 1.01 1.10 0.01 0.03 0.01 0.01 0.01 0.78 0.34 0.15 0.01 0.01 0.16 0.01

Note: expected count for 100 years and per unit-degree equatorial area

Table 4.17 - Estimated earthquake counts in Model D

Case 11

$I_o = 2$

184.	85.	5.	197.	264.	346.	188.	218.	1.	1.	1.	36.	1.	1.	1.	6.
60.	115.	168.	116.	229.	218.	466.	116.	83.	48.	51.	1.	49.	1.	54.	236.
1.	6.	1.	319.	490.	259.	256.	175.	1.	50.	112.	1.	1.	203.	127.	1.
1.	1.	1.	49.	1.	160.	68.	177.	53.	60.	248.	60.	1.	210.	306.	119.
1.	1.	6.	46.	47.	76.	96.	1.	61.	203.	140.	289.	39.	374.	353.	58.
1.	1.	1.	17.	1.	1.	297.	1.	2.	76.	516.	587.	180.	170.	1.	1.
17.	1.	1.	1.	90.	48.	286.	1.	72.	162.	176.	696.	876.	68.	1.	1.
1.	1.	1.	1.	1.	1.	49.	1.	229.	168.	2.	275.	587.	140.	6.	1.
1.	1.	1.	1.	1.	1.	114.	57.	342.	163.	53.	266.	125.	35.	47.	1.
29.	1.	64.	17.	32.	96.	83.	272.	167.	431.	92.	363.	36.	35.	1.	1.
1.	29.	1.	86.	36.	57.	229.	1.	1.	1.	1.	1.	1.	1.	1.	1.
1.	90.	58.	8.	55.	54.	84.	2.	1.	1.	1.	31.	1.	1.	1.	1.
1.	1.	59.	77.	57.	2.	34.	1.	1.	3.	1.	1.	41.	1.	1.	1.
1.	1.	1.	17.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.

$I_o = 4$

12.4	5.7	0.3	13.3	19.8	25.2	15.4	19.8	0.1	0.1	0.1	2.7	0.1	0.1	0.1	0.4
4.0	7.6	11.3	8.8	16.7	16.2	44.9	10.8	7.7	3.7	3.7	0.1	4.0	0.1	4.0	17.2
0.1	0.4	0.1	24.2	37.2	19.2	21.4	14.4	0.1	3.6	7.5	0.1	0.1	14.8	8.7	0.1
0.1	0.1	0.1	3.9	0.1	11.9	5.1	15.1	3.7	3.8	15.4	4.0	0.1	15.3	19.8	7.5
0.1	0.1	0.4	3.8	3.6	5.7	7.1	0.1	4.0	11.4	7.7	19.0	2.4	22.8	22.8	3.7
0.1	0.1	0.1	1.3	0.1	0.1	22.1	0.1	0.1	3.9	25.2	29.2	10.1	10.5	0.1	0.1
1.3	0.1	0.1	0.1	7.3	3.8	20.0	0.1	3.8	7.7	8.4	32.0	53.3	5.1	0.1	0.1
0.1	0.1	0.1	0.1	0.1	0.1	3.5	0.1	11.4	7.7	0.1	13.9	35.7	9.8	0.5	0.1
0.1	0.1	0.1	0.1	0.1	0.1	9.7	3.7	17.7	8.0	2.6	14.4	7.6	2.6	3.7	0.1
2.5	0.1	6.5	1.7	3.2	9.7	7.8	19.4	9.7	22.8	4.9	20.0	2.3	2.6	0.1	0.1
0.1	2.8	0.1	9.9	4.5	6.5	25.9	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.1	9.8	5.4	0.9	5.7	6.5	8.7	0.2	0.1	0.1	0.1	2.2	0.1	0.1	0.1	0.1
0.1	0.1	6.5	10.2	6.8	0.2	3.4	0.1	0.1	0.2	0.1	0.1	3.5	0.1	0.1	0.1
0.1	0.1	0.1	2.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Case 12

134.	103.	83.	192.	264.	356.	229.	192.	43.	24.	19.	21.	17.	20.	32.	46.
65.	86.	134.	149.	272.	280.	422.	128.	68.	37.	31.	20.	27.	28.	53.	127.
20.	24.	43.	229.	401.	272.	253.	165.	52.	61.	70.	35.	35.	131.	128.	84.
7.	8.	19.	55.	75.	160.	112.	128.	73.	84.	175.	74.	66.	180.	281.	139.
4.	4.	9.	23.	46.	80.	99.	54.	69.	163.	165.	234.	102.	292.	256.	91.
3.	3.	5.	14.	25.	55.	186.	58.	52.	109.	431.	500.	220.	158.	48.	35.
3.	3.	4.	8.	32.	55.	168.	63.	77.	142.	197.	599.	710.	85.	24.	16.
4.	4.	5.	8.	15.	34.	59.	74.	168.	155.	98.	275.	476.	115.	19.	12.
6.	7.	8.	11.	17.	30.	85.	95.	280.	167.	125.	220.	136.	45.	20.	9.
13.	12.	25.	21.	33.	75.	83.	205.	143.	283.	106.	223.	48.	22.	8.	6.
16.	24.	21.	56.	39.	61.	162.	38.	36.	31.	36.	27.	18.	8.	4.	3.
13.	35.	42.	28.	48.	50.	52.	17.	9.	8.	9.	11.	7.	4.	2.	2.
8.	10.	34.	46.	35.	16.	16.	7.	4.	4.	4.	5.	6.	3.	2.	2.
6.	8.	14.	22.	16.	11.	8.	5.	3.	3.	3.	4.	3.	3.	2.	2.

$I_o = 4$

9.2	6.9	5.4	12.9	19.6	25.9	18.8	17.4	3.7	1.9	1.4	1.6	1.2	1.5	2.3	3.4
4.3	5.7	9.0	11.3	19.8	20.8	41.5	11.9	6.2	2.8	2.2	1.4	2.1	2.0	3.9	9.4
1.3	1.6	3.1	18.1	31.7	20.2	21.2	13.5	4.1	4.3	4.7	2.4	2.6	9.7	9.0	5.7
0.5	0.6	1.4	4.3	5.7	11.9	8.3	11.2	5.1	5.3	11.1	5.0	4.6	13.7	17.6	9.0
0.3	0.3	0.7	1.9	3.5	6.1	7.4	3.9	4.5	9.2	9.3	15.7	6.3	18.1	17.2	5.8
0.2	0.2	0.4	1.0	2.0	4.2	14.4	3.7	2.9	5.7	21.4	25.9	12.6	9.8	3.1	2.3
0.2	0.2	0.3	0.7	2.6	4.4	12.2	3.9	4.1	6.8	9.4	28.7	45.8	6.3	1.7	1.1
0.3	0.3	0.4	0.7	1.2	2.7	4.3	4.6	8.5	7.3	4.6	14.2	30.1	8.2	1.4	0.9
0.5	0.6	0.7	1.0	1.5	2.6	7.6	6.1	15.1	8.3	6.1	12.4	8.5	3.3	1.5	0.7
1.1	1.0	2.7	2.2	3.4	8.0	8.0	15.6	8.5	15.6	5.7	12.7	3.1	1.6	0.6	0.5
1.5	2.4	2.3	6.9	4.9	7.4	19.8	3.2	2.4	1.9	2.2	1.7	1.3	0.6	0.3	0.2
1.3	4.1	5.1	3.5	6.1	6.2	5.7	1.5	0.6	0.5	0.6	0.8	0.5	0.3	0.2	0.2
0.8	1.1	4.0	6.6	4.4	1.7	1.6	0.6	0.3	0.3	0.3	0.4	0.5	0.2	0.1	0.1
0.6	0.8	1.5	2.7	1.7	1.1	0.7	0.4	0.3	0.2	0.2	0.3	0.3	0.2	0.1	0.1

Case 13

101.	106.	126.	216.	289.	374.	275.	180.	79.	48.	37.	34.	34.	41.	54.	66.
74.	90.	130.	179.	300.	346.	322.	155.	83.	55.	44.	37.	41.	50.	69.	94.
39.	46.	70.	163.	283.	286.	241.	165.	89.	76.	71.	56.	60.	100.	120.	113.
20.	23.	37.	70.	113.	172.	151.	122.	103.	104.	126.	96.	102.	142.	186.	154.
12.	13.	20.	34.	53.	97.	115.	97.	104.	155.	173.	182.	151.	210.	186.	130.
9.	9.	13.	23.	39.	68.	115.	96.	105.	155.	310.	374.	241.	168.	95.	79.
8.	8.	11.	17.	33.	59.	106.	100.	123.	175.	229.	448.	422.	117.	59.	46.
9.	9.	11.	17.	28.	49.	76.	107.	163.	179.	173.	266.	292.	106.	41.	30.
11.	11.	14.	20.	30.	47.	77.	113.	197.	165.	157.	175.	120.	53.	28.	20.
15.	15.	20.	26.	38.	60.	76.	131.	125.	158.	107.	117.	61.	26.	15.	13.
17.	19.	22.	33.	36.	51.	79.	55.	58.	52.	49.	35.	23.	13.	9.	8.
16.	20.	26.	28.	35.	39.	39.	30.	24.	21.	19.	16.	11.	7.	6.	5.
14.	16.	23.	26.	27.	24.	23.	18.	14.	12.	10.	9.	7.	5.	4.	4.
13.	15.	18.	22.	21.	20.	17.	15.	12.	10.	8.	7.	6.	5.	4.	4.

$I_o = 4$

6.9	7.1	8.3	14.2	21.4	27.8	22.5	16.4	6.3	3.5	2.5	2.4	2.4	2.9	3.9	4.7
4.8	5.6	8.7	13.5	21.9	25.7	33.6	14.1	7.2	3.9	2.9	2.5	3.0	3.5	5.0	7.0
2.5	3.0	4.9	13.1	23.2	21.7	20.6	13.5	6.6	5.1	4.5	3.8	4.3	7.6	8.5	7.7
1.3	1.5	2.6	5.4	8.5	12.7	11.2	10.4	6.9	6.3	8.0	6.3	7.0	11.2	13.3	10.1
0.8	0.9	1.4	2.7	4.6	7.2	8.4	6.7	6.5	8.5	9.7	12.7	9.3	13.5	12.7	8.2
0.6	0.6	0.9	1.6	2.9	5.2	9.1	6.1	5.7	7.7	16.0	20.6	14.4	10.6	6.0	5.0
0.5	0.6	0.8	1.3	2.6	4.6	7.9	6.1	6.3	8.2	11.2	23.2	30.7	8.5	3.9	3.0
0.6	0.6	0.8	1.3	2.2	3.9	5.6	6.7	8.3	8.4	8.1	14.4	20.4	7.7	2.8	2.1
0.8	0.9	1.2	1.7	2.7	4.2	7.0	7.4	11.1	8.4	7.8	10.2	7.7	3.8	2.0	1.4
1.3	1.3	2.1	2.6	4.0	6.8	7.7	10.8	7.7	9.3	5.9	7.0	3.4	1.9	1.1	0.9
1.6	2.0	2.5	4.4	4.8	6.6	11.1	4.8	4.0	3.2	3.0	2.2	1.6	0.9	0.6	0.6
1.6	2.4	3.3	3.6	4.8	5.2	4.6	2.6	1.7	1.4	1.2	1.1	0.8	0.6	0.4	0.4
1.4	1.6	2.8	4.0	3.6	2.6	2.2	1.4	1.0	0.8	0.7	0.6	0.6	0.4	0.3	0.3
1.2	1.4	2.0	2.7	2.5	2.0	1.5	1.1	0.8	0.7	0.6	0.5	0.4	0.4	0.3	0.3

Table 4.17 - (continued)

Case 11

$I_o = 6$

0.83	0.38	0.02	0.89	1.46	1.83	1.27	1.80	0.01	0.01	0.01	0.21	0.01	0.01	0.01	0.03
0.26	0.50	0.76	0.67	1.22	1.20	4.33	1.00	0.71	0.29	0.27	0.01	0.32	0.01	0.30	1.25
0.01	0.03	0.01	1.83	2.82	1.43	1.80	1.18	0.01	0.26	0.50	0.01	0.01	1.08	0.60	0.01
0.01	0.01	0.01	0.31	0.01	0.88	0.39	1.29	0.26	0.24	0.96	0.27	0.01	1.11	1.28	0.47
0.01	0.01	0.03	0.31	0.28	0.43	0.53	0.01	0.27	0.64	0.43	1.25	0.15	1.38	1.47	0.24
0.01	0.01	0.01	0.11	0.01	0.01	1.64	0.01	0.01	0.20	1.23	1.46	0.67	0.65	0.01	0.01
0.10	0.01	0.01	0.01	0.59	0.30	1.40	0.01	0.20	0.37	0.40	1.47	3.24	0.38	0.01	0.01
0.01	0.01	0.01	0.01	0.01	0.01	0.26	0.01	0.57	0.36	0.00	0.71	2.17	0.69	0.04	0.01
0.01	0.01	0.01	0.01	0.01	0.01	0.83	0.24	0.92	0.39	0.13	0.78	0.46	0.19	0.29	0.01
0.21	0.01	0.65	0.17	0.32	1.00	0.74	1.38	0.57	1.20	0.26	1.10	0.15	0.19	0.01	0.01
0.01	0.27	0.01	1.14	0.55	0.75	2.93	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.01	1.07	0.71	0.10	0.82	0.78	0.91	0.02	0.01	0.01	0.01	0.16	0.01	0.01	0.01	0.01
0.01	0.01	0.72	1.36	0.81	0.03	0.33	0.01	0.01	0.02	0.01	0.01	0.30	0.01	0.01	0.01
0.01	0.01	0.01	0.23	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Case 12

0.63	0.47	0.36	0.87	1.46	1.89	1.55	1.58	0.32	0.15	0.10	0.12	0.09	0.11	0.17	0.25
0.28	0.37	0.60	0.86	1.44	1.55	4.08	1.10	0.56	0.21	0.16	0.10	0.17	0.15	0.29	0.70
0.09	0.11	0.22	1.43	2.50	1.50	1.78	1.11	0.32	0.31	0.32	0.17	0.19	0.72	0.63	0.38
0.03	0.04	0.11	0.34	0.43	0.88	0.62	0.98	0.36	0.33	0.70	0.33	0.32	1.04	1.18	0.58
0.02	0.02	0.05	0.15	0.27	0.46	0.55	0.28	0.29	0.51	0.52	1.06	0.39	1.12	1.16	0.37
0.01	0.02	0.03	0.08	0.15	0.32	1.11	0.24	0.15	0.29	1.07	1.34	0.72	0.61	0.21	0.15
0.02	0.02	0.02	0.05	0.21	0.34	0.89	0.24	0.22	0.32	0.45	1.37	2.96	0.47	0.12	0.08
0.02	0.02	0.03	0.05	0.10	0.22	0.32	0.29	0.43	0.34	0.22	0.74	1.91	0.59	0.11	0.06
0.04	0.05	0.06	0.09	0.13	0.22	0.67	0.40	0.81	0.41	0.30	0.69	0.52	0.24	0.12	0.05
0.10	0.09	0.28	0.22	0.35	0.85	0.77	1.18	0.51	0.86	0.31	0.73	0.20	0.12	0.05	0.04
0.15	0.24	0.25	0.84	0.63	0.88	2.42	0.27	0.16	0.11	0.13	0.11	0.09	0.04	0.02	0.02
0.13	0.48	0.61	0.43	0.78	0.78	0.62	0.13	0.05	0.04	0.04	0.06	0.04	0.02	0.01	0.01
0.08	0.11	0.47	0.95	0.55	0.19	0.15	0.05	0.02	0.02	0.02	0.03	0.04	0.02	0.01	0.01
0.06	0.08	0.17	0.32	0.20	0.12	0.07	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01

Case 13

0.47	0.48	0.55	0.94	1.59	2.07	1.85	1.48	0.51	0.25	0.17	0.16	0.17	0.21	0.28	0.34
0.31	0.37	0.59	1.03	1.59	1.91	3.51	1.28	0.53	0.27	0.19	0.17	0.22	0.26	0.37	0.52
0.16	0.19	0.34	1.06	1.91	1.64	1.76	1.11	0.49	0.34	0.29	0.25	0.31	0.57	0.61	0.53
0.08	0.10	0.19	0.42	0.55	0.95	0.83	0.89	0.47	0.38	0.50	0.42	0.48	0.88	0.95	0.67
0.05	0.06	0.10	0.20	0.34	0.54	0.61	0.46	0.40	0.47	0.55	0.89	0.58	0.87	0.87	0.52
0.04	0.04	0.06	0.12	0.22	0.39	0.72	0.38	0.31	0.38	0.83	1.13	0.86	0.67	0.38	0.32
0.04	0.04	0.05	0.09	0.21	0.35	0.59	0.37	0.32	0.38	0.55	1.20	2.24	0.62	0.26	0.19
0.04	0.05	0.06	0.10	0.17	0.31	0.40	0.41	0.42	0.39	0.38	0.78	1.43	0.56	0.20	0.14
0.06	0.07	0.10	0.14	0.23	0.37	0.63	0.49	0.62	0.43	0.39	0.60	0.50	0.28	0.15	0.10
0.11	0.12	0.22	0.27	0.41	0.77	0.79	0.88	0.48	0.54	0.32	0.42	0.22	0.14	0.08	0.07
0.16	0.21	0.28	0.58	0.64	0.86	1.56	0.42	0.27	0.19	0.18	0.14	0.11	0.07	0.05	0.04
0.16	0.30	0.43	0.47	0.67	0.69	0.52	0.23	0.12	0.09	0.08	0.08	0.06	0.04	0.03	0.03
0.13	0.16	0.34	0.60	0.46	0.28	0.21	0.11	0.07	0.05	0.05	0.05	0.05	0.03	0.02	0.02
0.11	0.14	0.23	0.33	0.29	0.20	0.14	0.09	0.06	0.05	0.04	0.04	0.04	0.03	0.02	0.02

Table 4.17 - (End)

COMPLETENESS REGION 1

I_0		1625-1780 C	1625-1780 E	1780-1860 C	1780-1860 E	1860-1910 C	1860-1910 E	1910-1950 C	1910-1950 E	1950-1980 C	1950-1980 E	TOTAL C
4	W	0.0	0.0	13.0	11.7	34.0	32.9	44.0	45.8	43.0	47.7	134.0
	U	0.0	0.0	12.3	10.3	30.2	31.7	38.2	40.0	38.1	41.2	118.8
5	W	0.0	0.0	2.0	3.1	7.0	8.8	18.0	15.6	20.0	12.8	47.0
	U	0.0	0.0	0.9	2.8	7.9	11.2	16.7	14.4	18.5	11.2	44.0
6	W	0.0	0.0	3.0	2.7	3.0	4.1	1.0	4.2	7.0	3.4	14.0
	U	0.0	0.0	3.6	3.4	2.8	4.4	2.4	3.9	6.5	3.1	15.3
7	W	1.0	2.8	1.0	2.2	3.0	1.5	1.0	1.2	1.0	0.9	7.0
	U	1.1	2.9	1.3	2.0	3.2	1.4	1.2	1.1	1.2	0.8	8.0
8	W	1.0	1.3	0.0	0.7	0.0	0.4	1.0	0.3	0.0	0.3	2.0
	U	0.9	1.2	0.0	0.6	0.0	0.4	0.9	0.3	0.0	0.2	1.8

COMPLETENESS REGION 2

I_0		1625-1780 C	1625-1780 E	1780-1860 C	1780-1860 E	1860-1910 C	1860-1910 E	1910-1950 C	1910-1950 E	1950-1980 C	1950-1980 E	TOTAL C
4	W	22.0	19.3	27.0	26.3	36.0	37.5	38.0	38.6	26.0	32.2	149.0
	U	19.1	17.8	22.5	22.9	30.9	32.6	34.0	35.1	24.9	31.2	131.4
5	W	2.0	4.9	5.0	6.7	11.0	11.9	11.0	10.6	21.0	8.2	50.0
	U	3.2	4.7	5.5	7.0	12.9	12.6	11.8	10.7	19.7	8.3	53.1
6	W	2.0	3.2	2.0	4.1	2.0	3.2	3.0	2.8	4.0	2.1	13.0
	U	1.6	2.8	3.4	4.7	2.6	3.4	2.6	2.9	4.5	2.2	14.7
7	W	2.0	2.6	1.0	1.4	2.0	0.9	1.0	0.7	0.0	0.6	6.0
	U	1.9	2.7	1.2	1.5	1.9	1.0	1.3	0.8	0.0	0.6	6.3
8	W	1.0	0.8	0.0	0.4	0.0	0.3	0.0	0.2	0.0	0.2	1.0
	U	1.4	0.9	0.0	0.4	0.0	0.3	0.0	0.2	0.0	0.2	1.4

W : without considering uncertainty on earthquakes size
U : considering uncertainty on earthquake size

C : earthquake count (for U, a-posteriori)
E : expected earthquake count

Table 4.18 - Observed and expected count for the reference case and the case without uncertainty on earthquake size

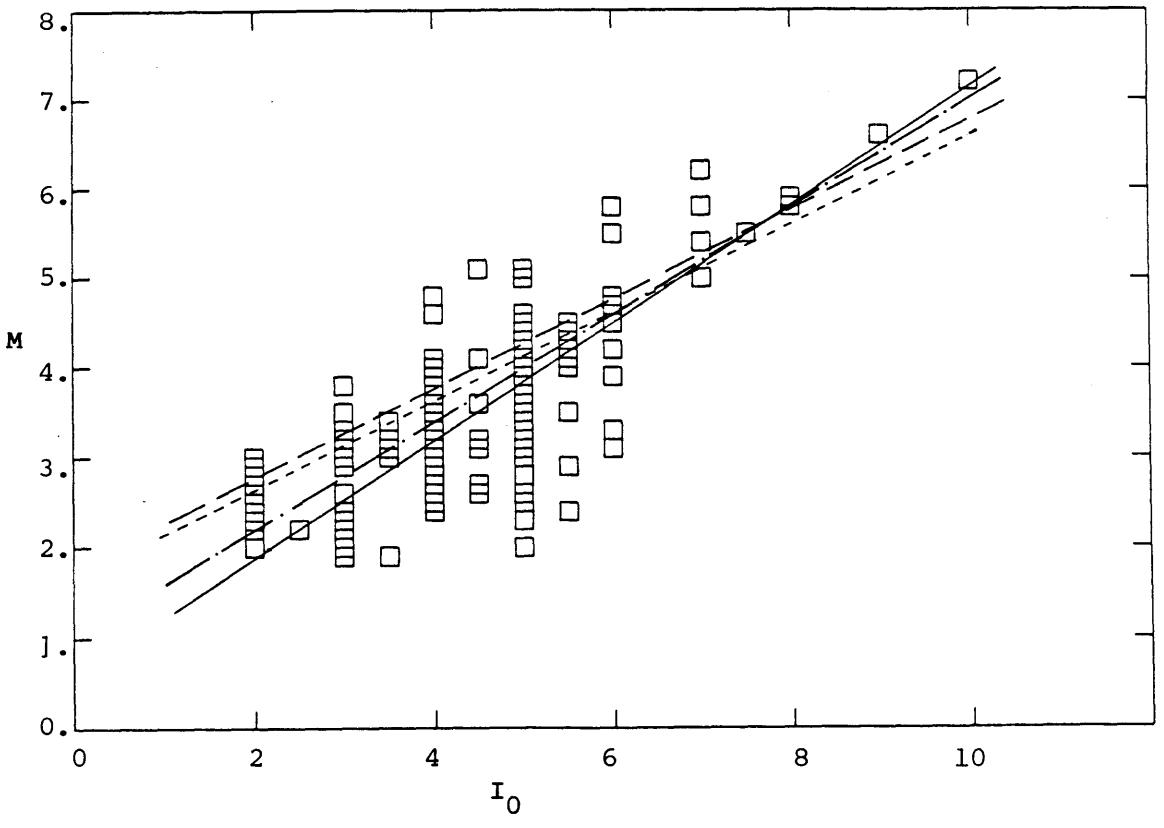
a. Empirical bootstrapping results

average exp(a)	7.006	8.536	20.214	23.783	2.563	1.565	0.839	5.725
	0.134	7.439	16.719	15.208	2.950	6.788	7.761	9.584
	0.091	1.505	2.236	7.368	4.910	19.619	11.157	7.053
	0.398	0.091	3.080	6.116	7.865	13.372	25.727	0.156
	0.640	2.221	3.139	9.885	14.547	10.346	3.951	1.031
	3.187	4.451	5.878	8.771	0.091	0.570	0.091	0.091
	0.091	4.853	2.082	1.260	0.107	0.091	1.084	0.091
average b	1.341	1.343	1.311	1.202	1.238	1.302	1.305	1.314
	1.319	1.308	1.303	1.268	1.314	1.355	1.320	1.368
	1.288	1.273	1.303	1.330	1.398	1.413	1.438	1.356
	1.253	1.260	1.266	1.359	1.456	1.463	1.310	1.335
	1.205	1.178	1.229	1.272	1.414	1.410	1.327	1.288
	1.119	1.131	1.100	1.151	1.288	1.322	1.295	1.275
	1.117	1.062	1.107	1.170	1.241	1.277	1.254	1.269
standard deviation exp(a)	2.584	2.737	3.947	4.498	1.369	0.987	0.821	2.308
	0.064	2.952	4.433	4.056	1.571	2.348	2.451	2.459
	0.000	1.095	1.424	2.572	2.409	3.869	2.261	2.299
	0.295	0.000	1.320	2.694	2.831	3.107	3.800	0.096
	0.665	1.168	1.780	2.752	2.860	2.710	1.575	0.872
	1.460	1.694	1.675	2.435	0.000	0.451	0.000	0.000
	0.000	1.571	0.953	0.808	0.028	0.000	0.947	0.000
standard deviation b	0.061	0.073	0.059	0.109	0.069	0.041	0.052	0.062
	0.048	0.057	0.073	0.066	0.051	0.036	0.059	0.053
	0.031	0.040	0.039	0.068	0.032	0.057	0.040	0.052
	0.034	0.032	0.042	0.033	0.034	0.041	0.081	0.042
	0.040	0.064	0.042	0.049	0.042	0.033	0.034	0.036
	0.061	0.059	0.061	0.074	0.029	0.025	0.029	0.031
	0.059	0.077	0.052	0.044	0.032	0.026	0.054	0.037

b. Parametric bootstrapping results

average exp(a)	6.951	8.410	17.546	20.417	3.064	1.623	1.028	5.394
	0.186	6.945	16.301	13.839	2.408	6.692	7.521	8.808
	0.190	1.263	2.068	6.840	4.449	17.910	9.990	6.245
	0.506	0.140	3.024	5.135	6.952	11.672	24.044	0.233
	0.642	2.225	2.822	9.678	13.377	9.424	3.742	1.459
	3.297	3.869	5.939	8.194	0.219	0.514	0.202	0.161
	0.122	3.877	1.551	0.687	0.118	0.139	0.840	0.204
average b	1.355	1.337	1.296	1.260	1.300	1.336	1.360	1.370
	1.341	1.340	1.323	1.303	1.324	1.355	1.379	1.377
	1.305	1.309	1.318	1.344	1.374	1.411	1.420	1.390
	1.251	1.259	1.284	1.335	1.401	1.445	1.392	1.406
	1.184	1.197	1.227	1.284	1.384	1.412	1.421	1.409
	1.128	1.130	1.167	1.230	1.321	1.381	1.409	1.413
	1.112	1.112	1.138	1.218	1.297	1.362	1.398	1.409
standard deviation exp(a)	2.716	2.810	5.366	4.961	1.655	1.301	1.060	2.948
	0.288	2.344	4.069	4.220	1.718	2.576	3.079	3.063
	0.269	1.021	1.076	3.023	2.080	3.629	2.830	2.769
	0.641	0.196	1.768	1.897	2.280	2.882	3.625	0.327
	0.777	1.247	1.830	2.404	2.962	2.482	1.717	1.037
	1.504	1.480	1.885	2.473	0.296	0.580	0.398	0.239
	0.106	1.475	0.806	0.689	0.135	0.193	0.767	0.461
standard deviation b	0.175	0.166	0.126	0.121	0.126	0.144	0.177	0.185
	0.155	0.145	0.121	0.111	0.110	0.129	0.149	0.165
	0.129	0.120	0.102	0.102	0.094	0.107	0.104	0.123
	0.132	0.123	0.114	0.106	0.105	0.102	0.117	0.114
	0.152	0.143	0.131	0.133	0.110	0.128	0.141	0.145
	0.180	0.168	0.158	0.156	0.132	0.145	0.163	0.167
	0.185	0.187	0.177	0.156	0.147	0.151	0.169	0.173

Table 4.19 - Sample statistics of bootstrapping for Model D



——— 0.44+0.67 I_0 (WGC, 1982)

——·—— 1.00+0.60 I_0 (Chiburis, 1981)

— — — 1.75+0.50 I_0 (Nuttli, 1974)

------- 1.66+0.49 I_0 (Street and Turcotte, 1977)

□ data in Chiburis (1981) catalog

Figure 2.1 - Comparison of proposed relationships between magnitude M and Modified Mercalli Intensity I_0 and the data in the Chiburis catalog

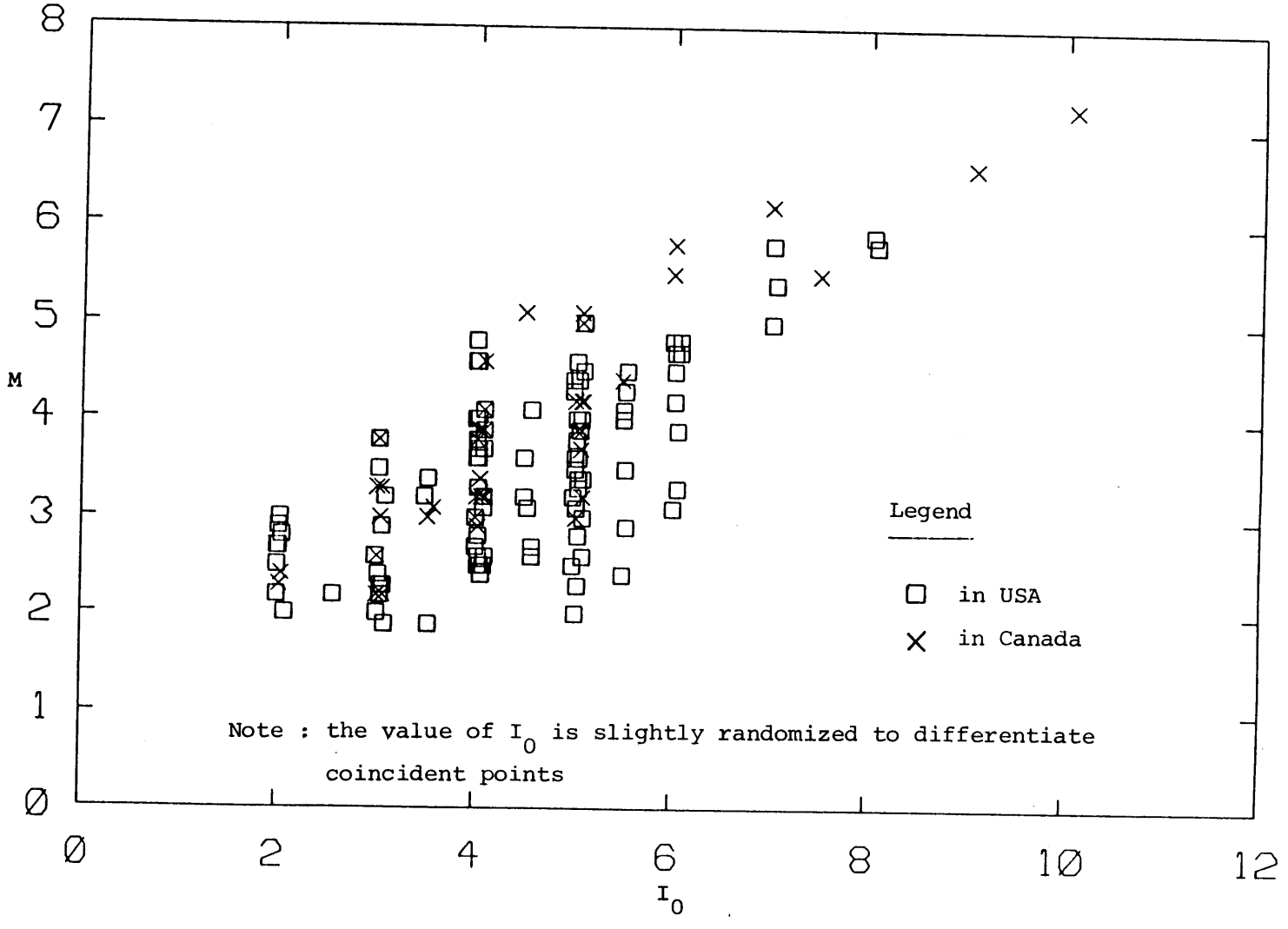
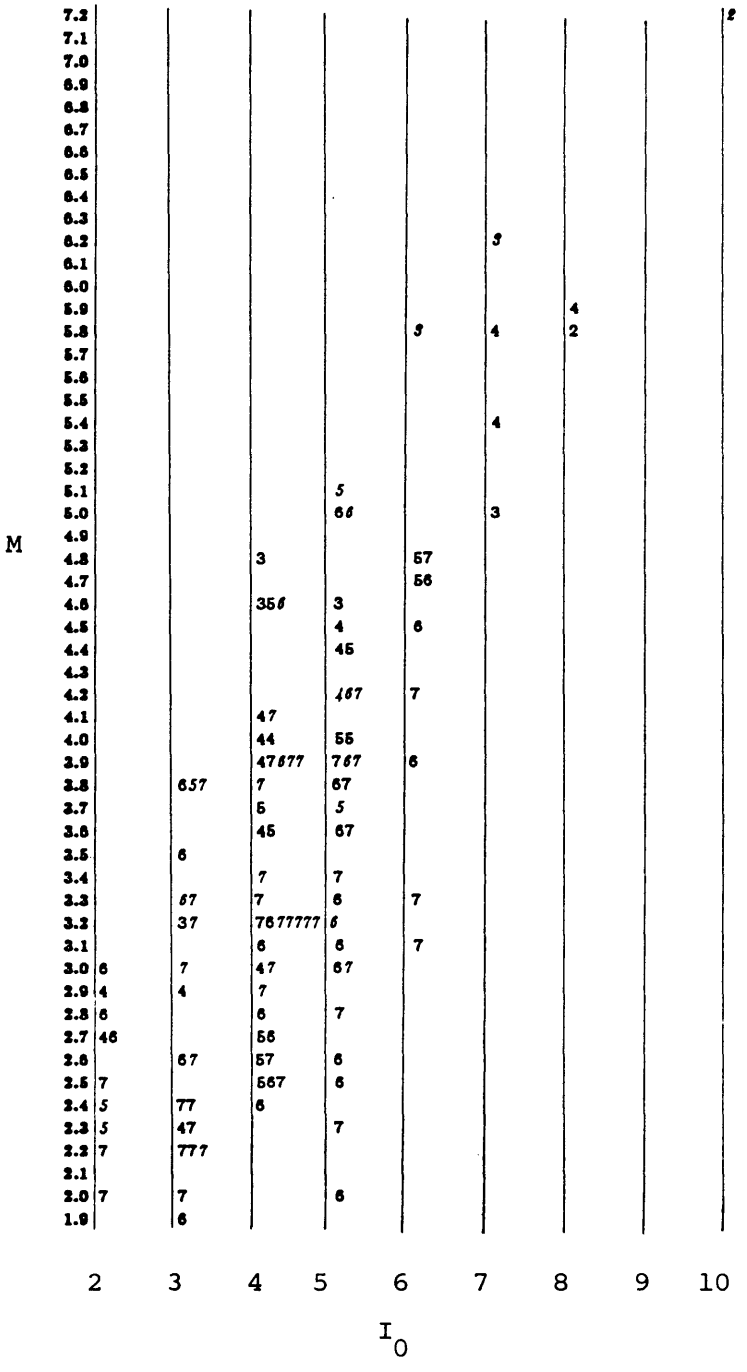


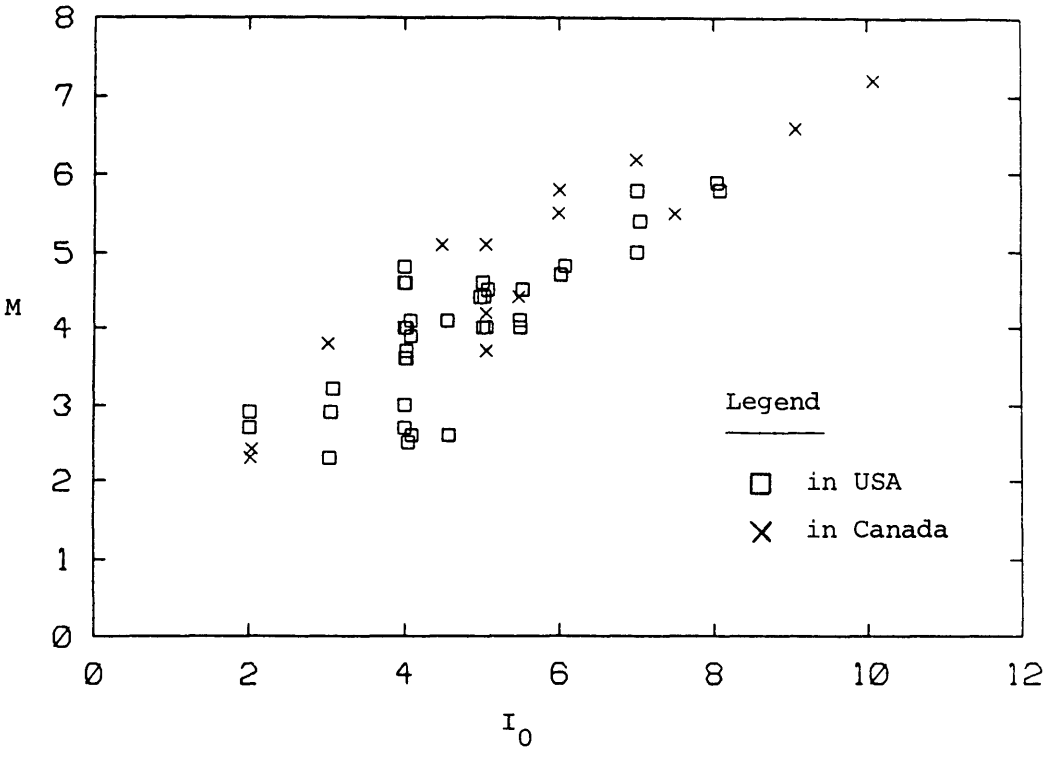
Figure 2.2 - I_0 versus M in the Chiburis (1981) catalog



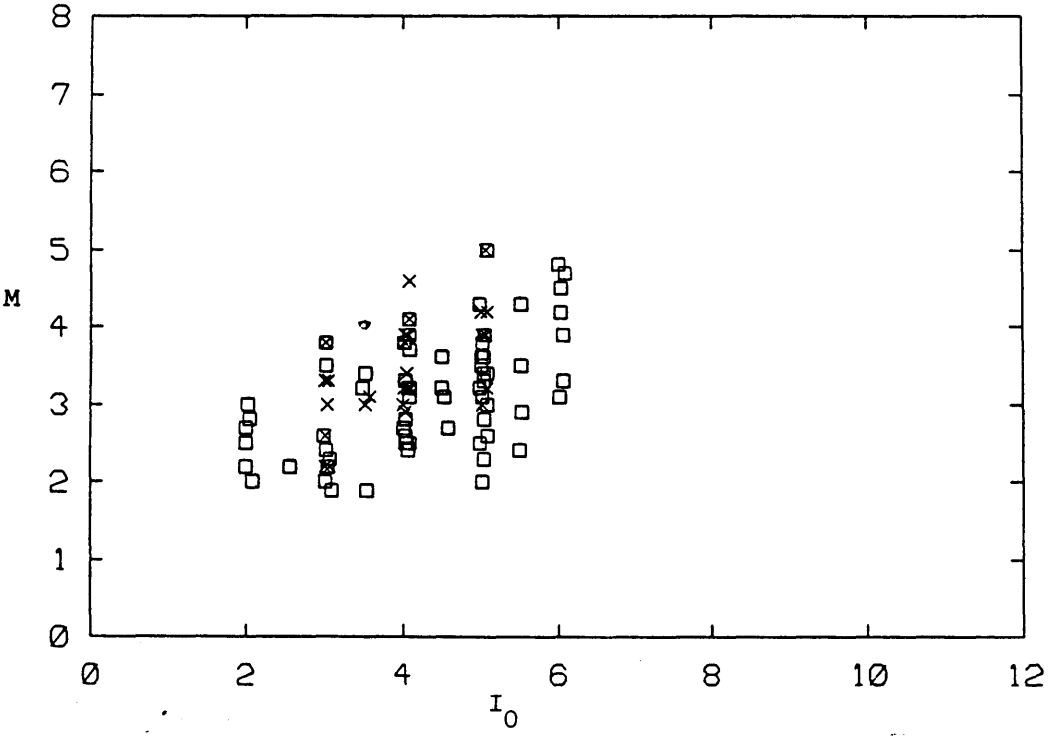
Note : - the value associated with each datapoint indicates the decade since 1900 when the earthquake occurred, e.g. '2' corresponds to 1920-1929

- italics indicate Canadian earthquakes

Figure 2.3 - Value of M, time of occurrence and geographical location versus I_0 for the Chiburis data with accurate estimates of I_0



a. Prior to 1960



b. Since 1960

Note : to differentiate coincident points, the values of I_0 are slightly randomized

Figure 2.4 - M versus I_0 prior to, and since 1960 in the Chiburis catalog

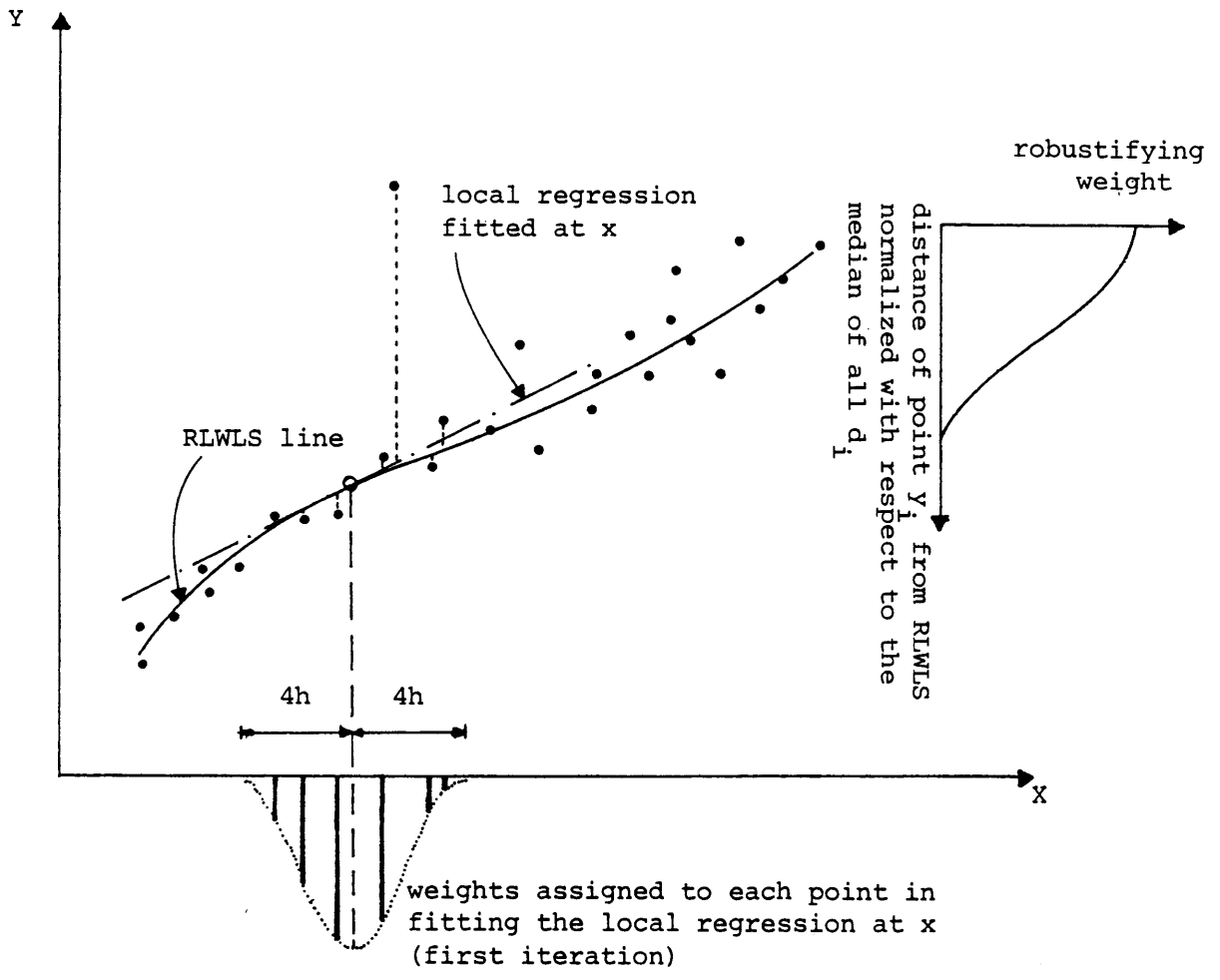
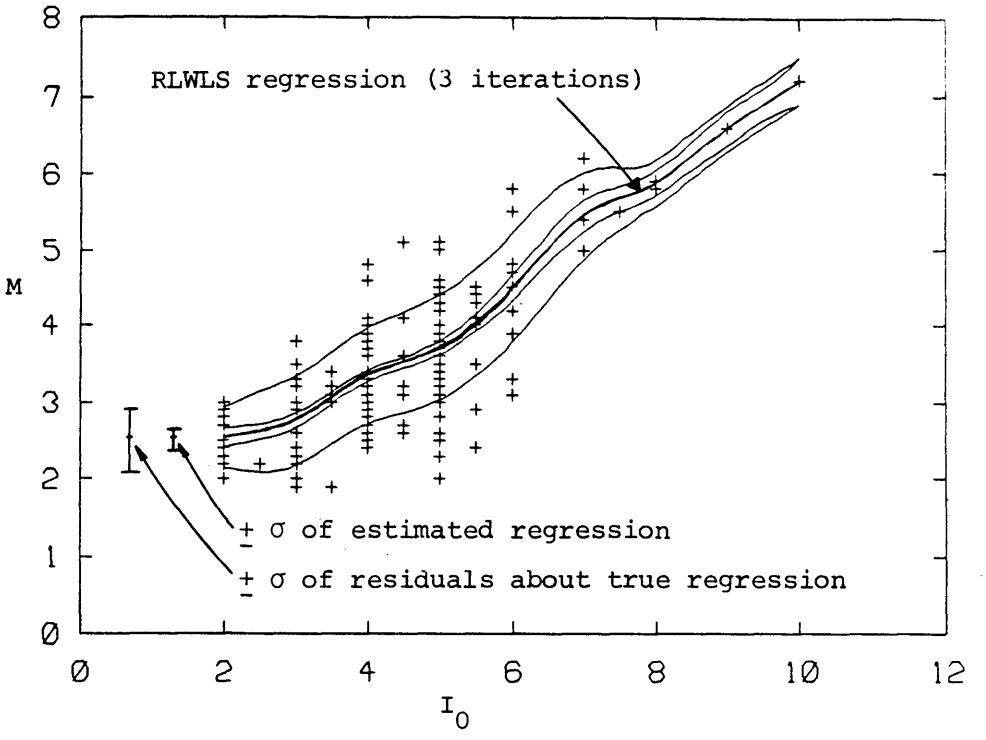
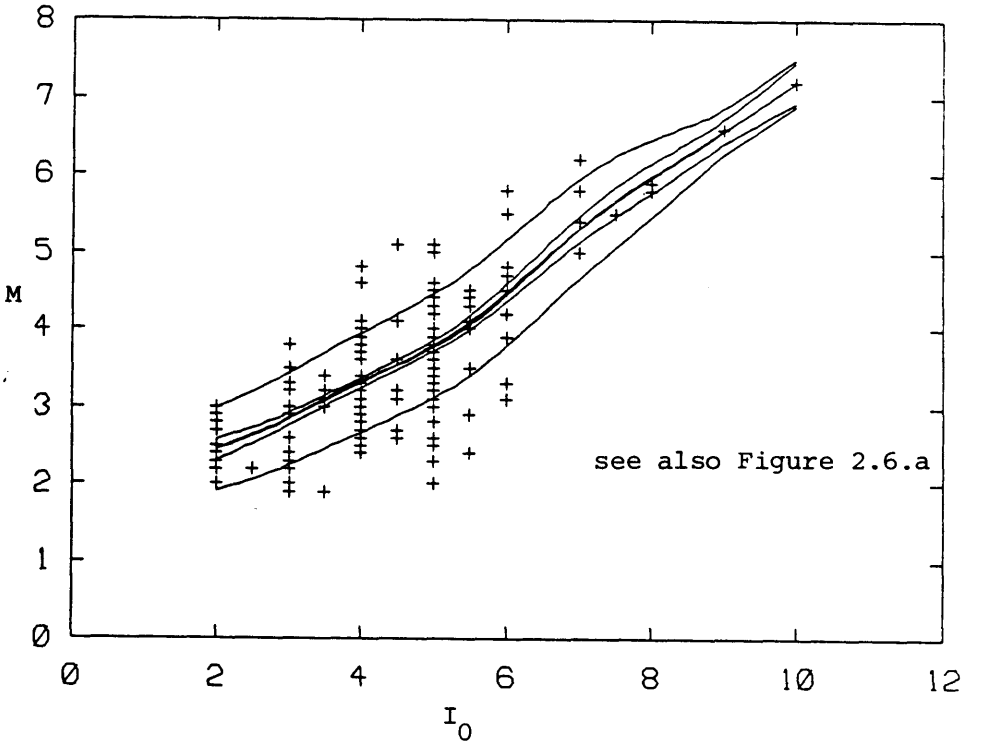


Figure 2.5 - Illustration of the robust locally-weighted least-squares method (RLWLS)

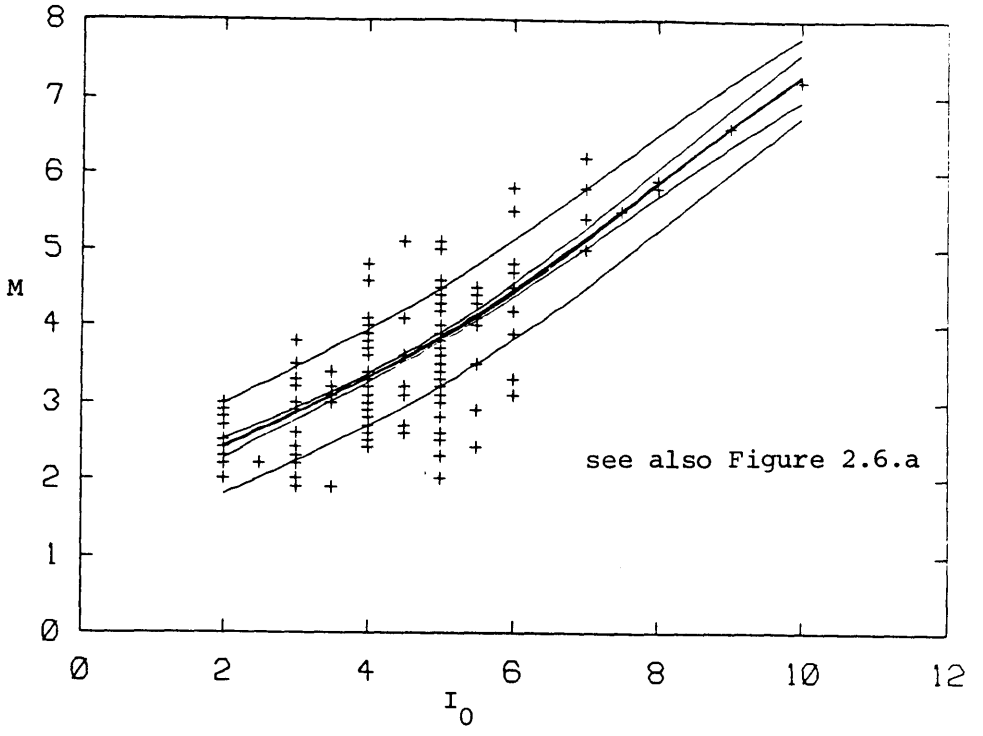


a. h = 0.5

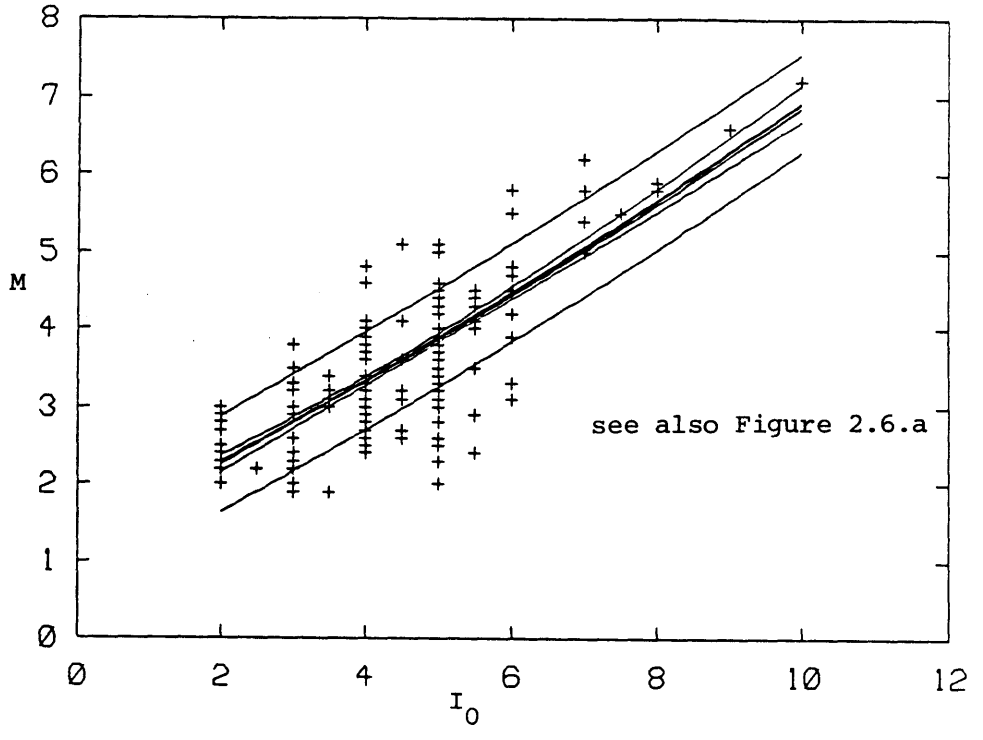


b. h = 1.

Figure 2.6 - Application of RLWLS to the estimation of the regression of M against I_0 for the Chiburis data

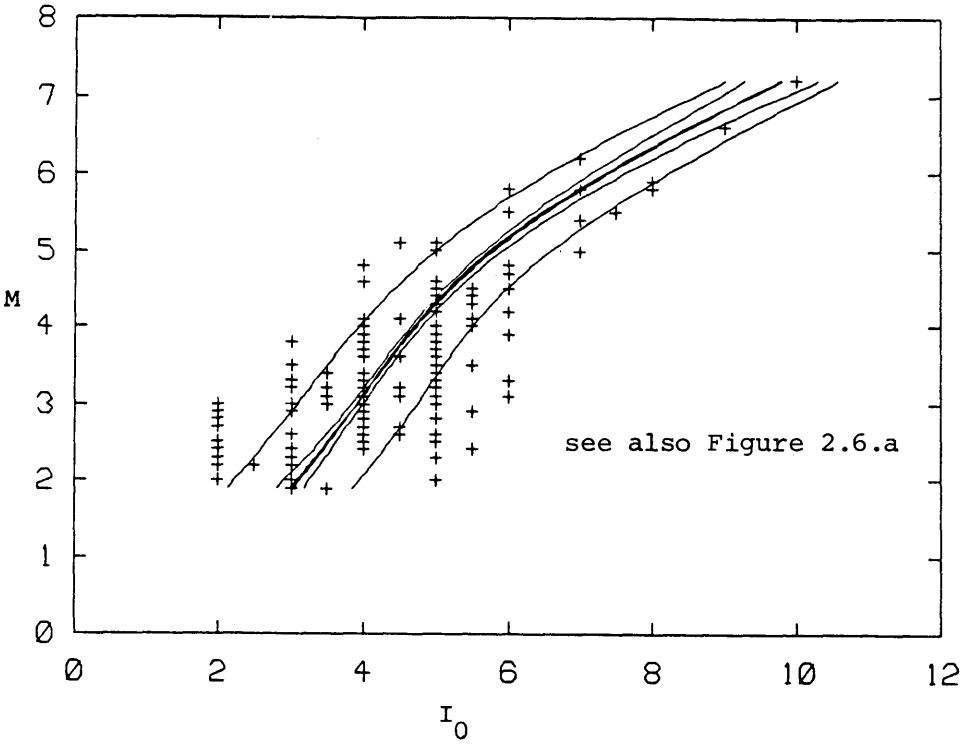


c. h = 2.

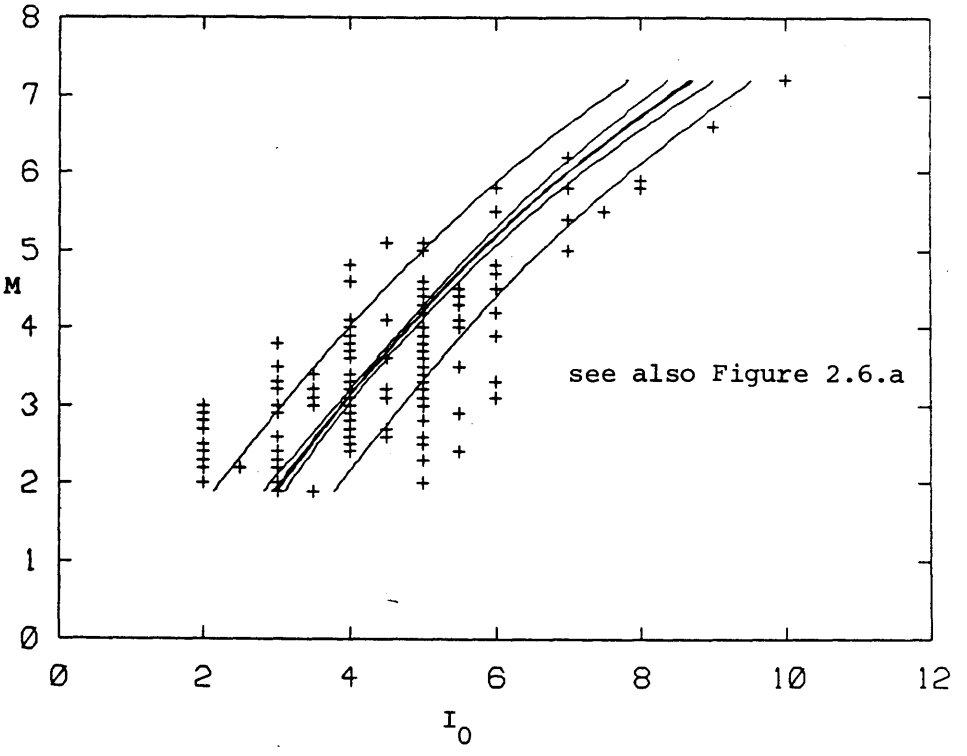


d. h = 5.

Figure 2.6 - (End)

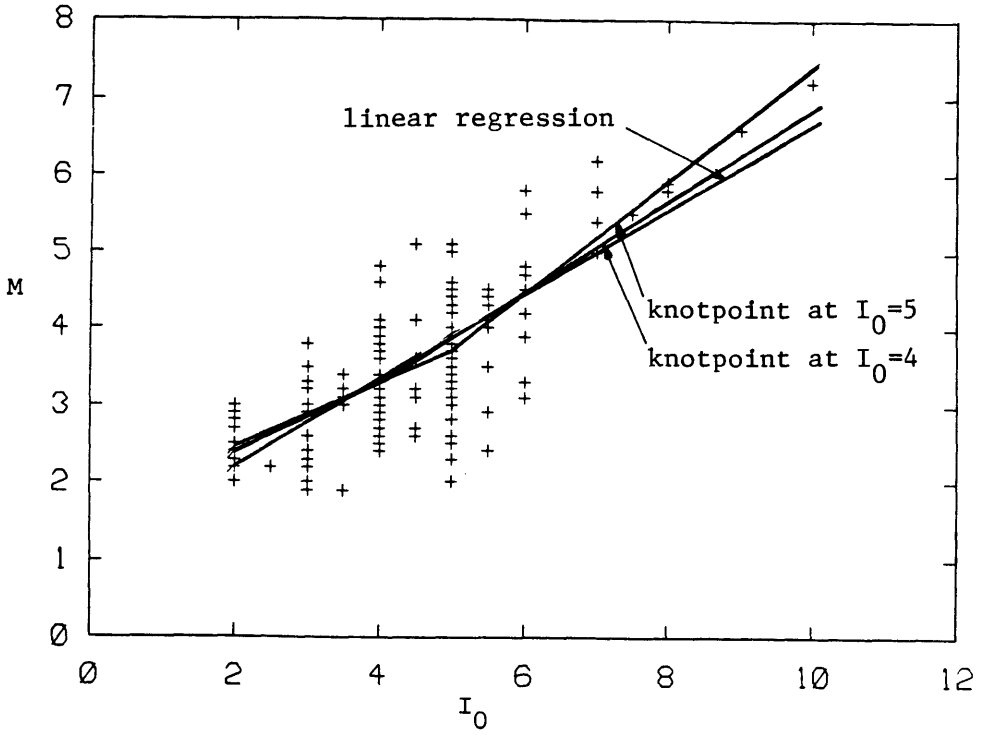


a. h = 1.



b. h = 2.

Figure 2.7 - Application of RLWLS method to the estimation of the regression of I_0 versus M for the Chiburis data



a. Comparison of 3 fitted linear splines of M on I_0
(Data from Chiburis, 1981)

FROM TO	YSPL	BETA	RMSE	UNCA	UNCB	CORAB	NRES	SRES	SSRES	ALPHA
[2.0-10.]	1.11	0.55	0.69	0.19	0.041	-0.956	151	0.00	71.81	-
[2.0-5.0]	1.64	0.42	0.67	0.26	0.064	-0.975	123	1.20	54.92	-
(5.0-10.]	-0.02	0.75	0.70	0.45	0.081	-0.986	28	-1.20	13.23	0.005
[2.0-4.0]	1.55	0.43	0.61	0.33	0.093	-0.982	80	5.14	30.03	-
(4.0-10.]	0.87	0.60	0.76	0.29	0.060	-0.976	71	-7.68	40.70	0.182

Notation

YSPL : intercept

BETA : slope

RMSE : root-mean-square of residuals

UNCA : standard deviation of YSPL

UNCB : standard deviation of BETA

CORAB : correlation of BETA and YSPL

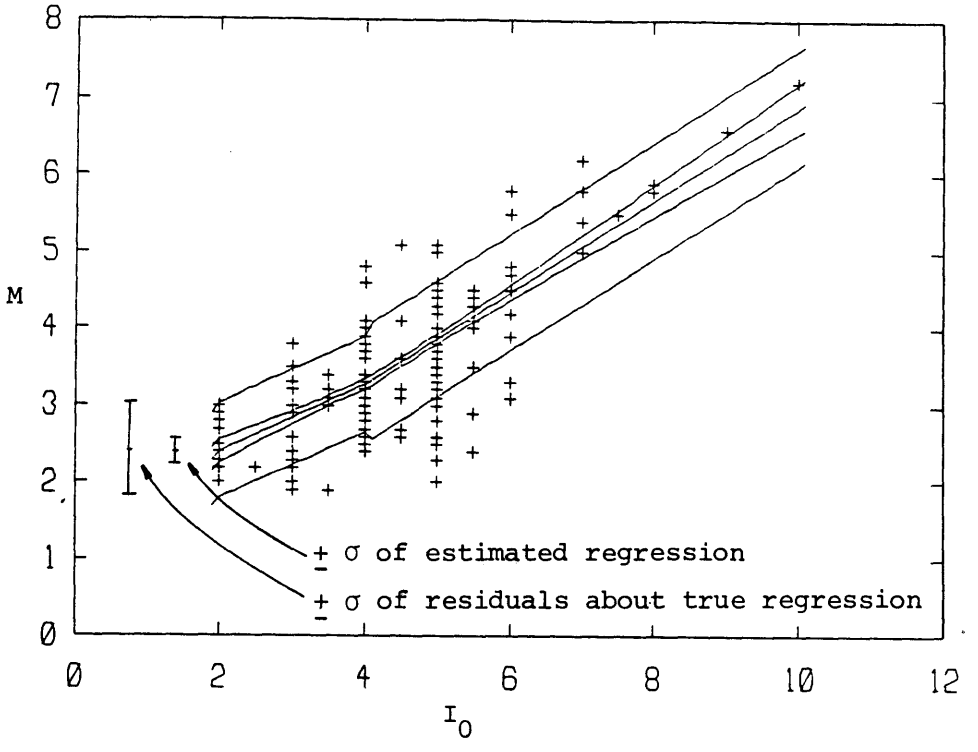
NRES : number of residuals

SRES : sum of residuals

SSRES : sum of squares of residuals

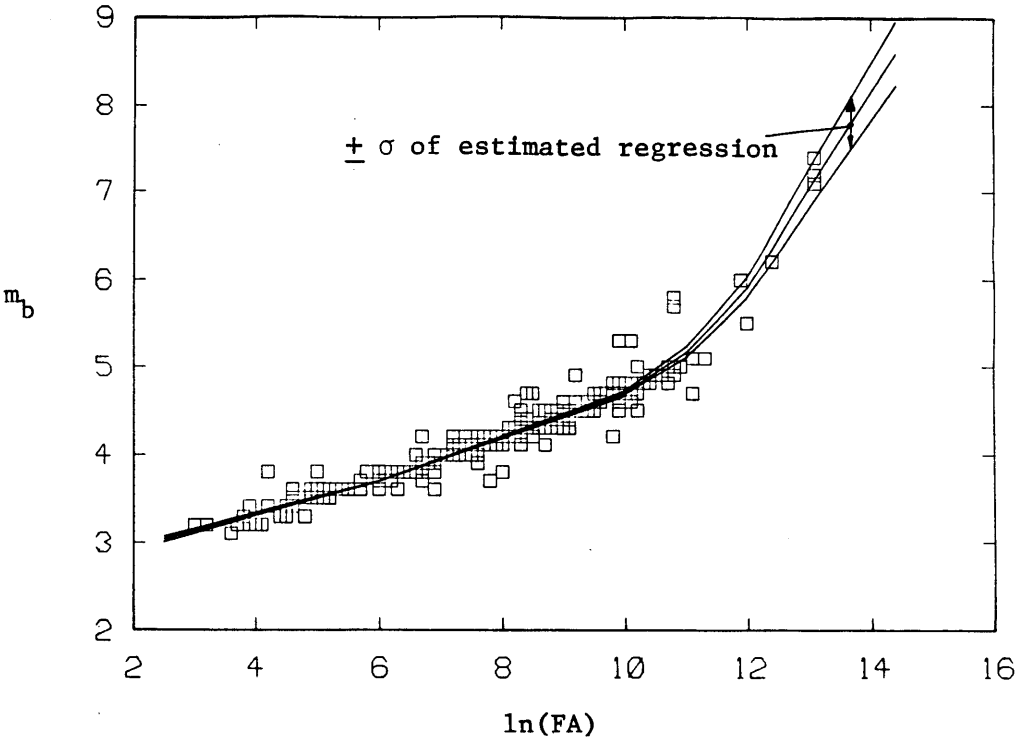
ALPHA : probability that an equal or larger change of slope is due to random error

Figure 2.8 - Illustration of linear spline regression



b. Linear spline with knot-point at I_0=4
(Chiburis data, 1981)

Figure 2.8 - (Continued)



c. Illustration of linear spline regression of bodywave magnitude m_b on the logarithm of felt area ln(FA). Data from Epri (1985)

Figure 2.8 - (End)

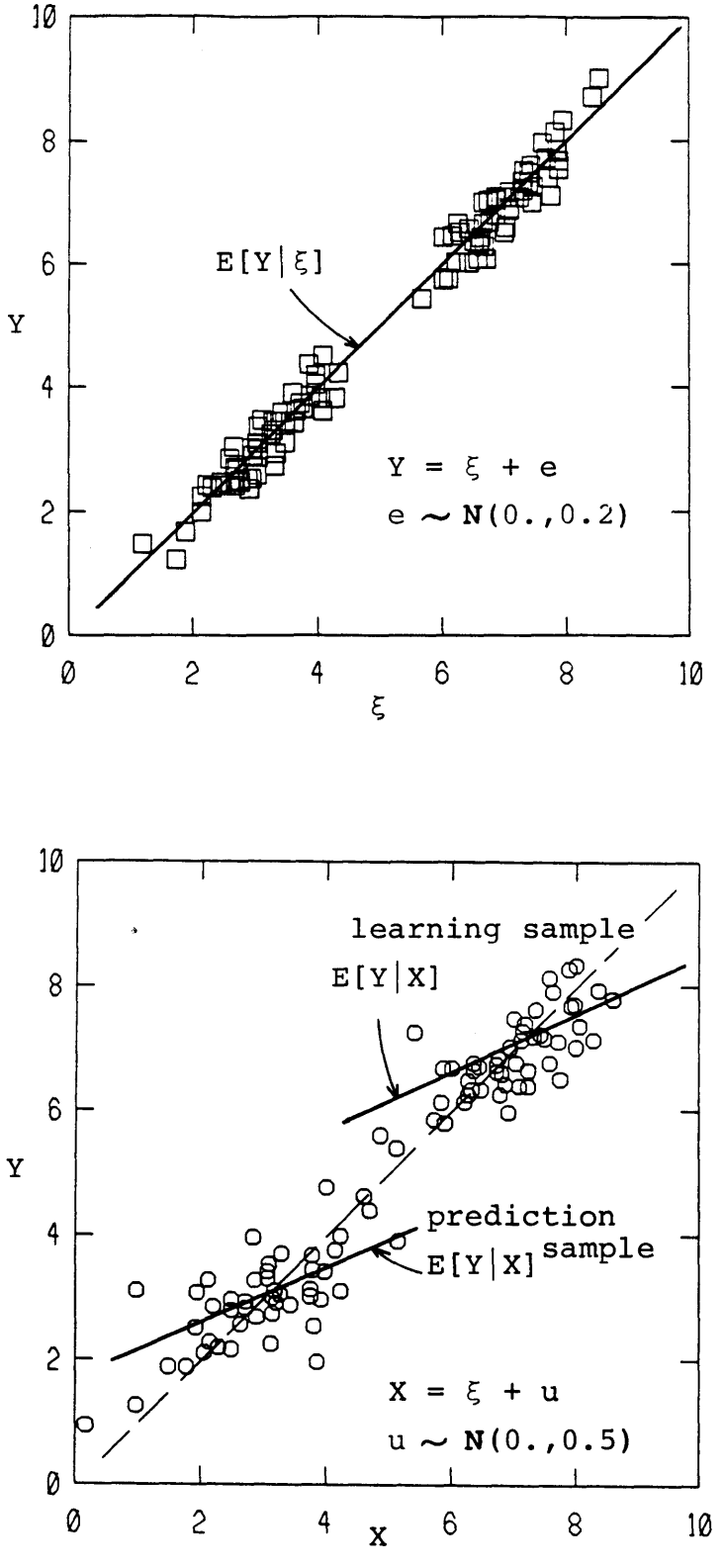
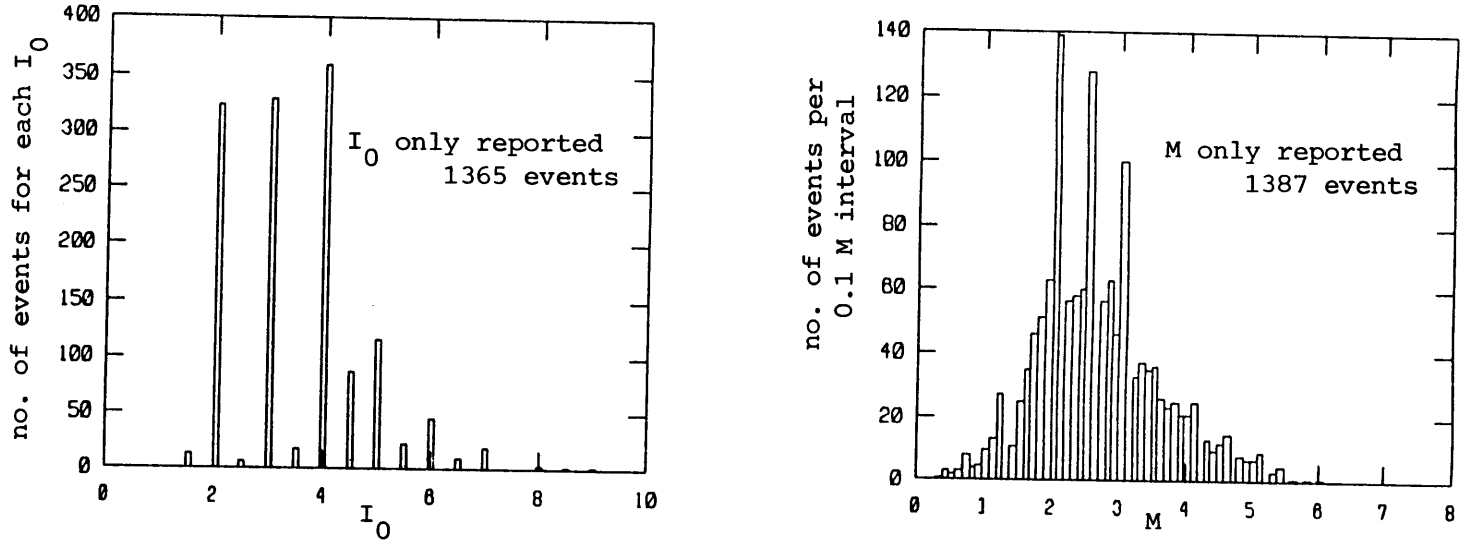
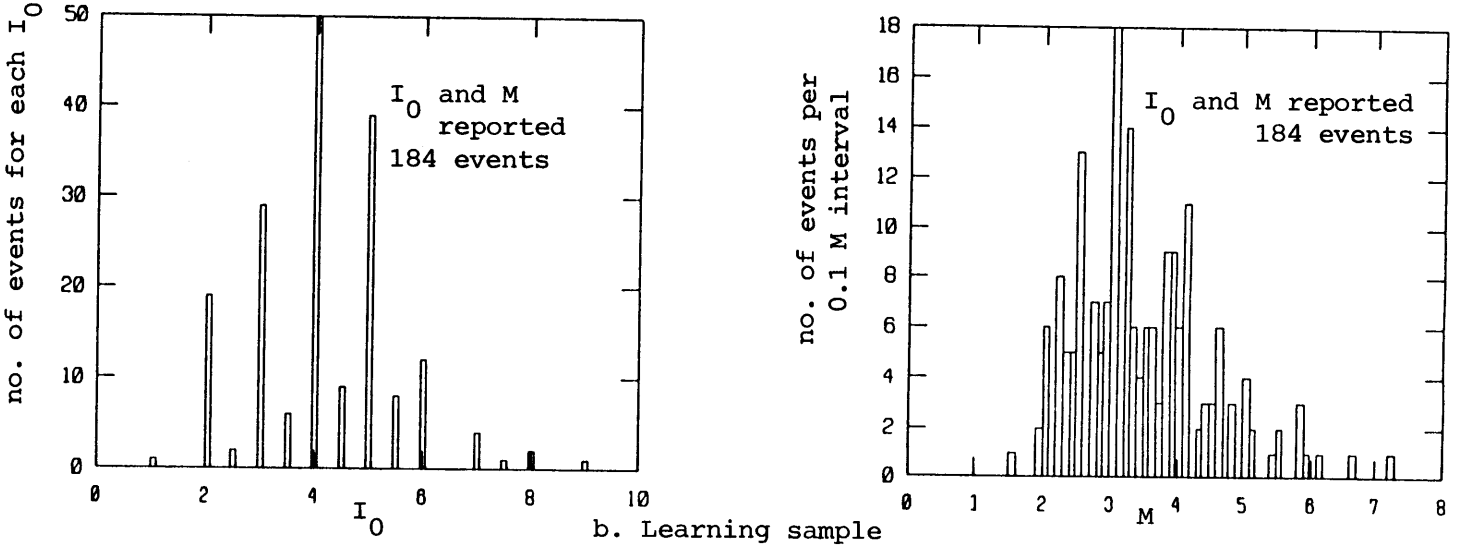


Figure 2.9 - Illustration of the effect of the marginal distribution of X on the regression estimate when X is subject to estimation error

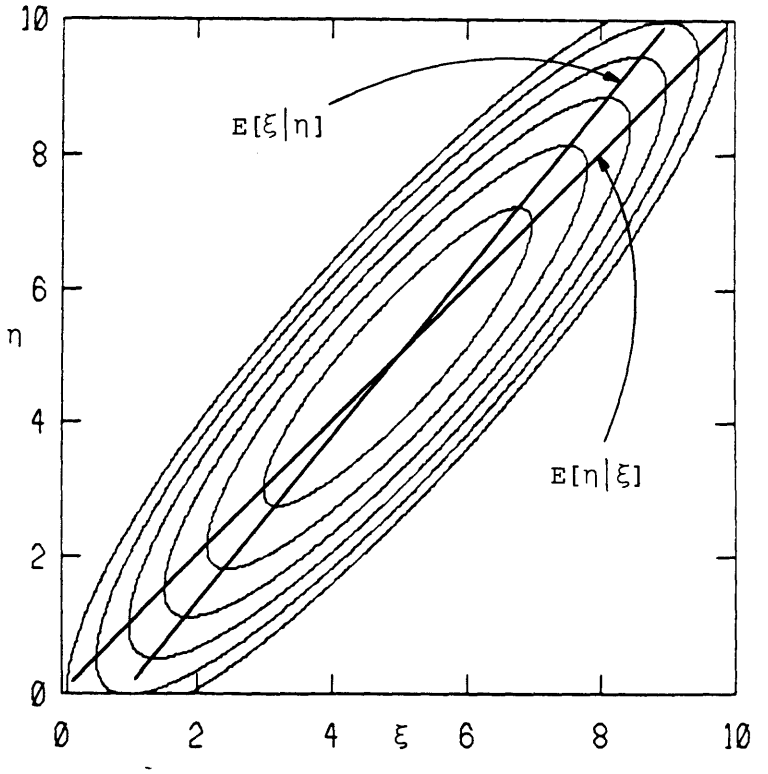


a. Prediction sample

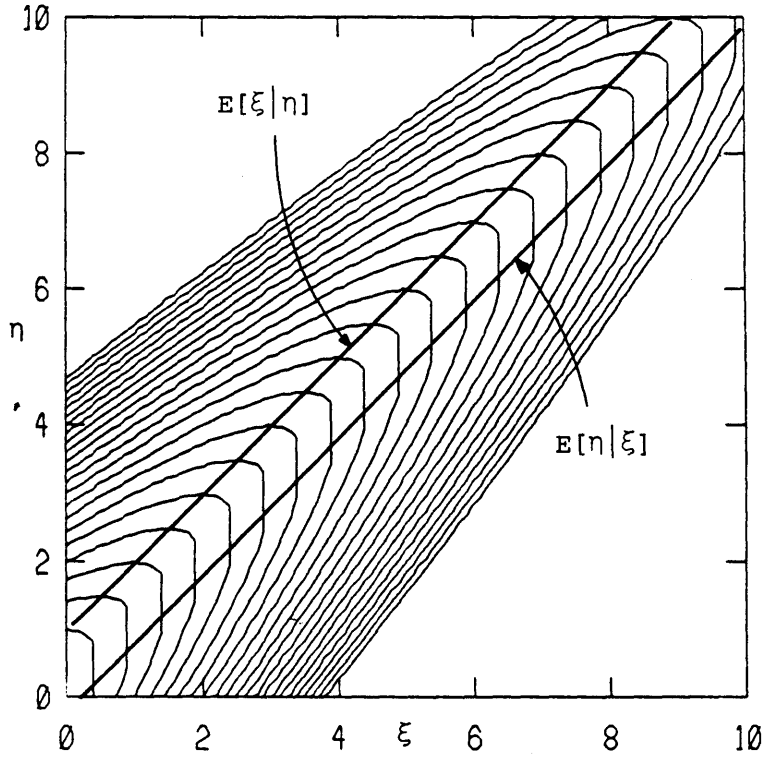


b. Learning sample

Figure 2.10 - Histogram of I_0 and M in the prediction and learning sample of the updated Chiburis catalog



a. Contourlines for a bivariate normal distribution



b. Contourlines for a marginally-exponential, conditionally normal distribution

Figure 2.11 - Illustration of a bivariate normal and a marginally-exponential, conditionally-normal distribution

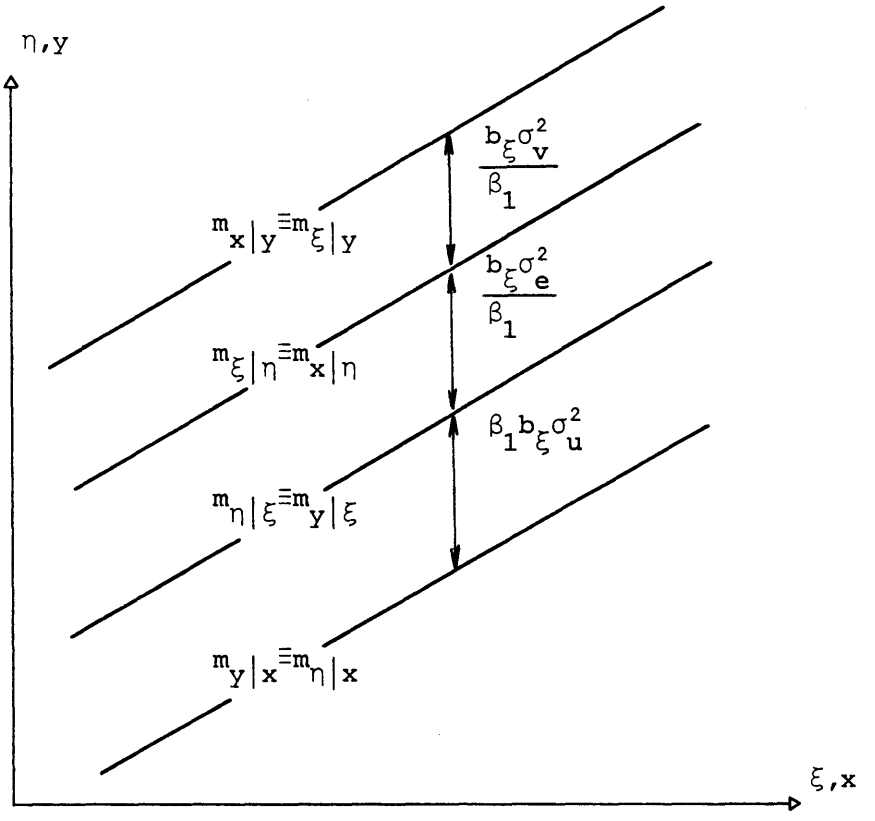


Figure 2.12 - Various regressions for the marginally-exponential, conditionally-normal bivariate distribution

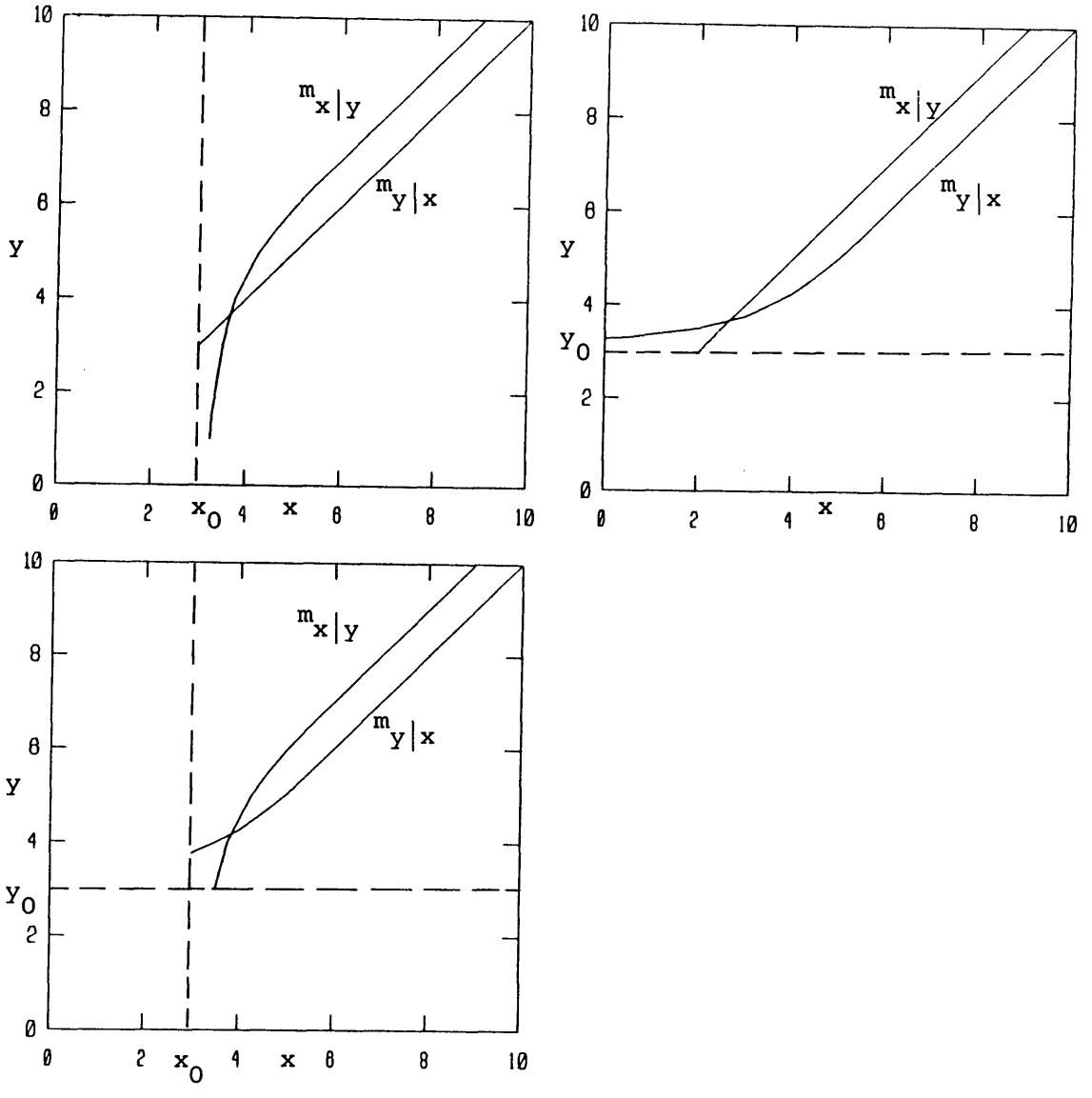


Figure 2.13 - Effect of truncation of the marginally-exponential conditionally-normal distribution on the regressions

m^*_2	M	I_0 $\Delta I_0=0$	I_0 $\Delta I_0=1$	I_0 $\Delta I_0=2$
0.3	1			
0.4	3			
0.5	2			
0.6	3			
0.7	8			
0.8	4			
0.9	5			
1.0	10			
1.1	13			
1.2	14			
1.3	13			
1.4	11			
1.5	25			
1.6	35			
1.7	46			
1.8	51			
1.9	63			
2.0	59		13	
2.1	80			
2.2	56			
2.3	58			
2.4	60	324		
2.5	64			
2.6	64		7	
2.7	56			4
2.8	63			
2.9	46			
3.0	56	325		
3.1	44			
3.2	33		19	
3.3	37			2
3.4	35			
3.5	36			
3.6	26	357		
3.7	23			
3.8	25		87	
3.9	21			4
4.0	21			
4.1	15			
4.2	10	112		
4.3	13			
4.4	10		23	
4.5	12			4
4.6	10			
4.7	5			
4.8	8	42		
4.9	7			
5.0	7		10	
5.1	6			2
5.2	3			
5.3	3			
5.4	5	18		
5.5				
5.6			1	
5.7	1			1
5.8	1			
5.9				
6.0	1	3		
6.1				
6.2			3	
6.3				
6.4				
6.5				
6.6		2		
6.7				
6.8			1	
6.9				
7.0				
7.1				
7.2		1		

Figure 2.14 - Number of datapoints in each magnitude interval after conversion

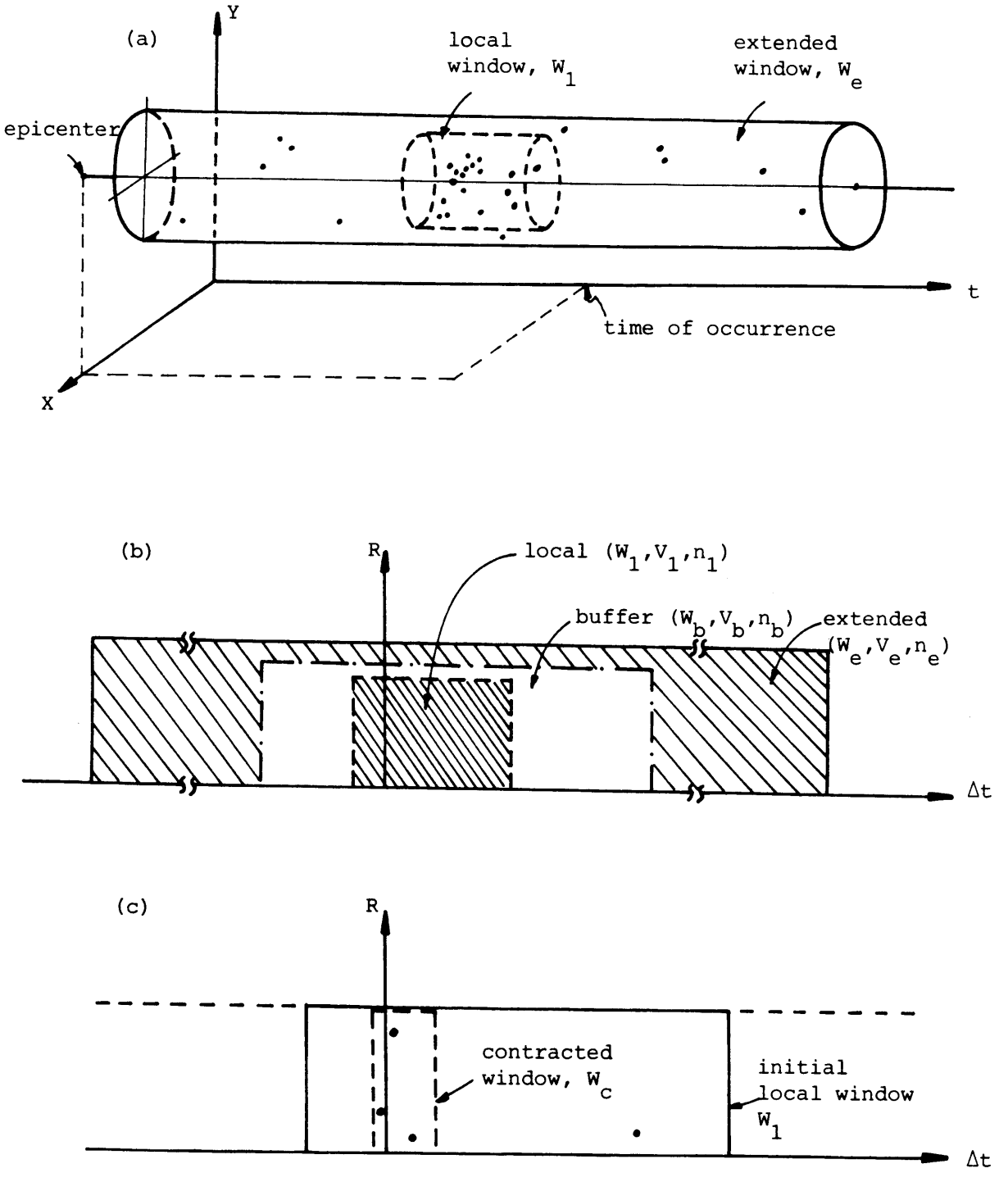


Fig. 3.1 - Windows used in Sec. 3.3.2 for the test of clustering: (a) local and extended windows in 3D, (b) buffer window, and (c) contracted window.

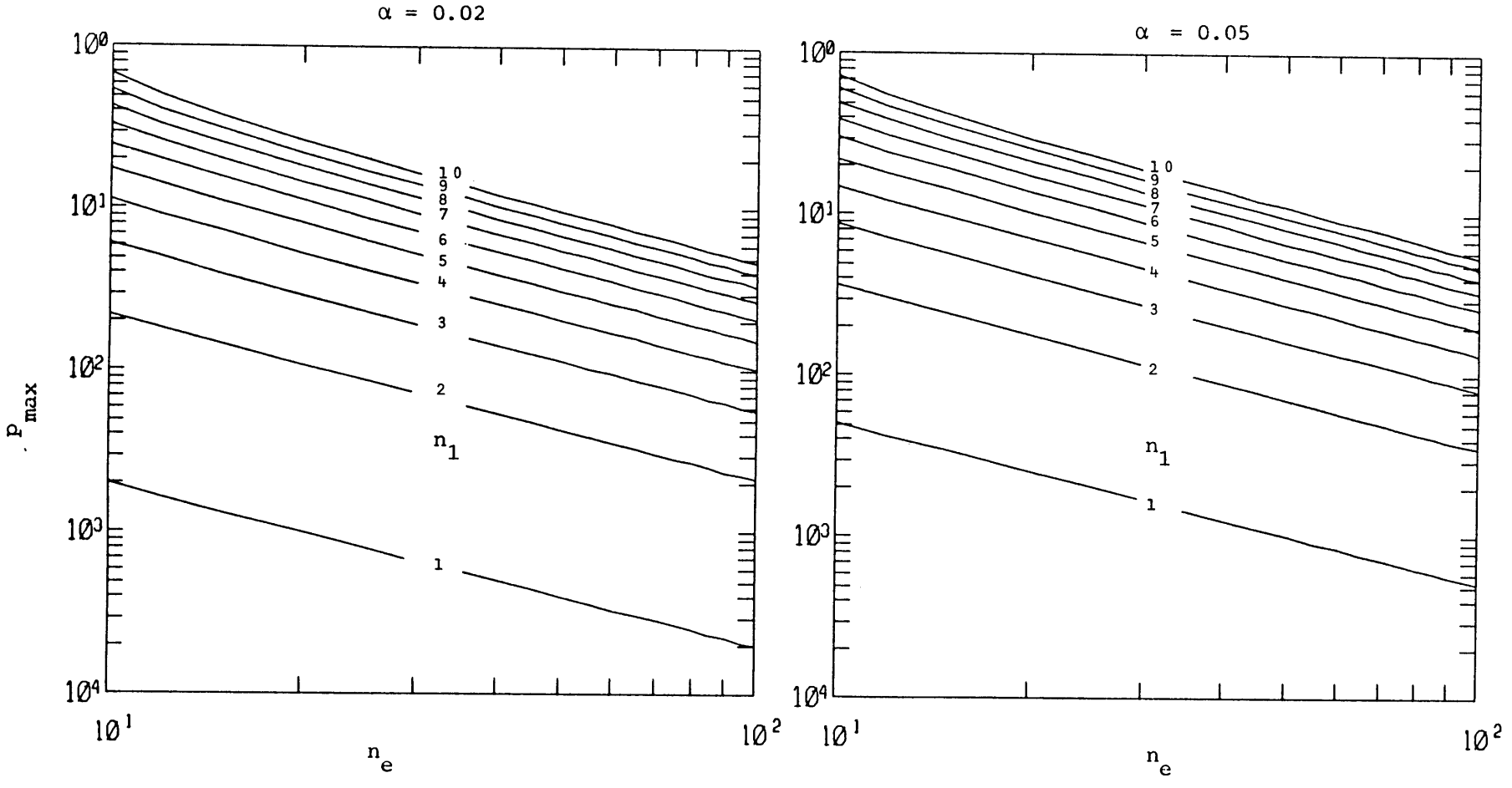


Fig. 3.2 - Maximum value of p for which clustering is detected as derived from Eqs. 3.5 and 3.6.

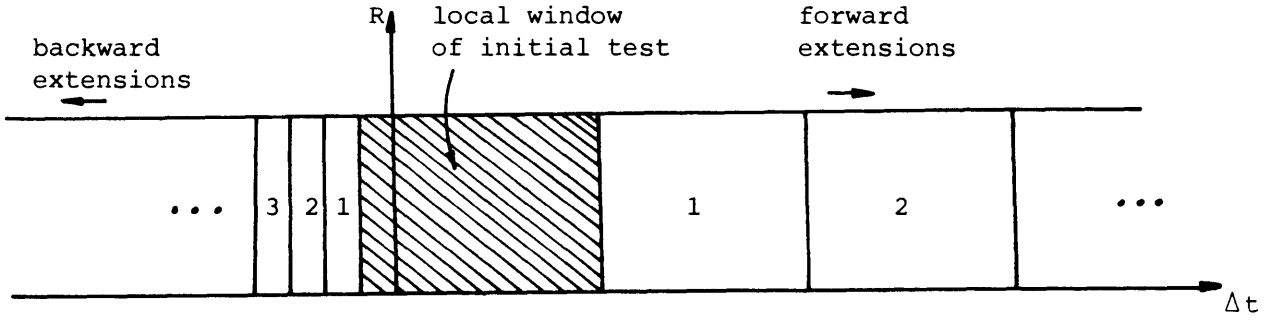


Fig. 3.3 - Estimation of cluster region in the one-dimensional scheme.

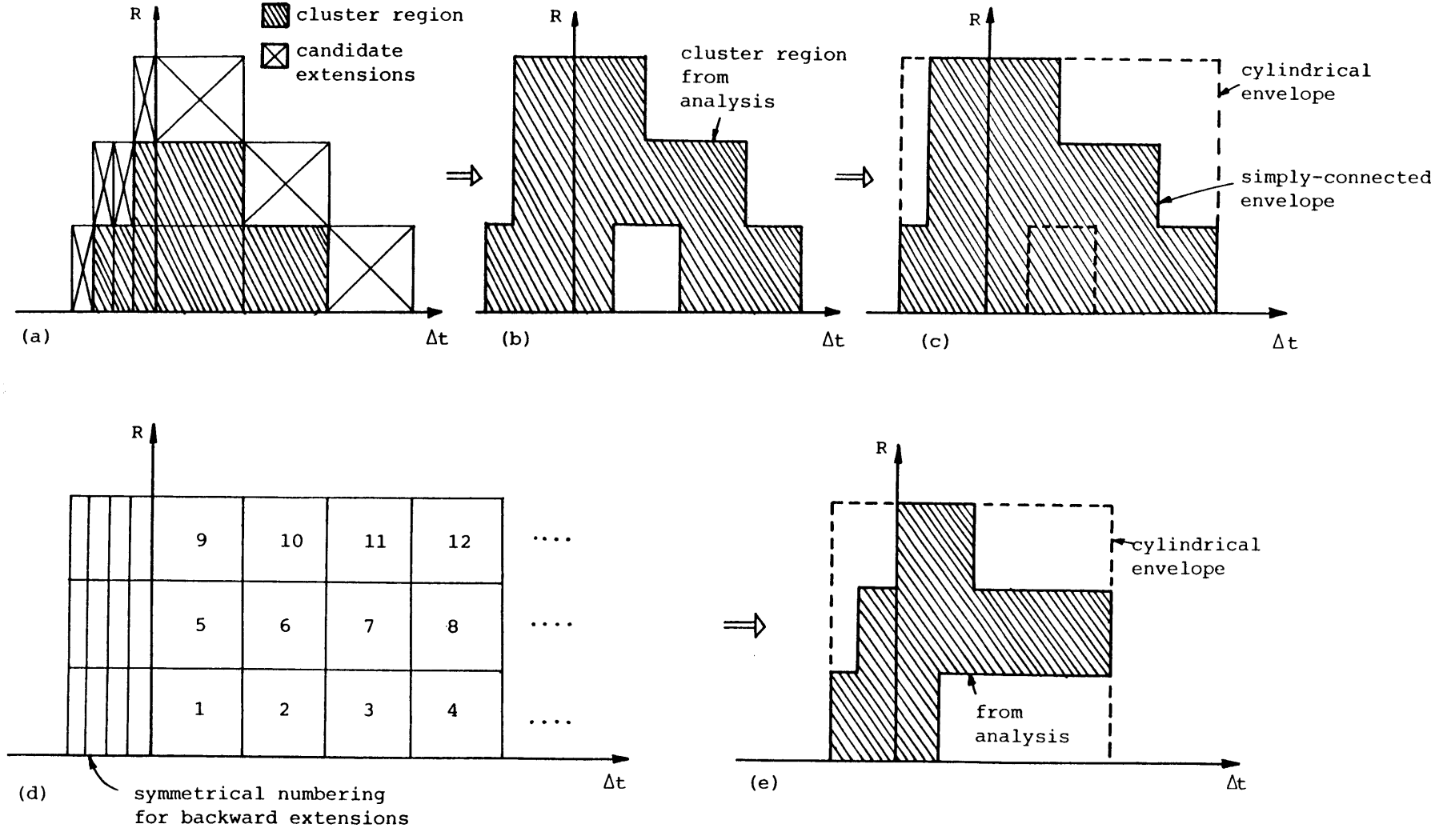


Fig. 3.4 - Estimation of cluster region using two-dimensional schemes.

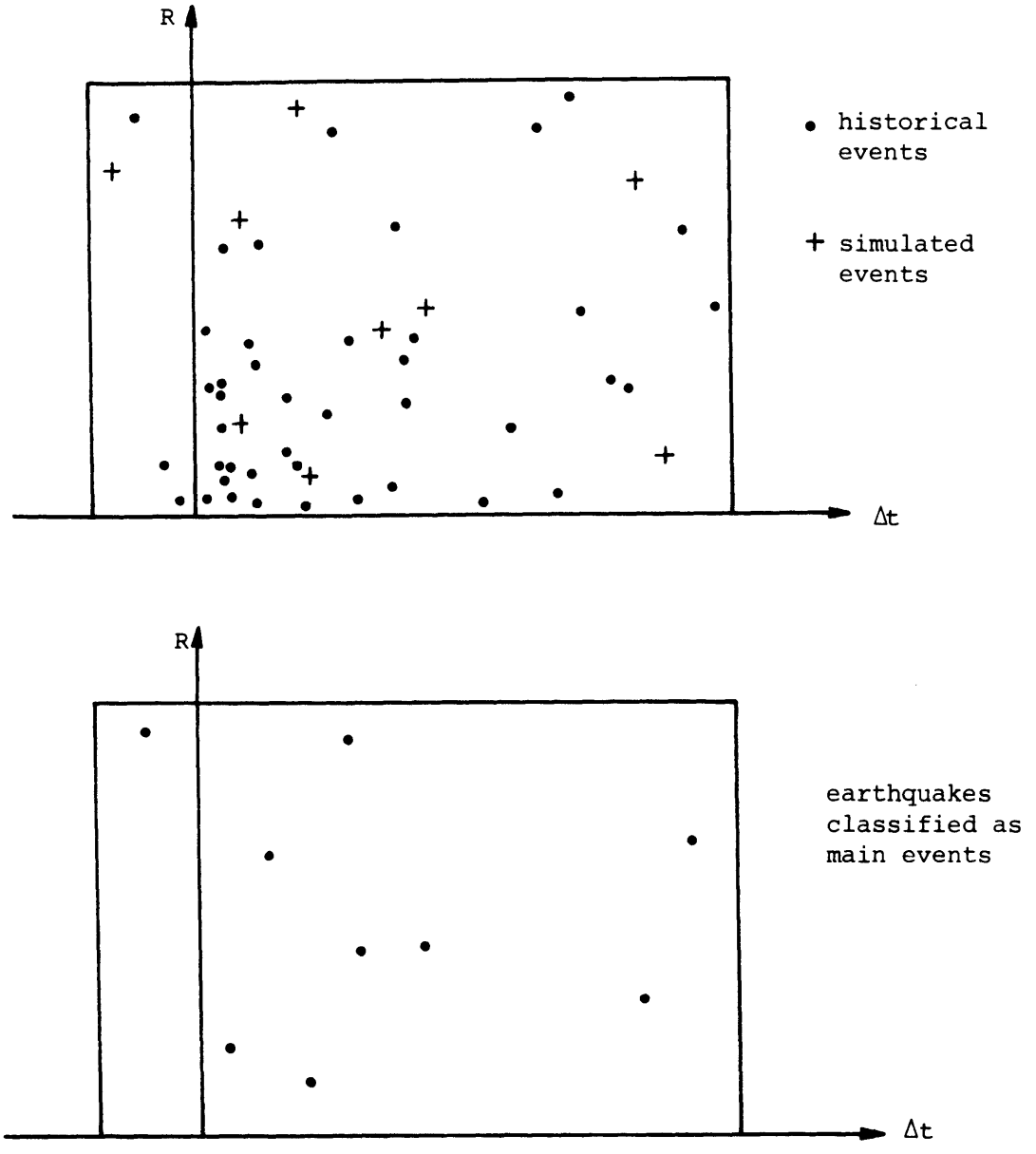


Figure 3.5 - Identification of secondary events inside the cluster region through Poisson thinning.

All Events 1625-1980

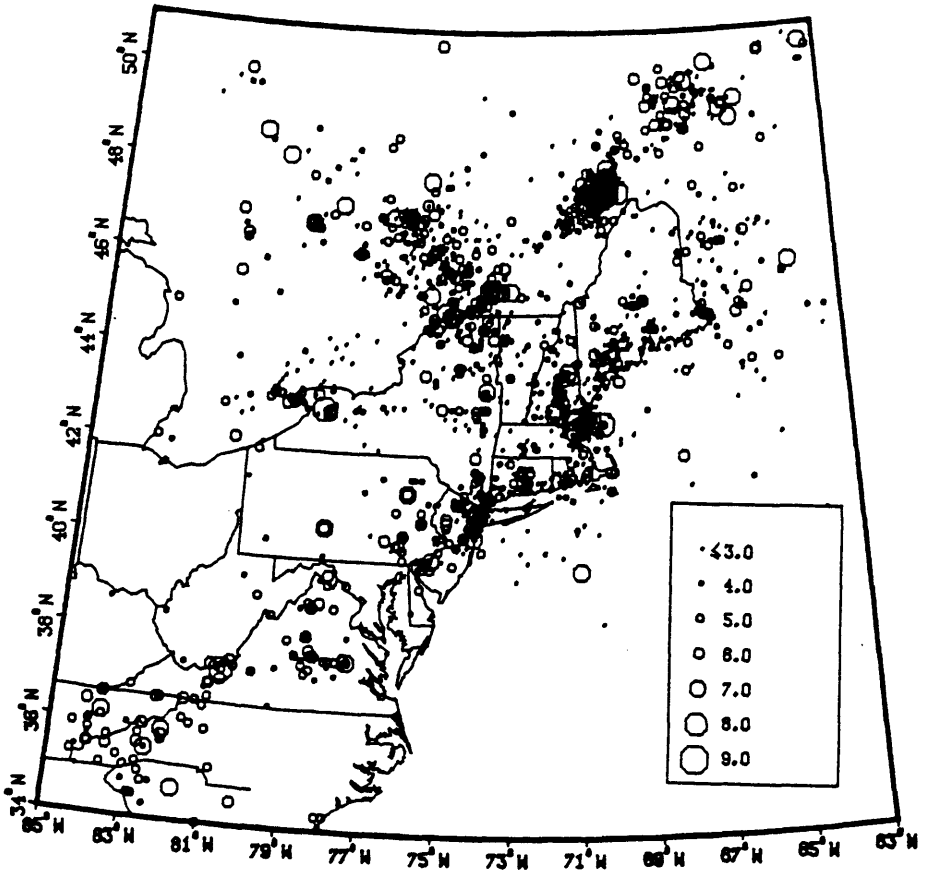


Figure 3.6 - Events of MM Intensity 1 or greater included in the Weston Observatory Catalog. Events not originally reported in the MMI scale have been converted using Eq. 3.10.

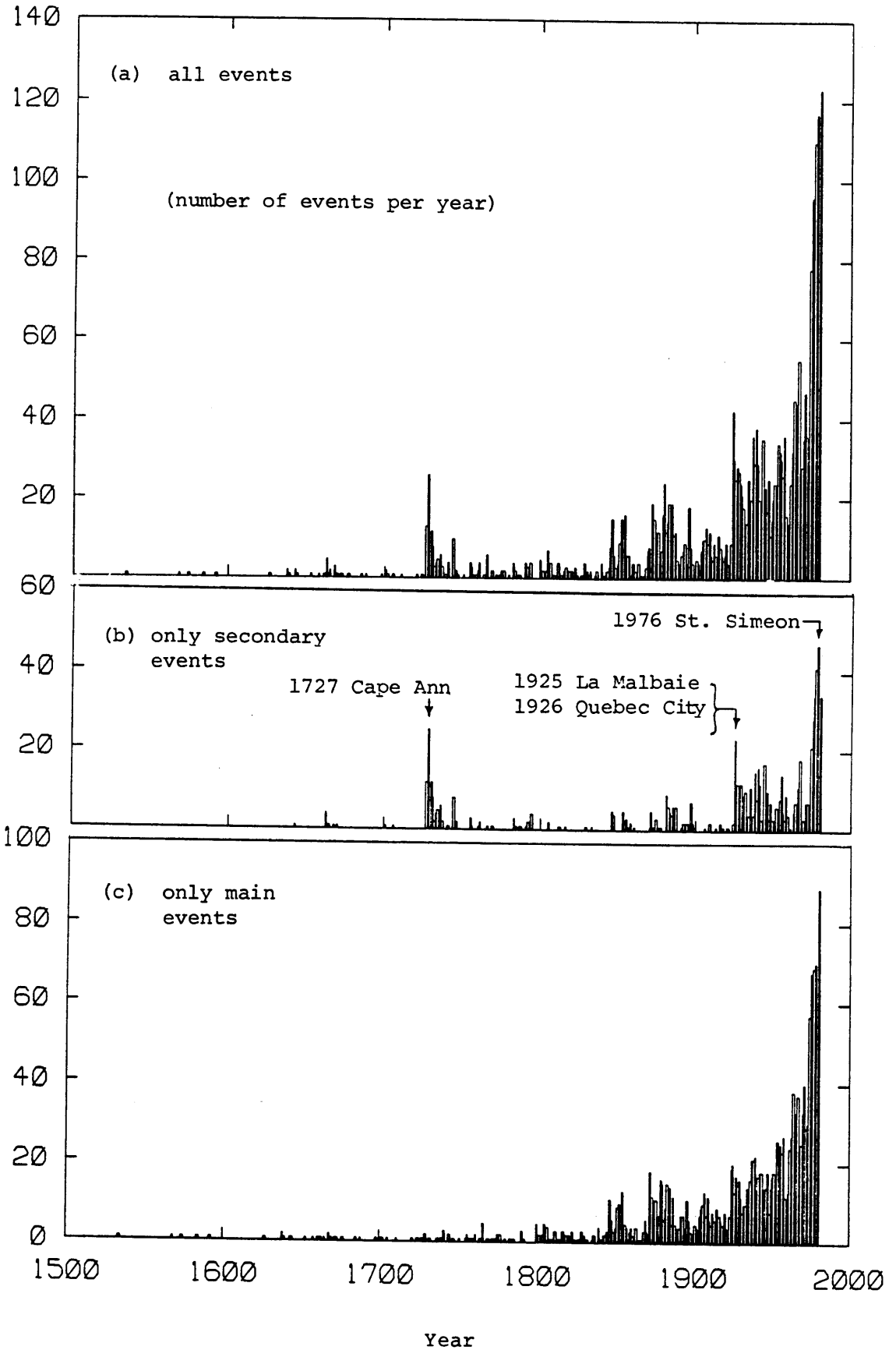
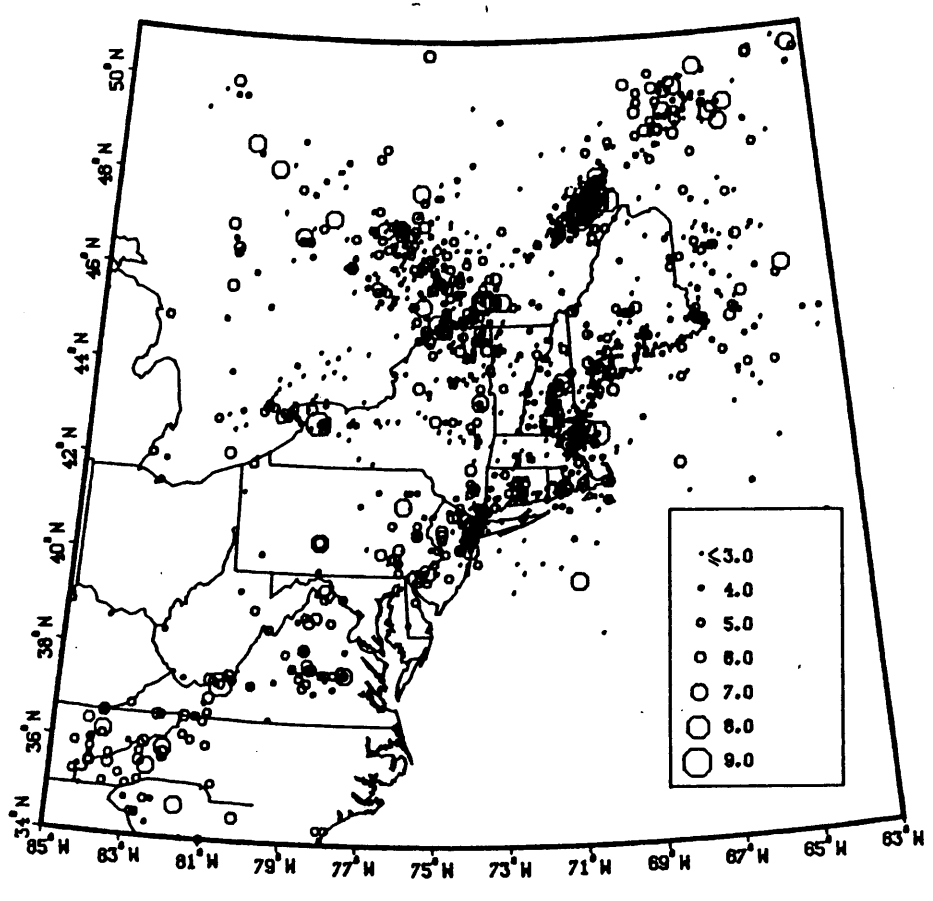
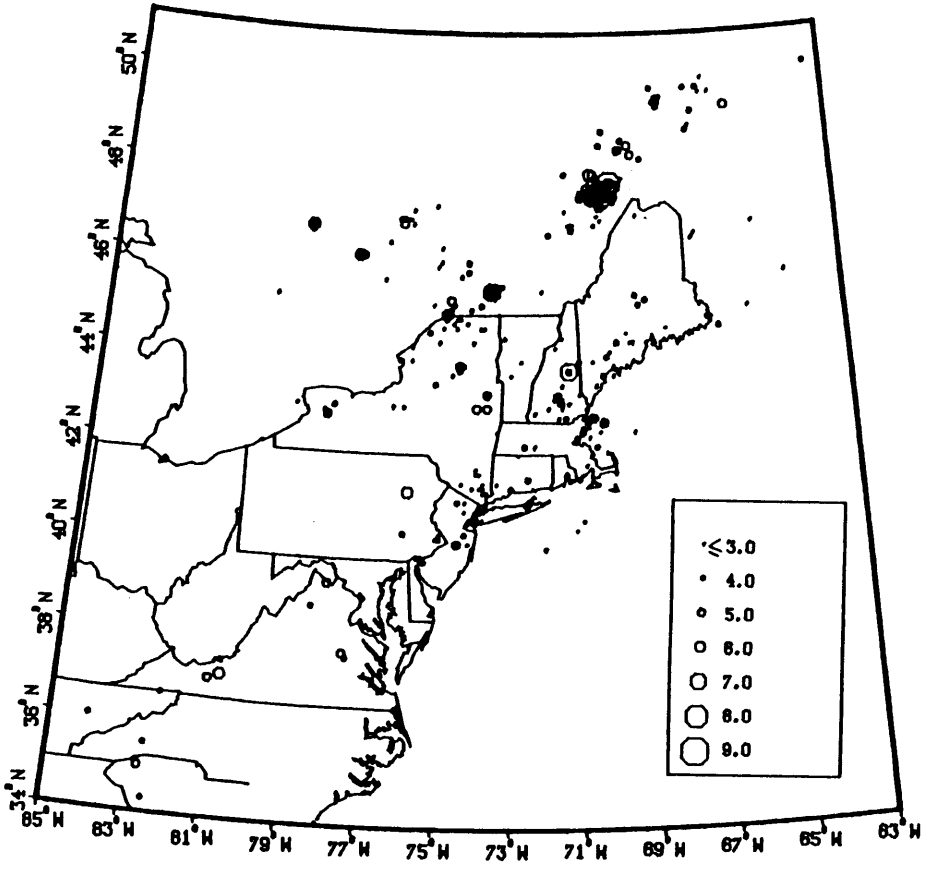


Figure 3.7 - Count plots in time (a) of all the events in the catalog, (b) of the secondary events identified by the procedure, and (c) of the remaining main events.



Main events 1625-1980



Secondary events 1625-1980

Figure 3.8:- Geographical distribution of main and secondary events identified by the present method.

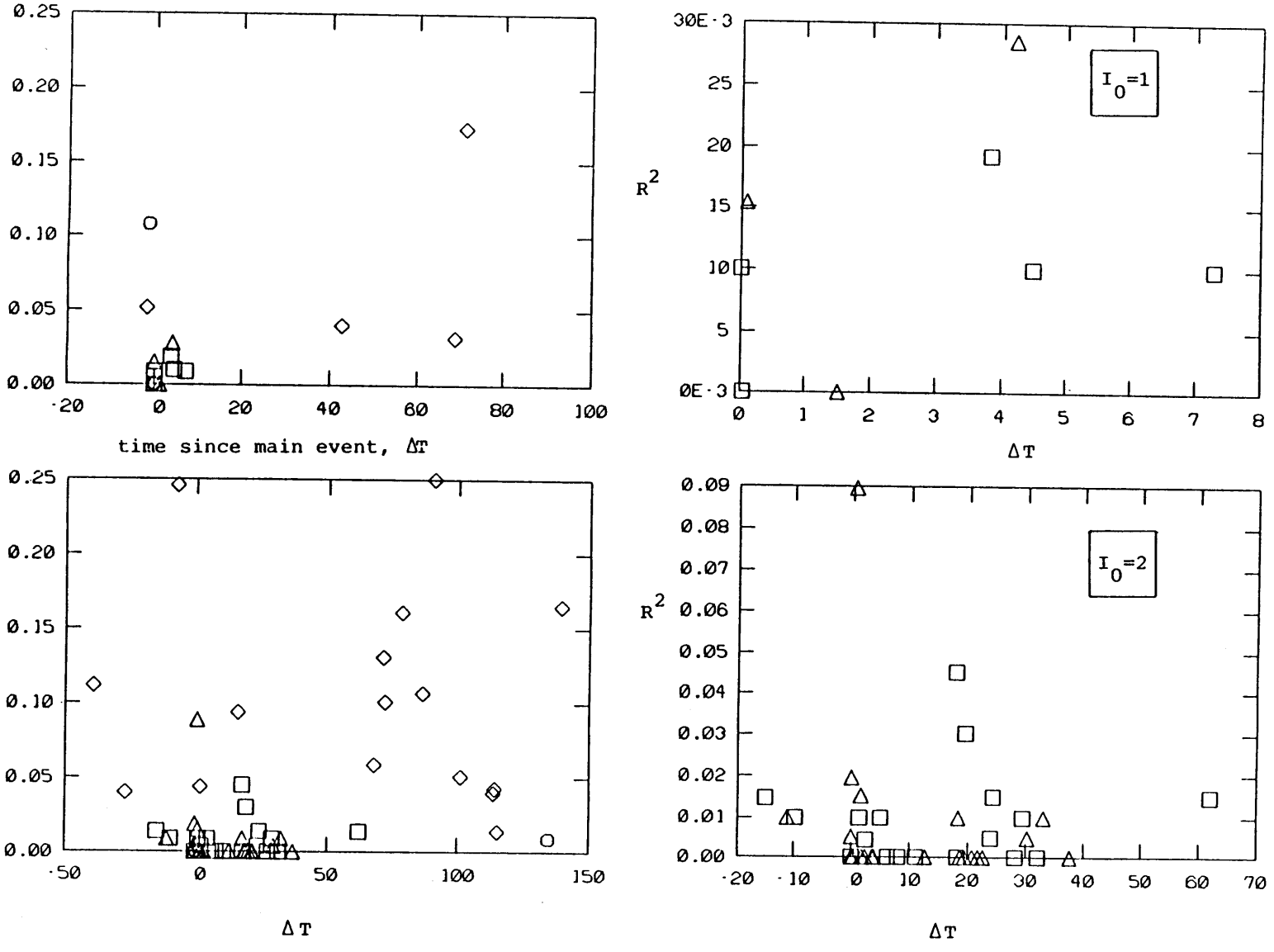


Figure 3.9 - Distribution of secondary events and background activity around earthquakes with associated clusters. For each intensity I_o, the plot on the right contains only the secondary events, whereas the plot on the left includes main events in the background of intensity at most I_o, R is in degrees, ΔT in days. Symbols are defined in Equation 3.12.

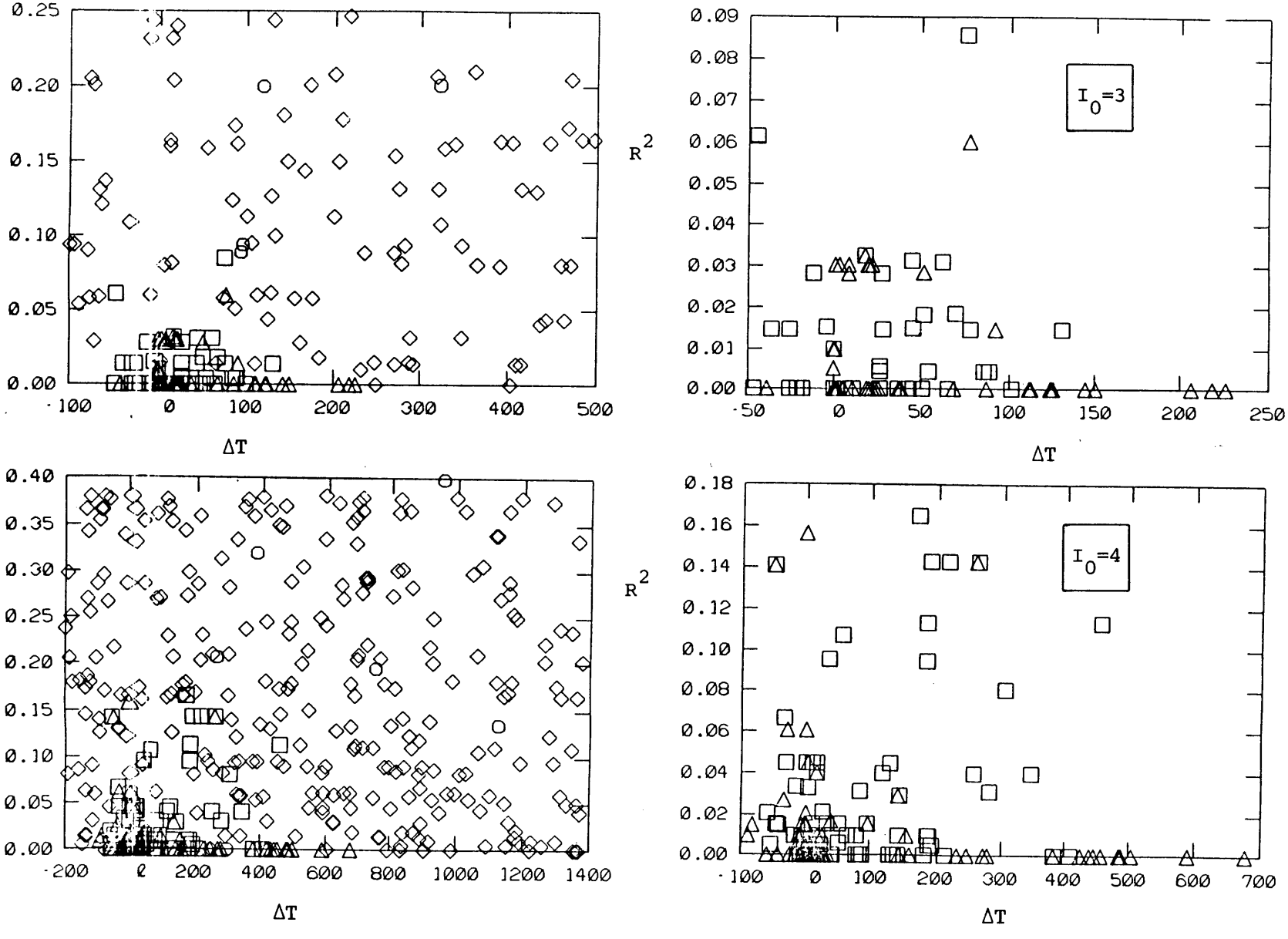


Figure 3.9 (continued)

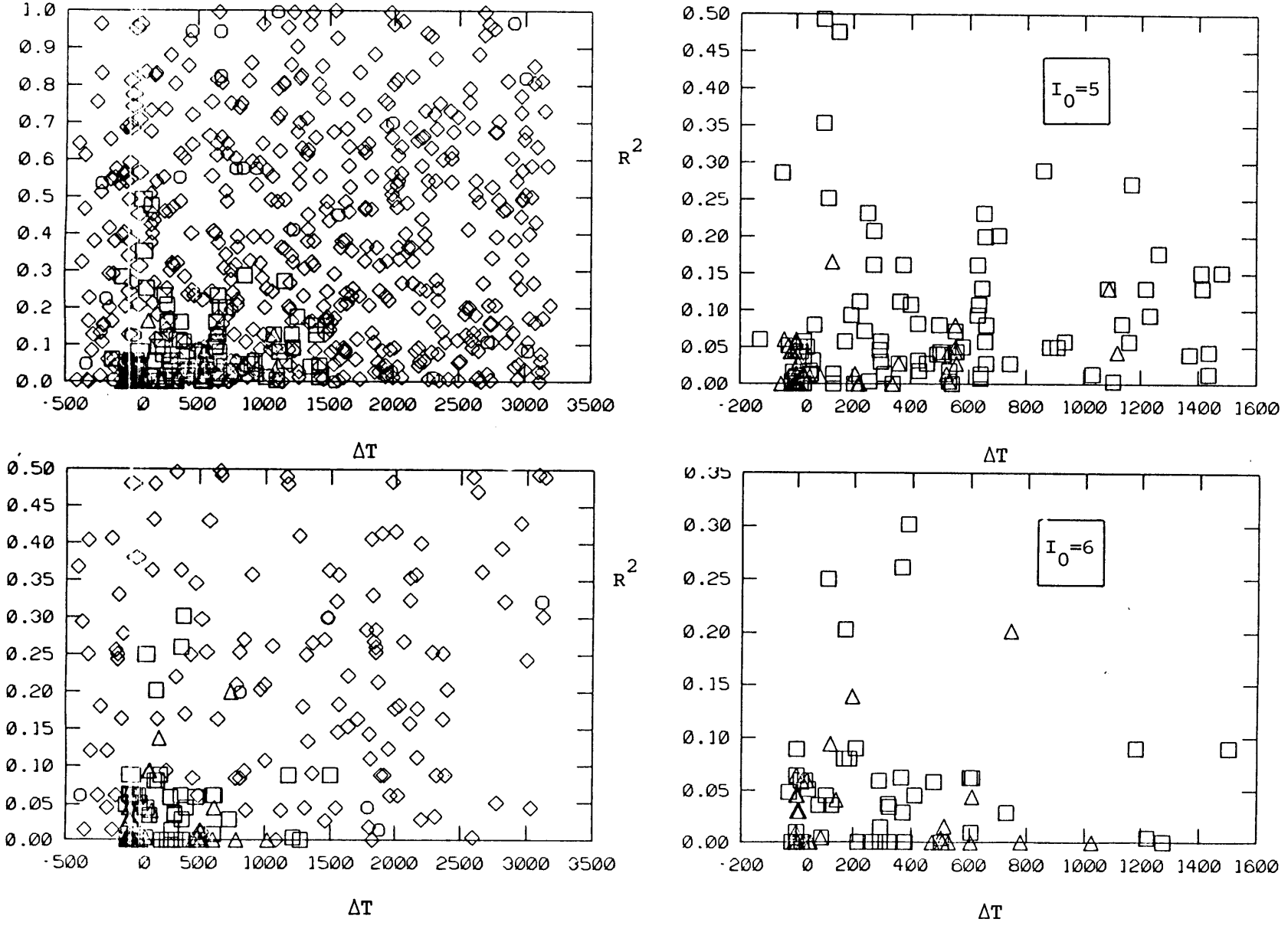
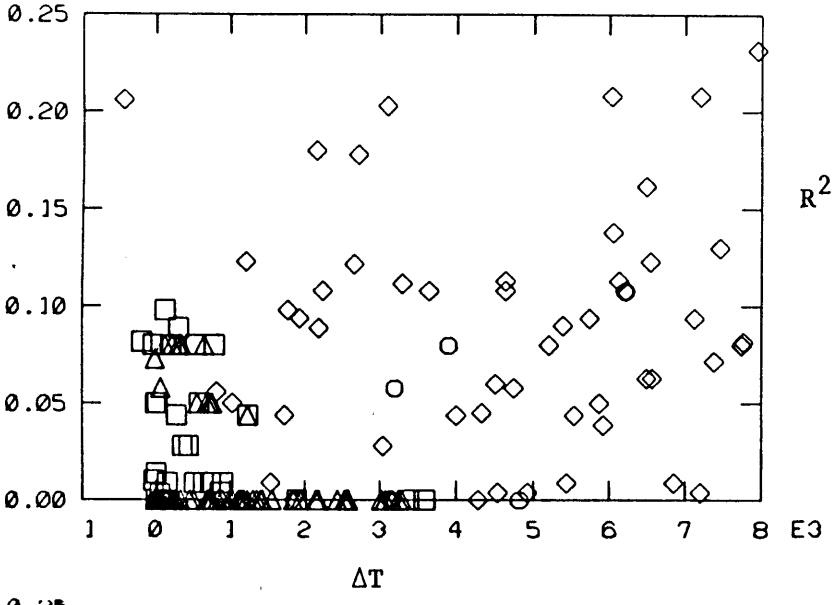
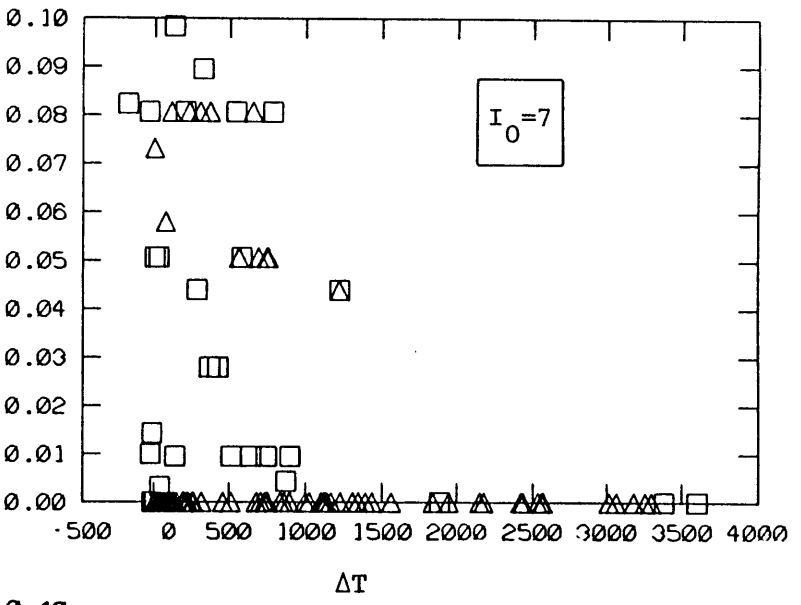
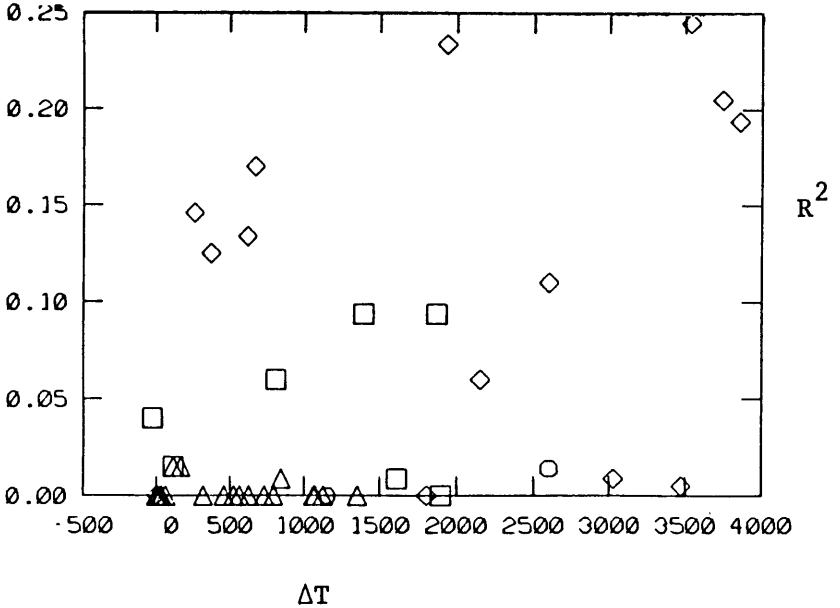


Figure 3.9 (continued)







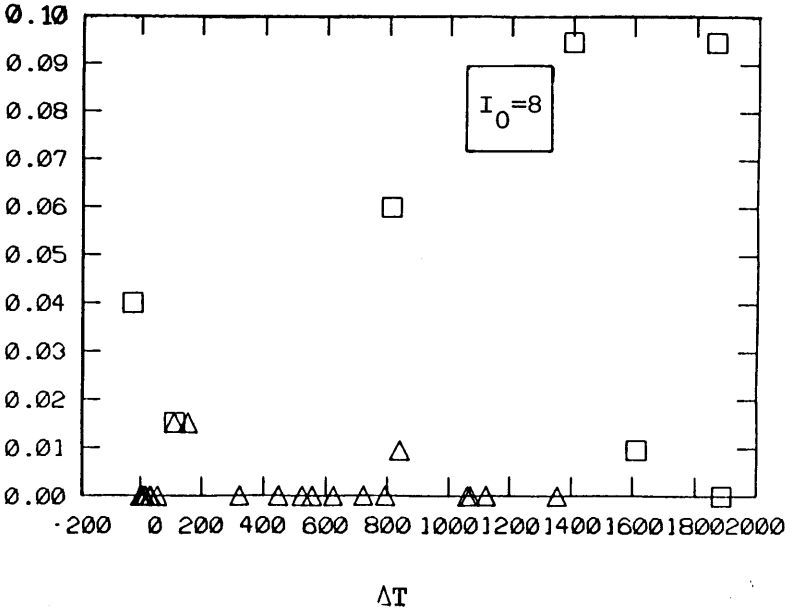
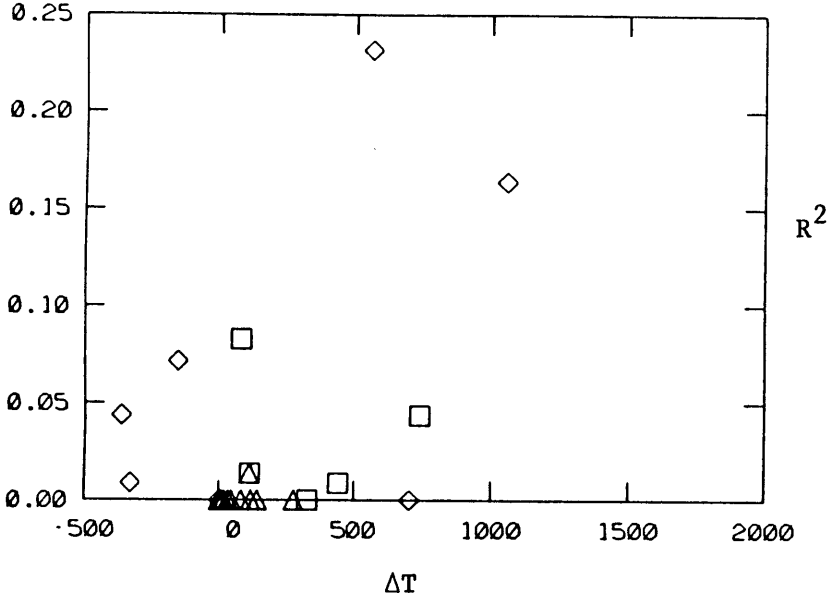


Figure 3.9 (continued)



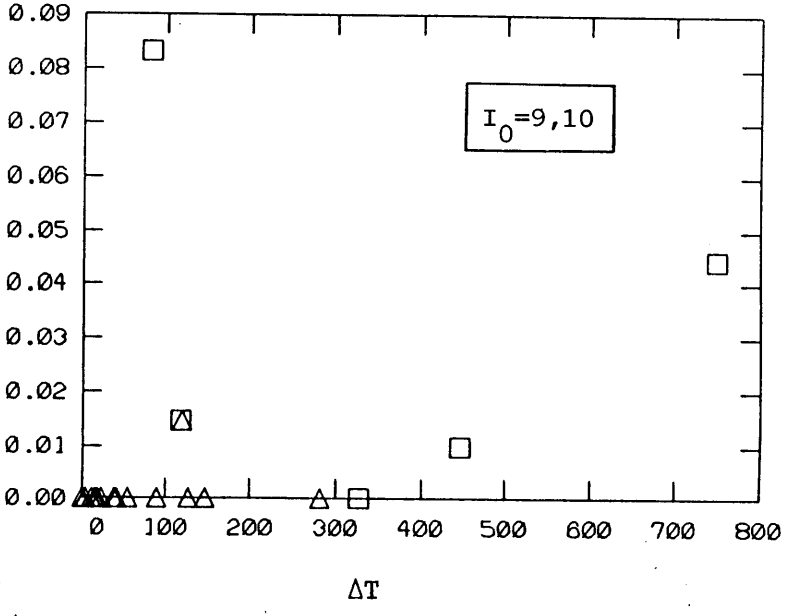


Figure 3.9 (End)

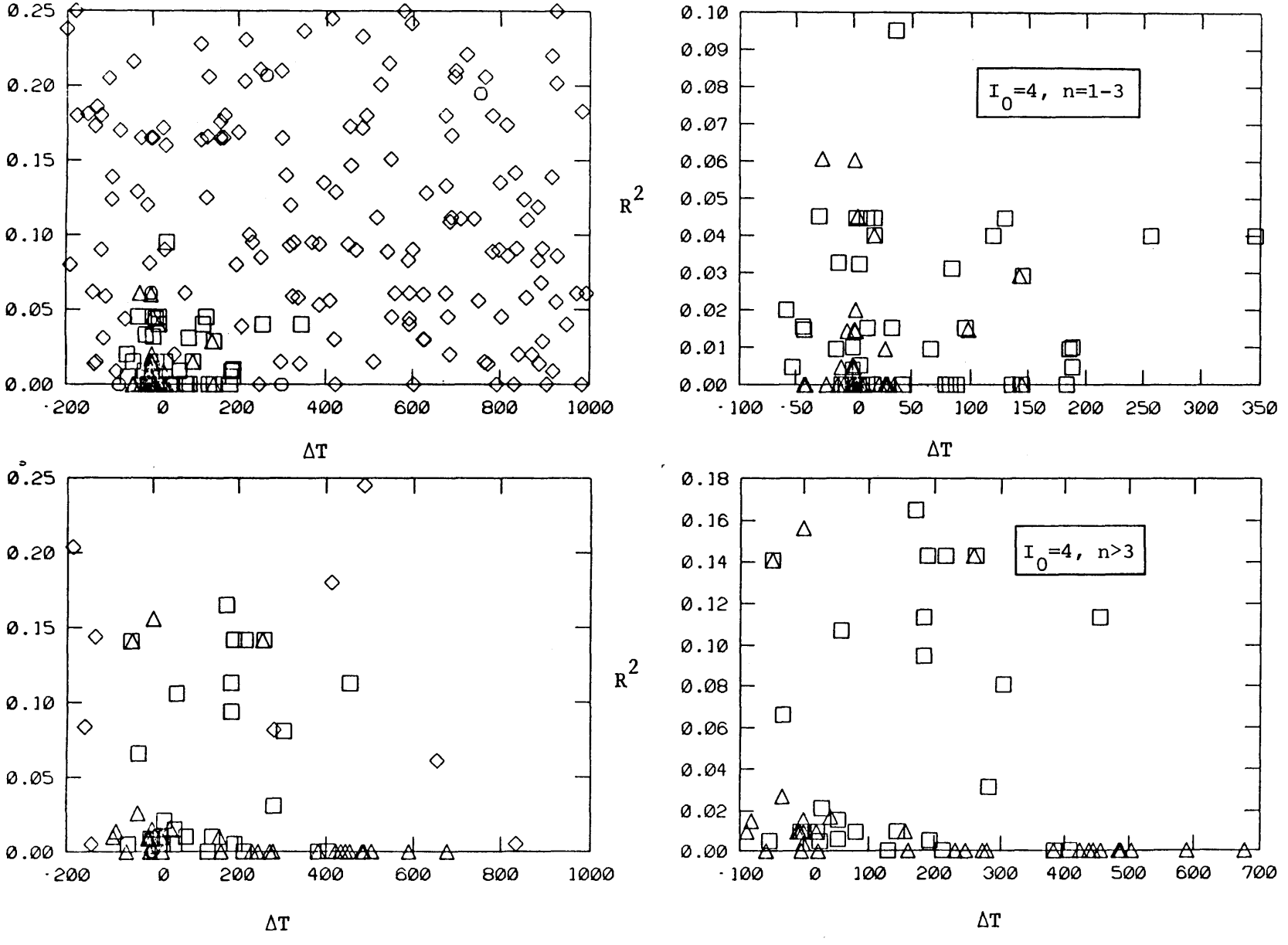


Figure 3.10 - Selected clusters for main events of intensity 4,5,6 and 7. Same format as Fig. 3.9.

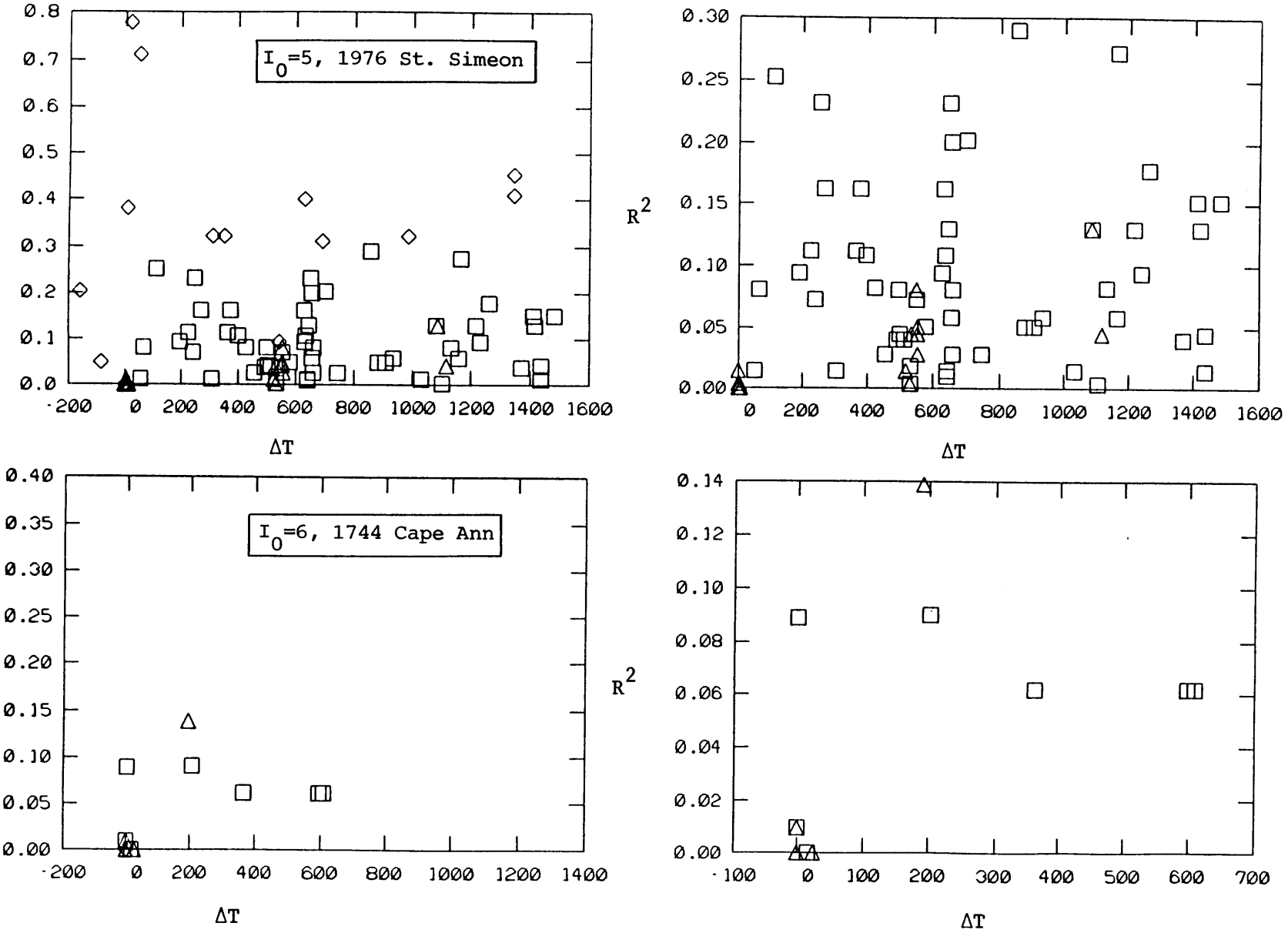


Figure 3.10 (continued)

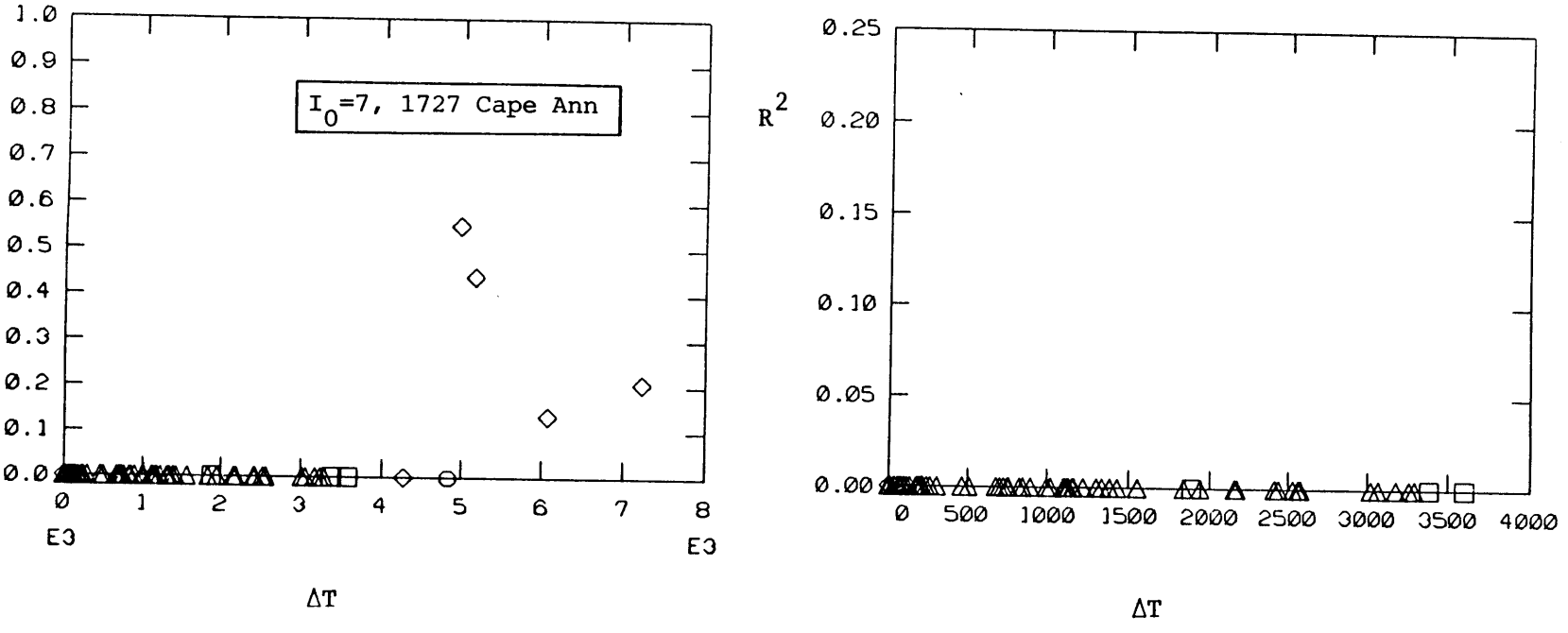


Figure 3.10 (End)

Note: Only secondary events are shown.
The same format as in Fig. 3.9 is used

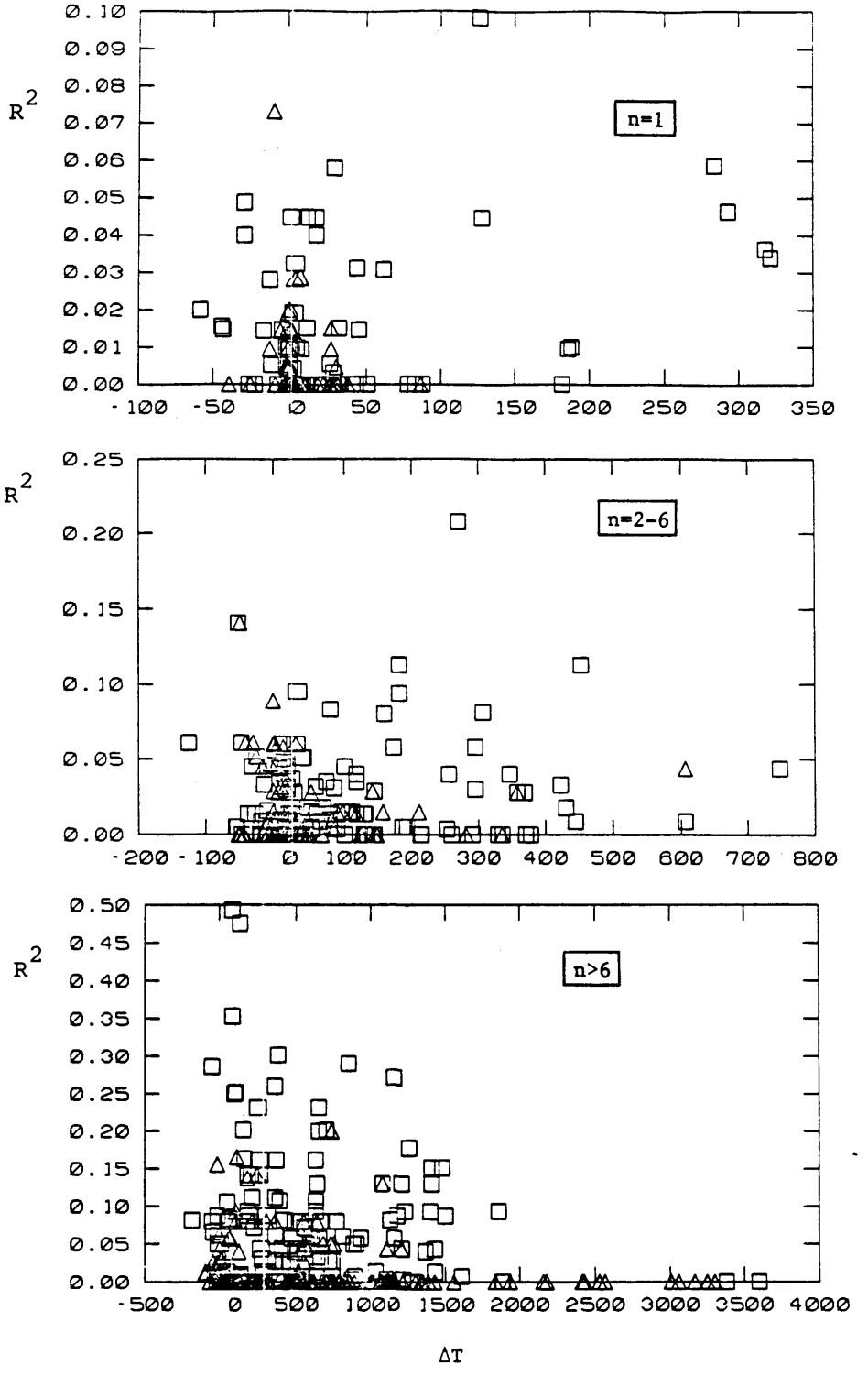


Figure 3.11 - Aggregation of clusters by cluster size.

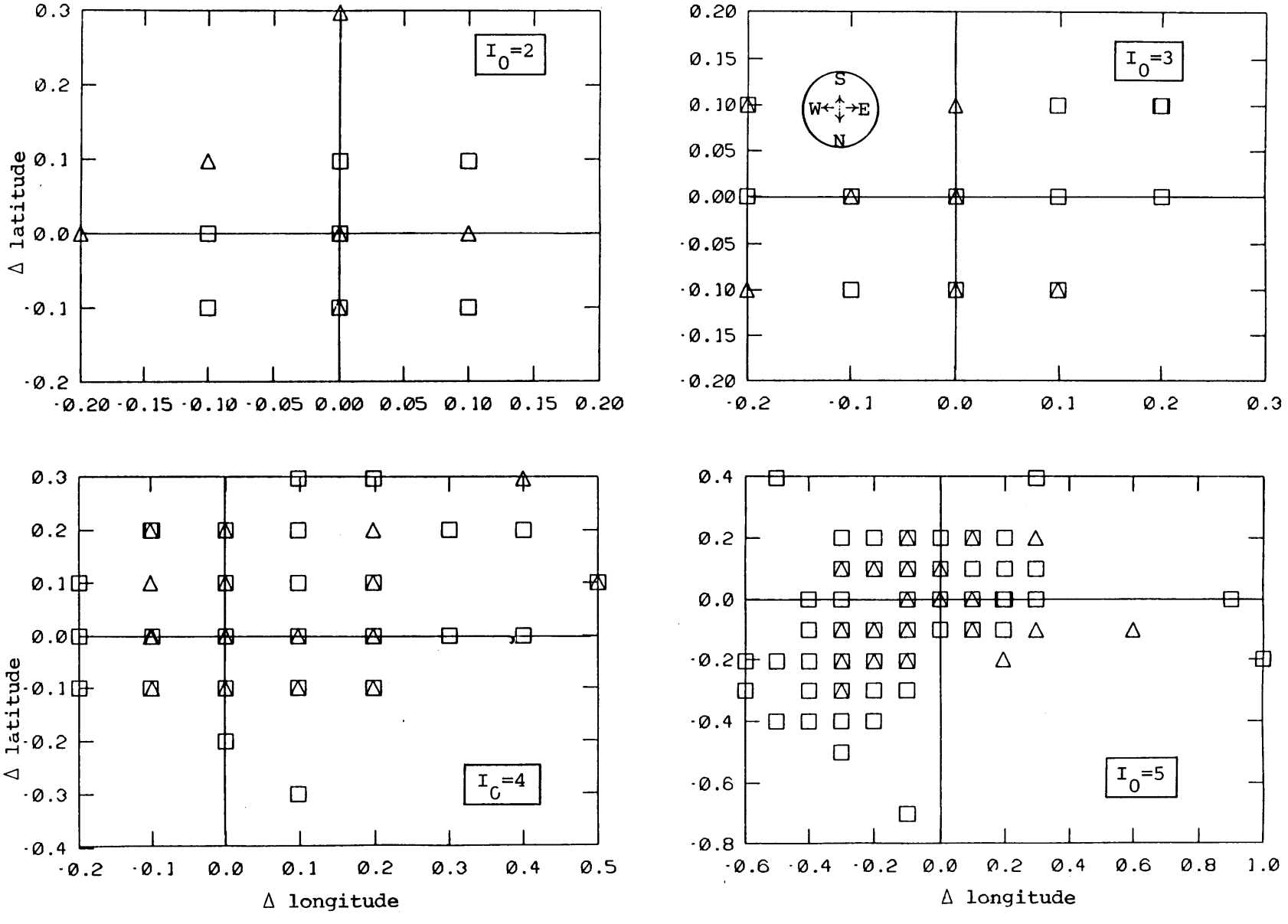


Figure 3.12 - Spatial distribution of secondary earthquakes around the associated main event. Symbols according to Eq. 3.12.

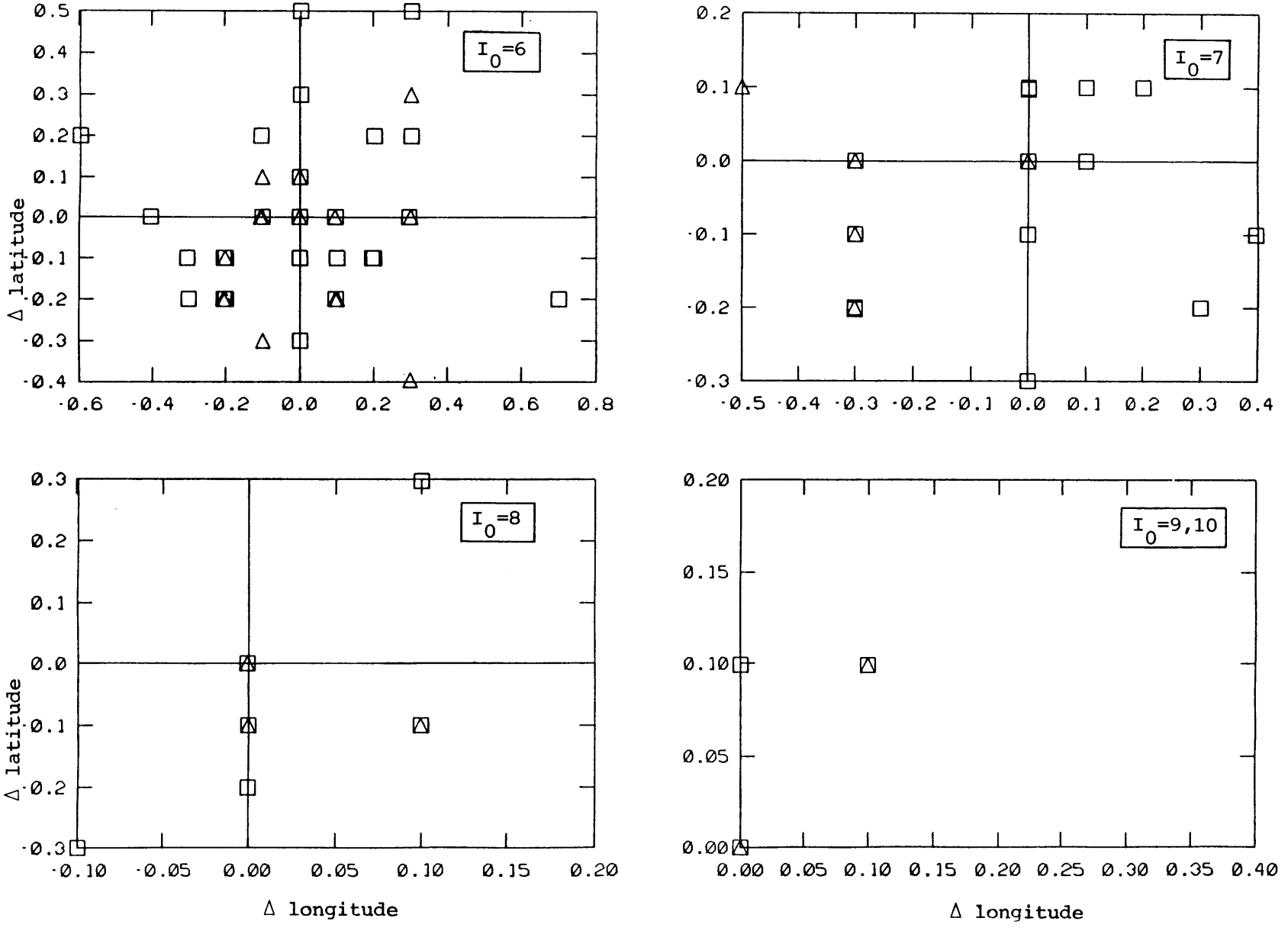


Figure 3.12 (End)

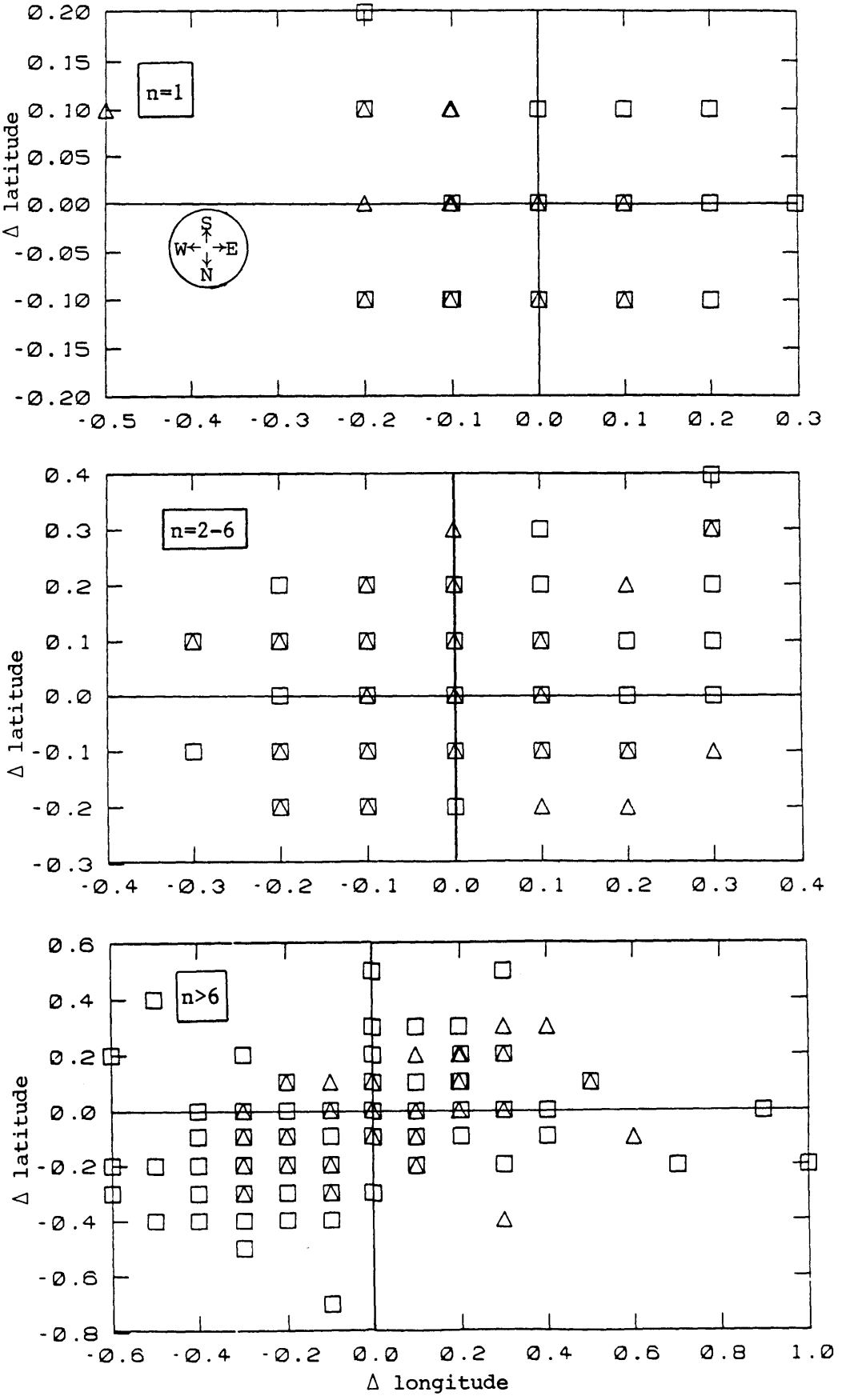


Figure 3.13 Spatial distribution of secondary events grouped according to cluster size

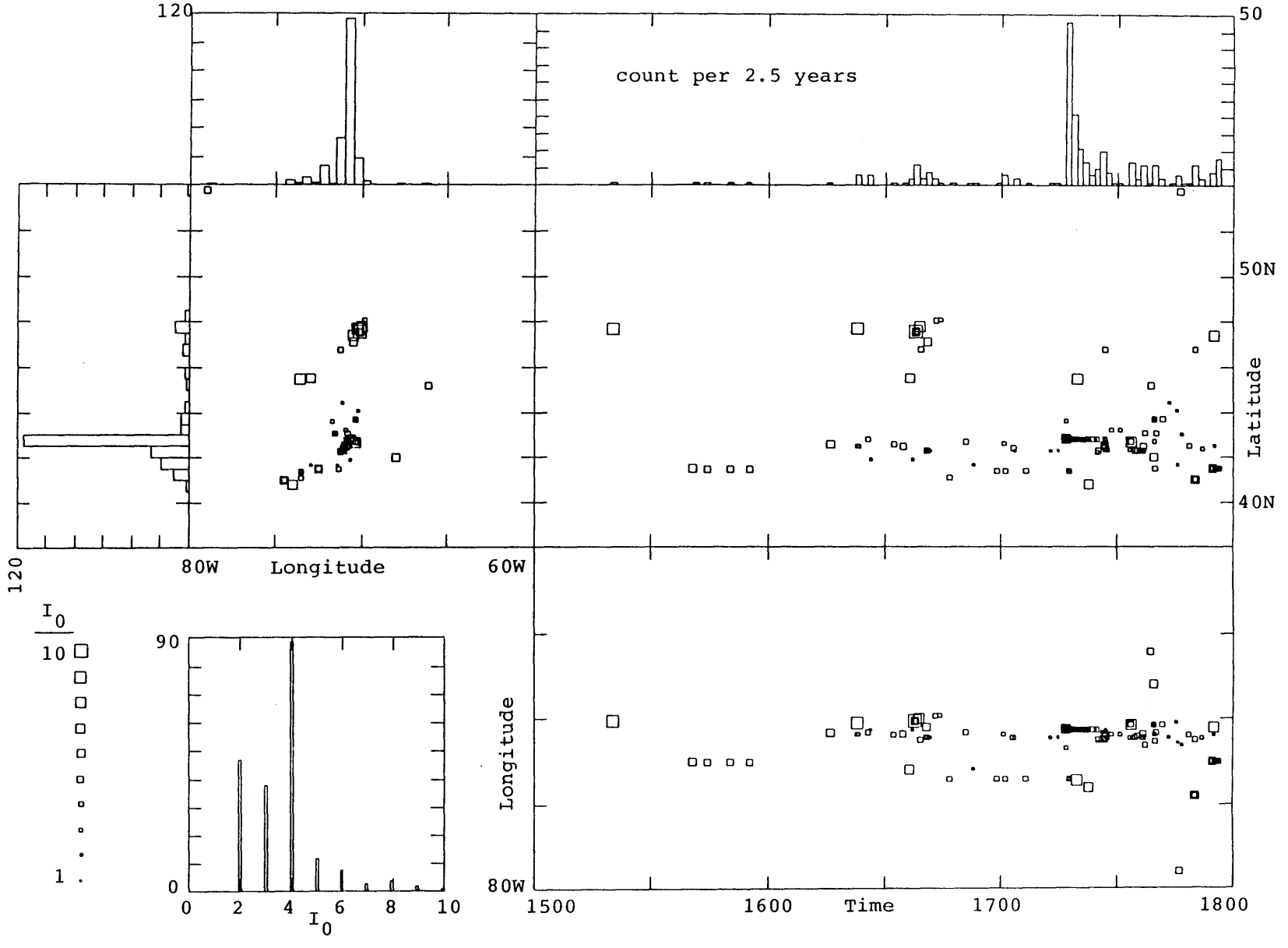


Figure 3.14a - 1. All events (1500-1800)

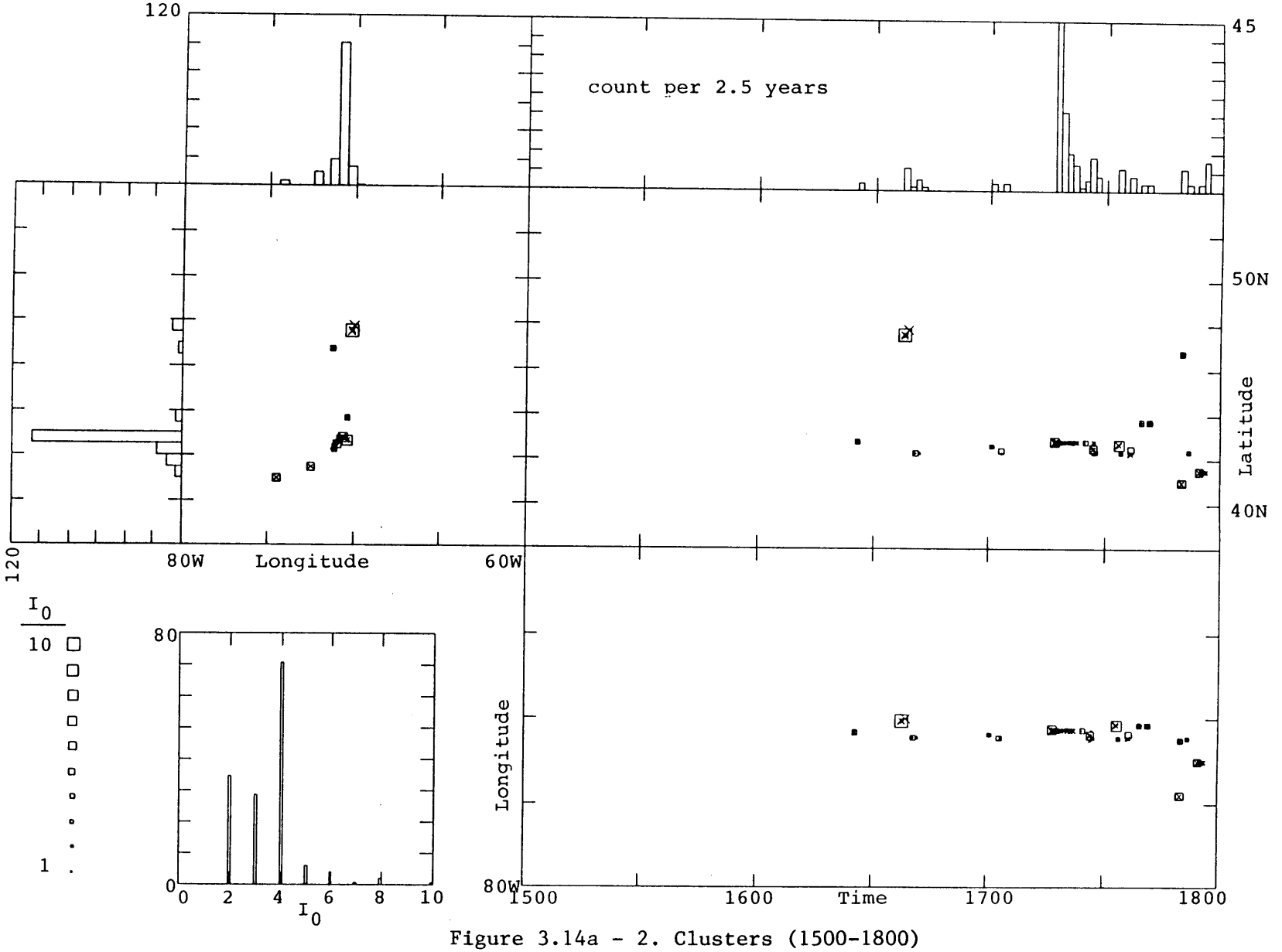


Figure 3.14a - 2. Clusters (1500-1800)

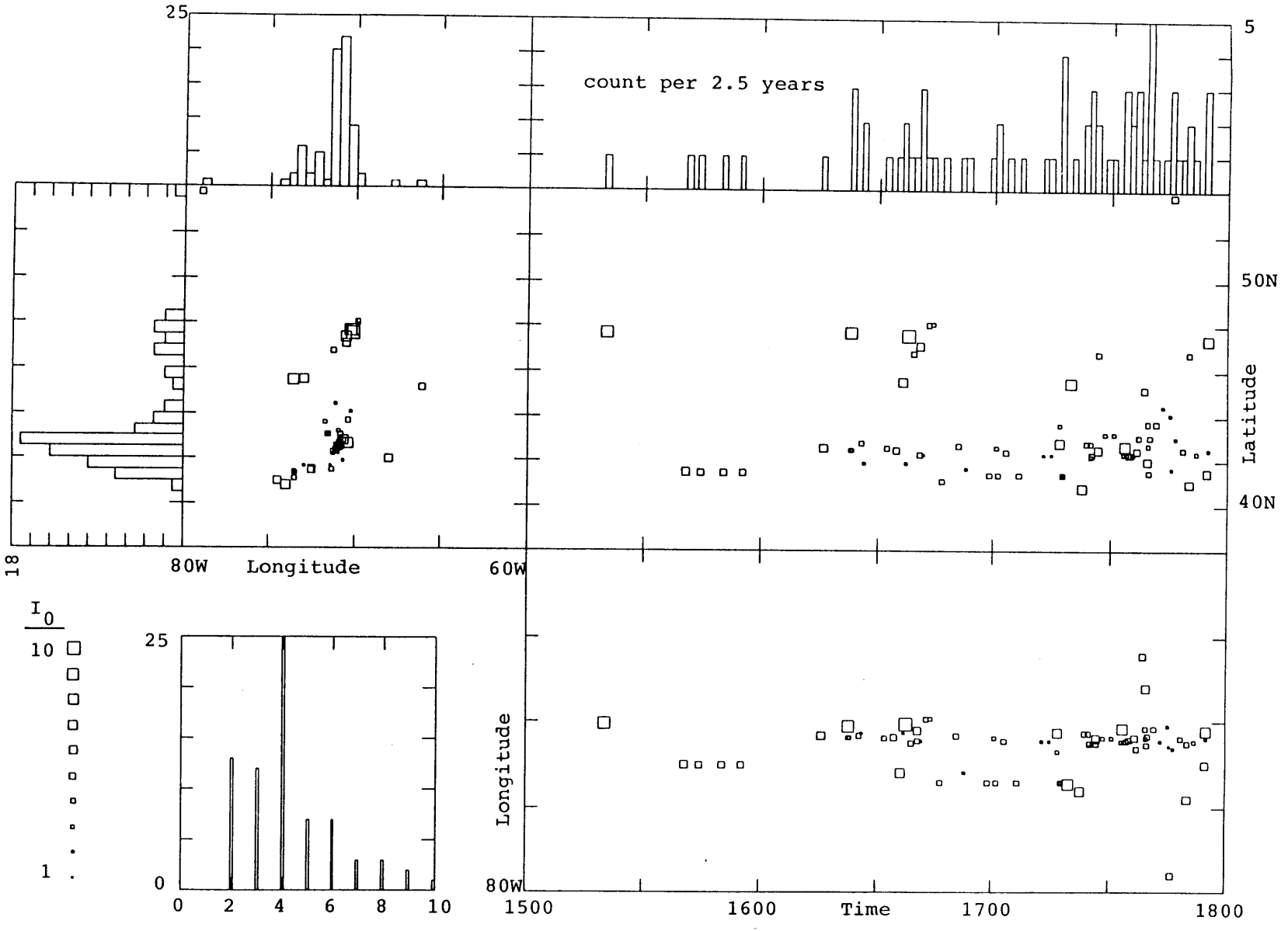


Figure 3.14a - 3. Main events (1500-1800)

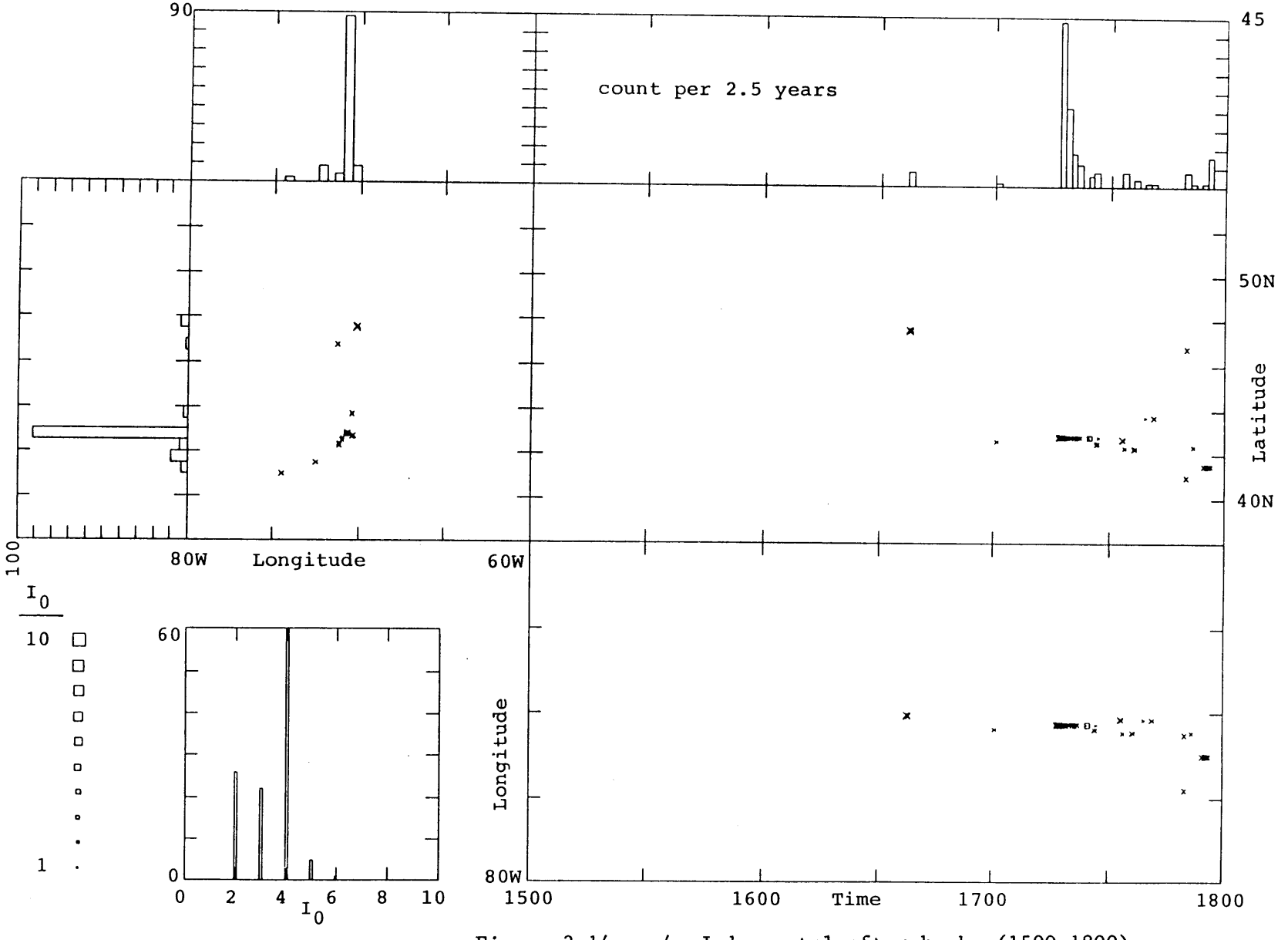


Figure 3.14a - 4. Judgemental aftershocks (1500-1800)

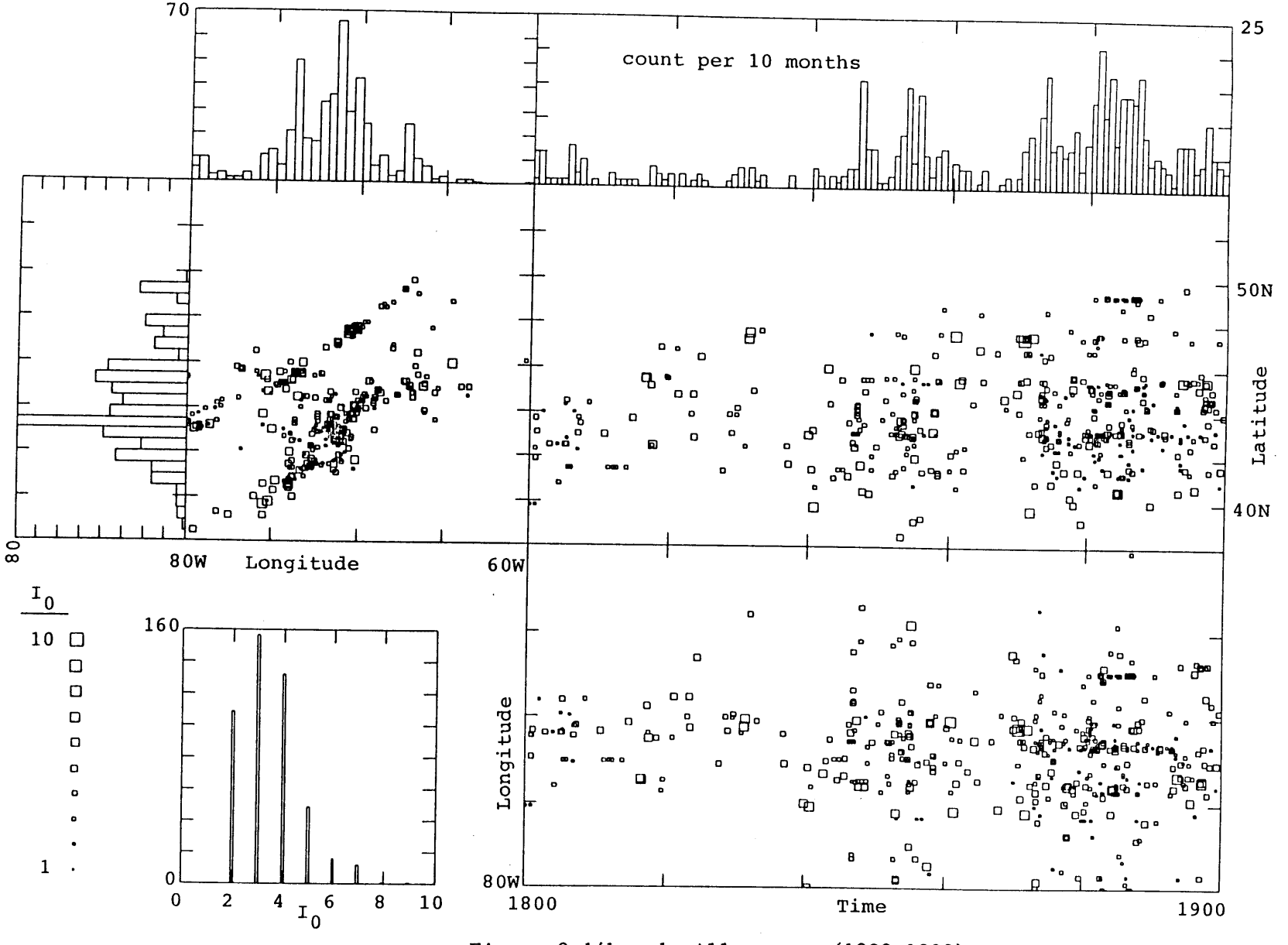


Figure 3.14b - 1. All events (1800-1900)

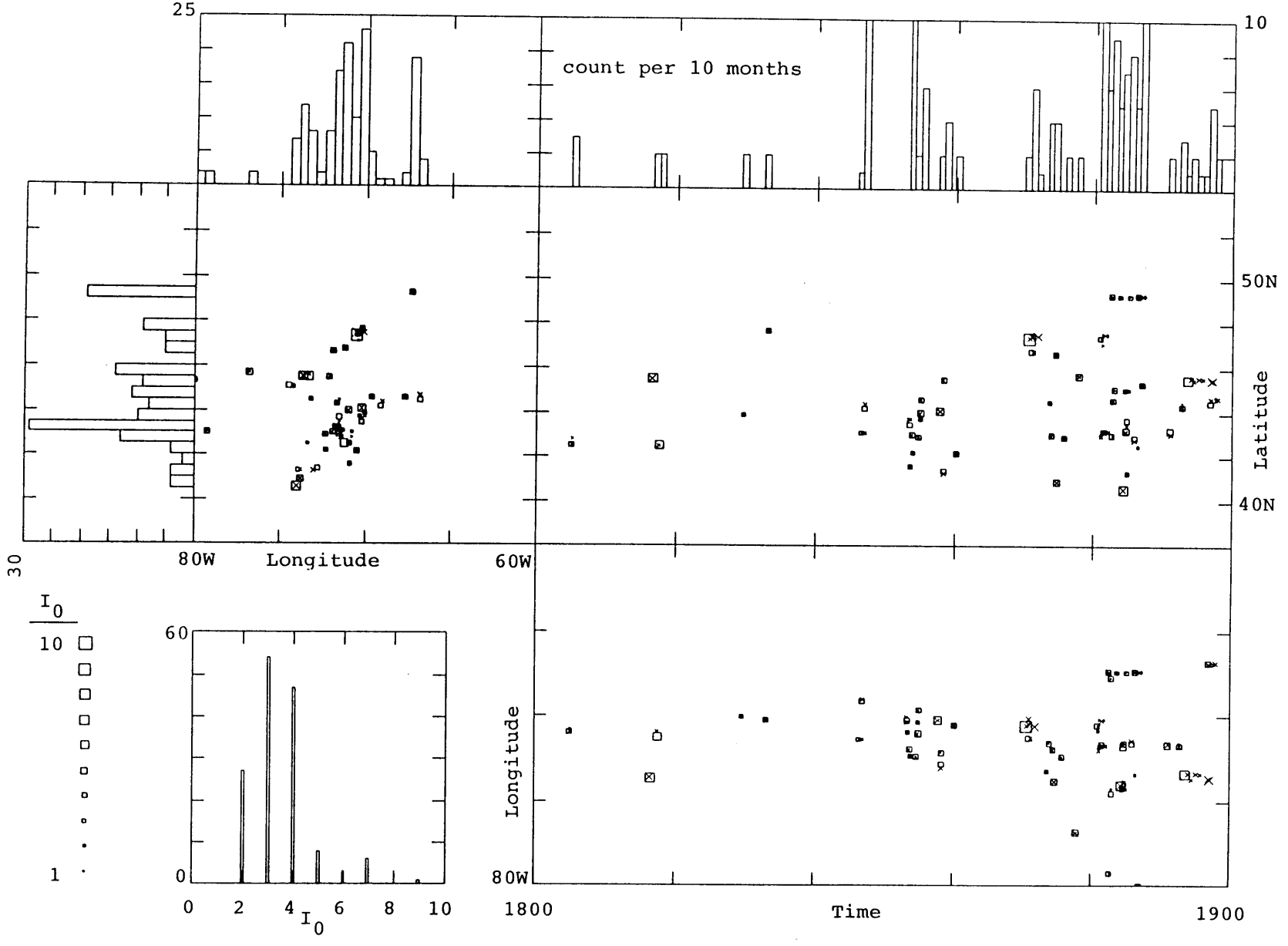


Figure 3.14b - 2. Clusters (1800-1900)

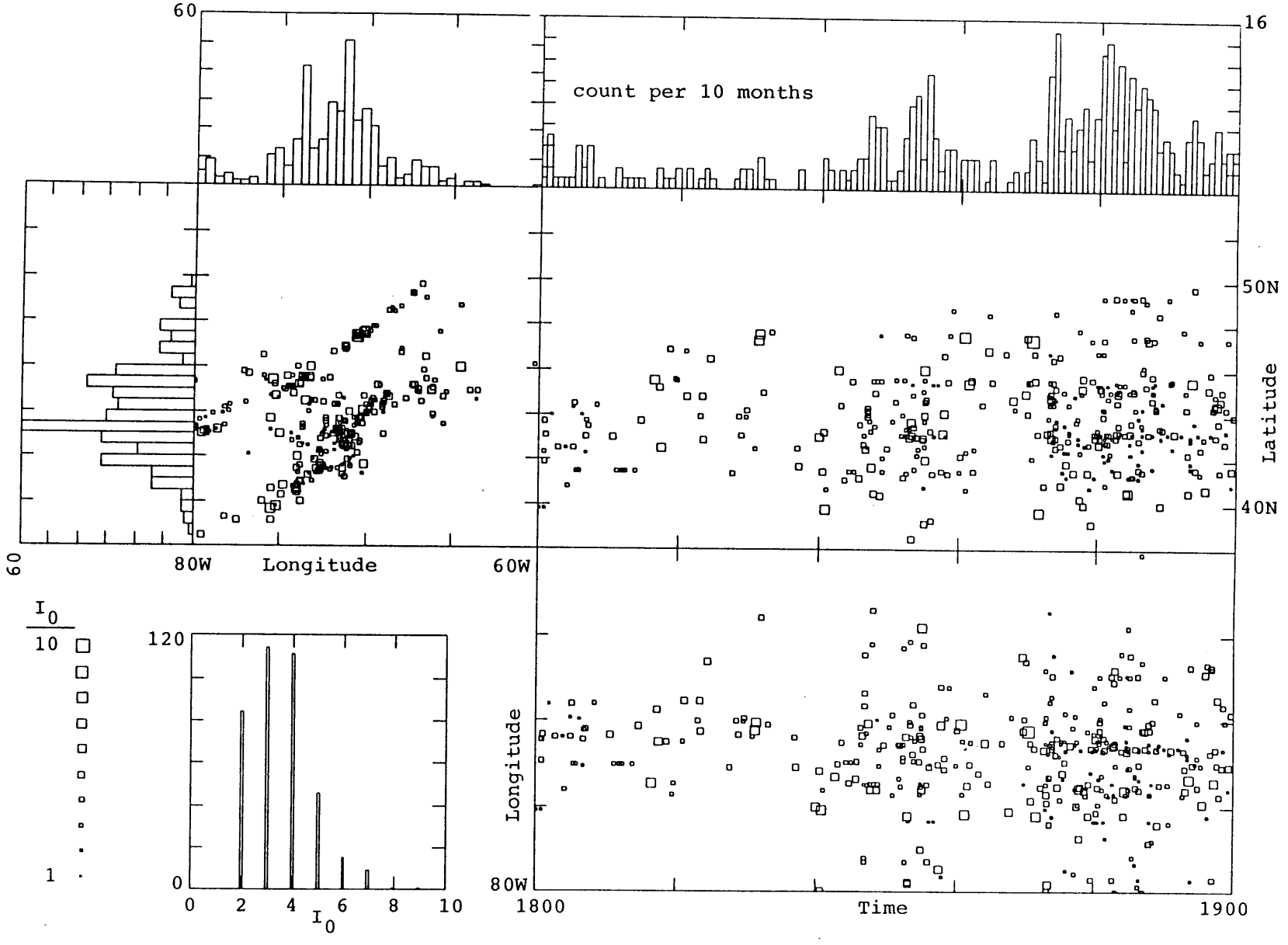


Figure 3.14b - 3. Main events (1800-1900)

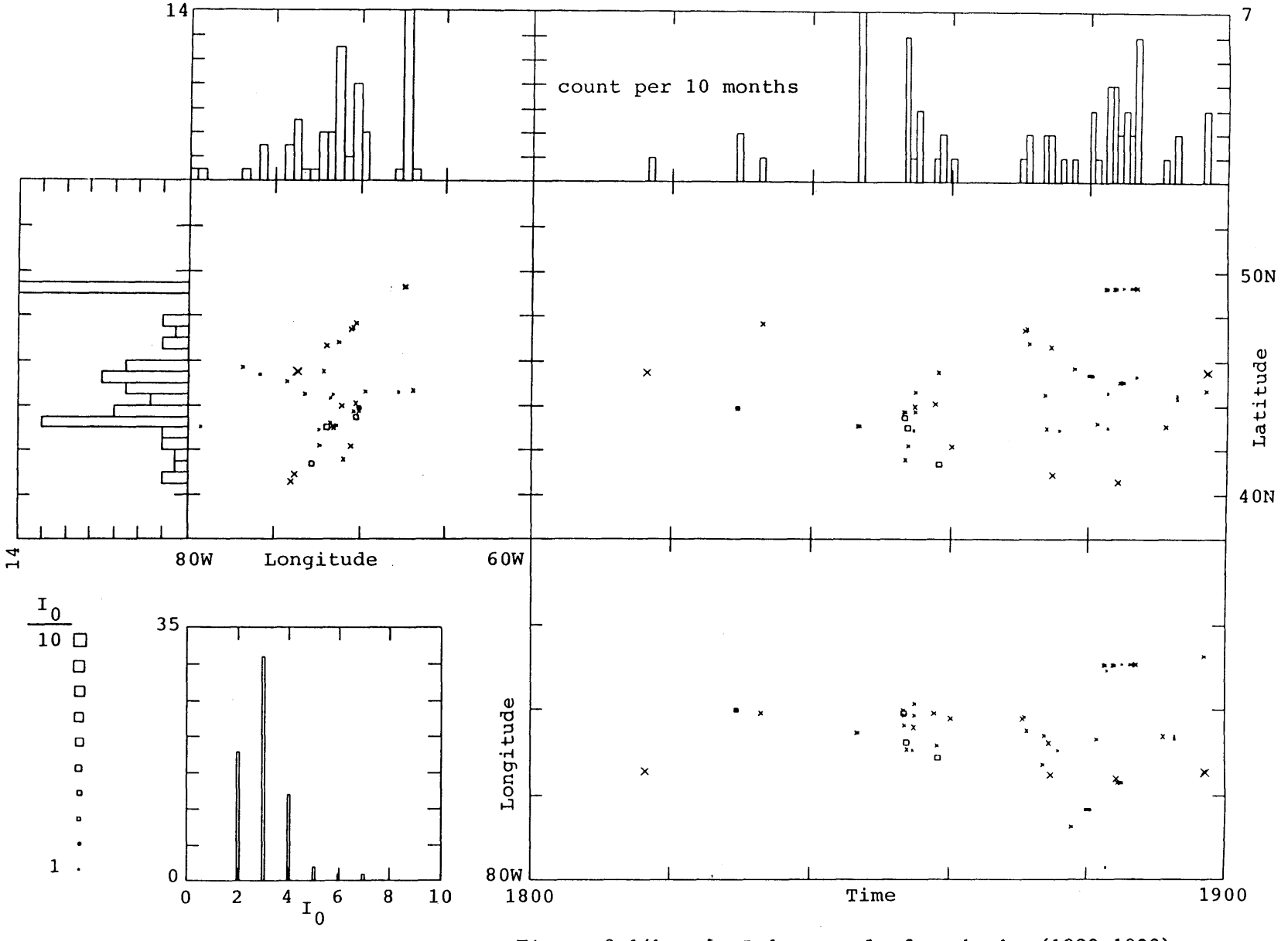


Figure 3.14b - 4. Judgemental aftershocks (1800-1900)

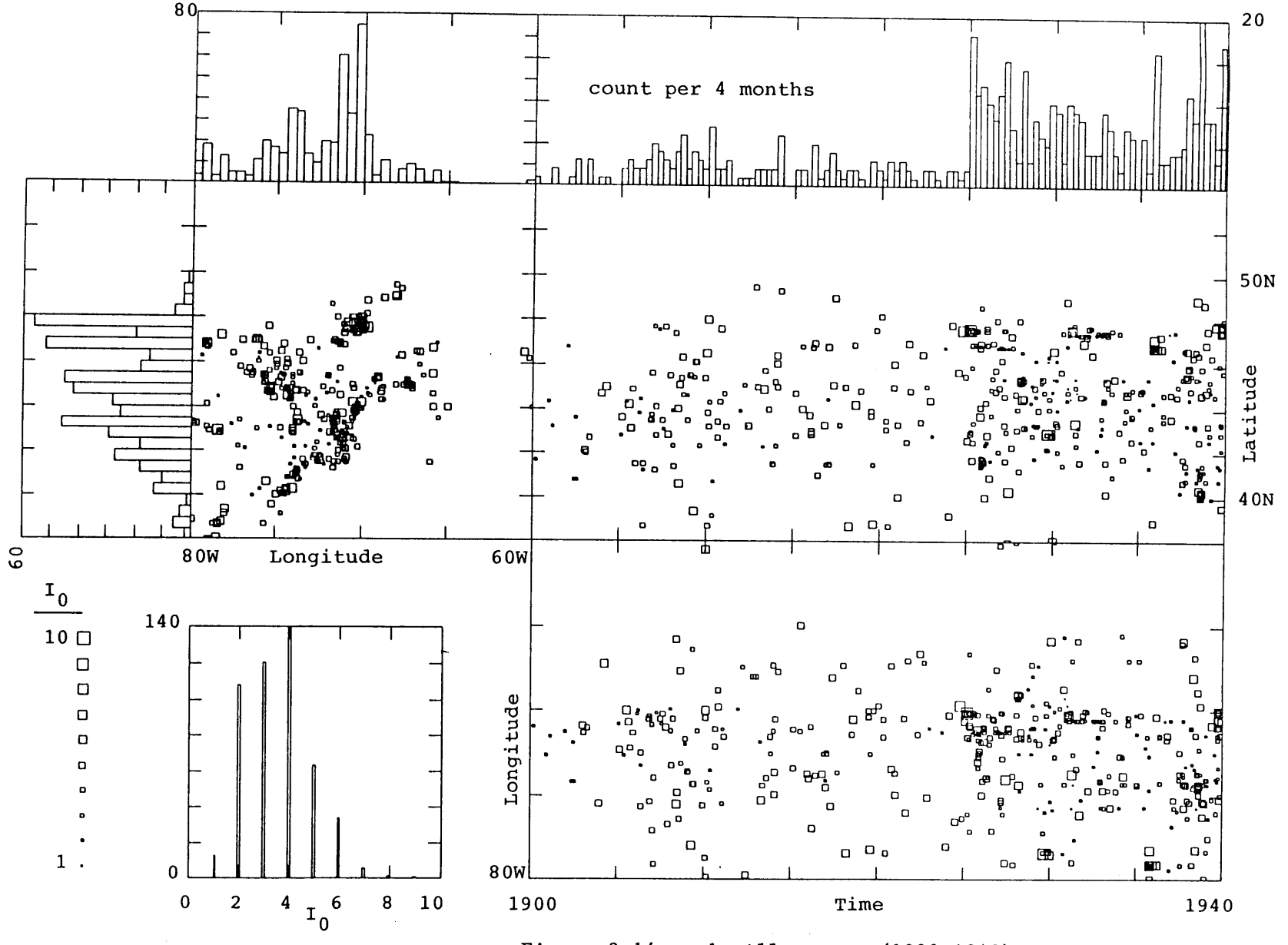


Figure 3.14c - 1. All events (1900-1940)

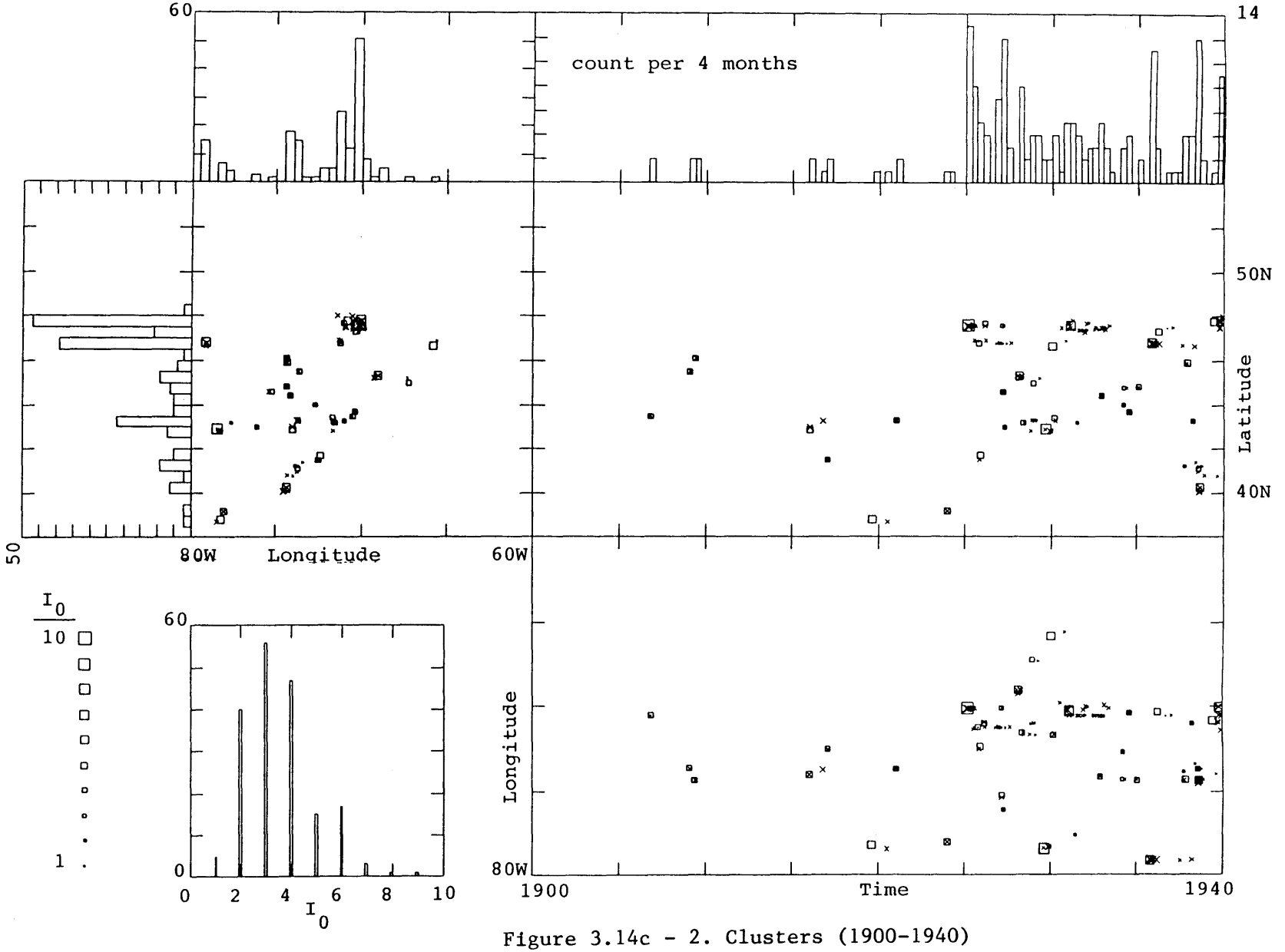


Figure 3.14c - 2. Clusters (1900-1940)

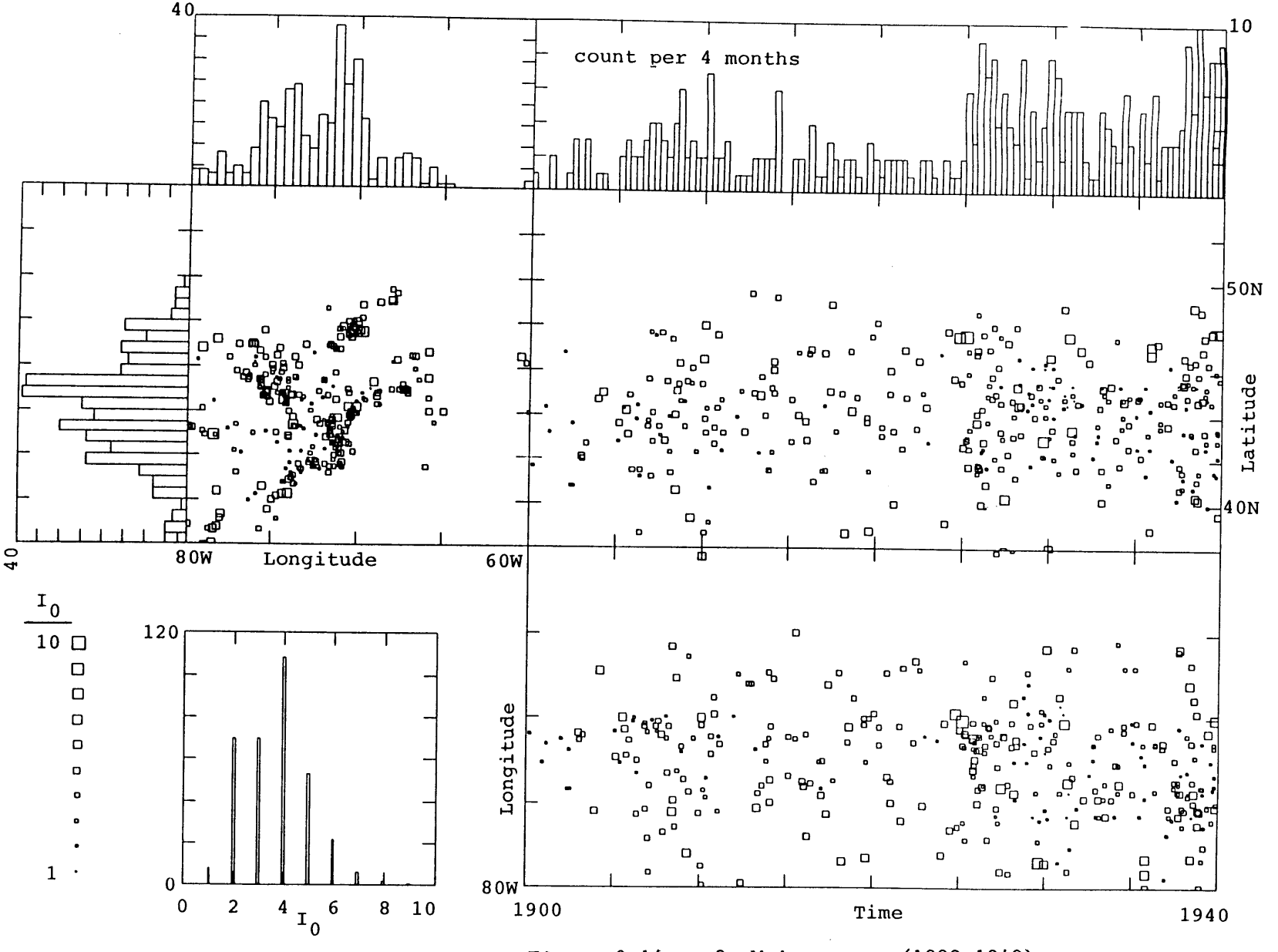


Figure 3.14c - 3. Main events (1900-1940)

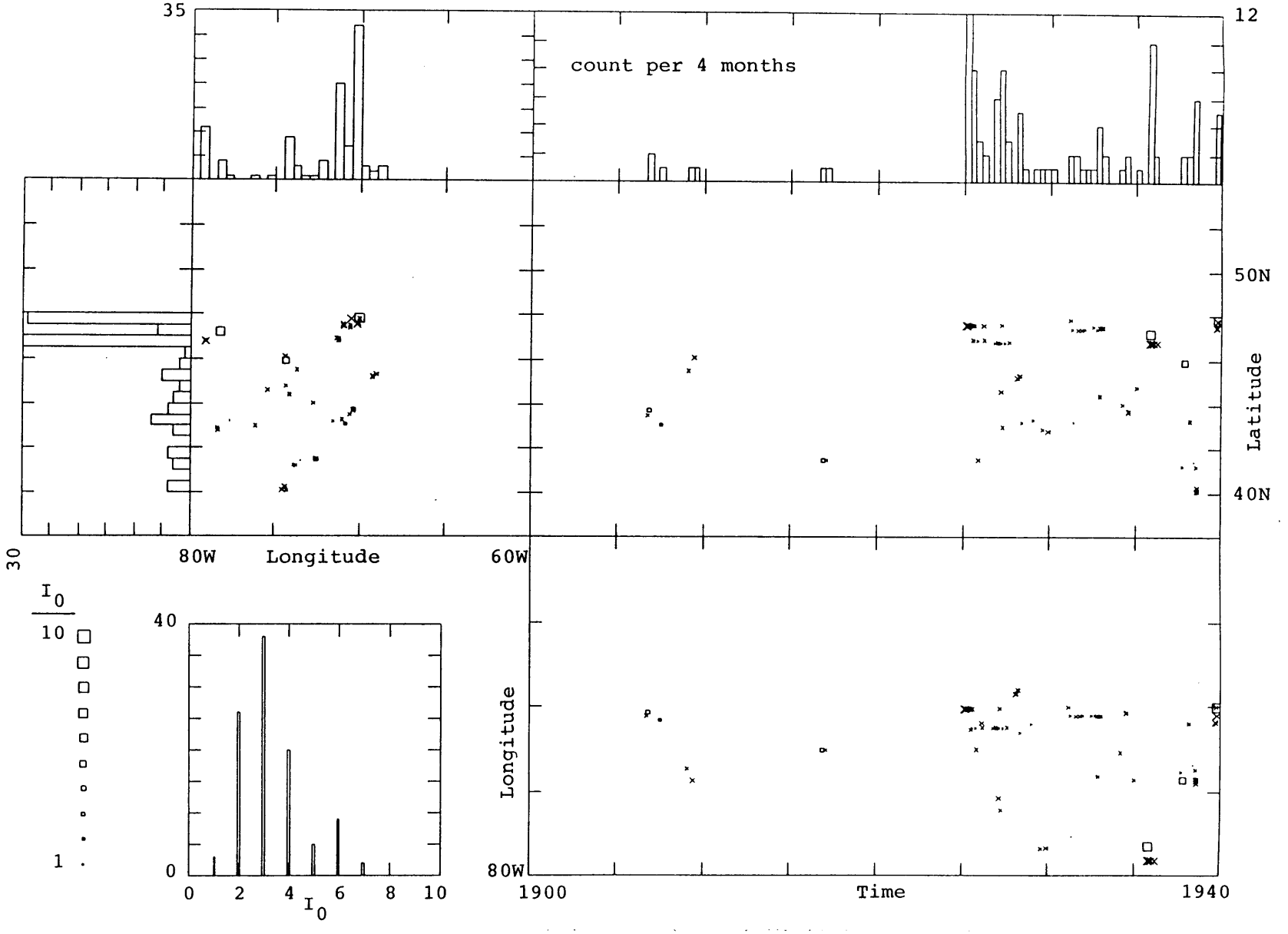


Figure 3.14c - 4. Judgemental aftershocks (1900-1940)

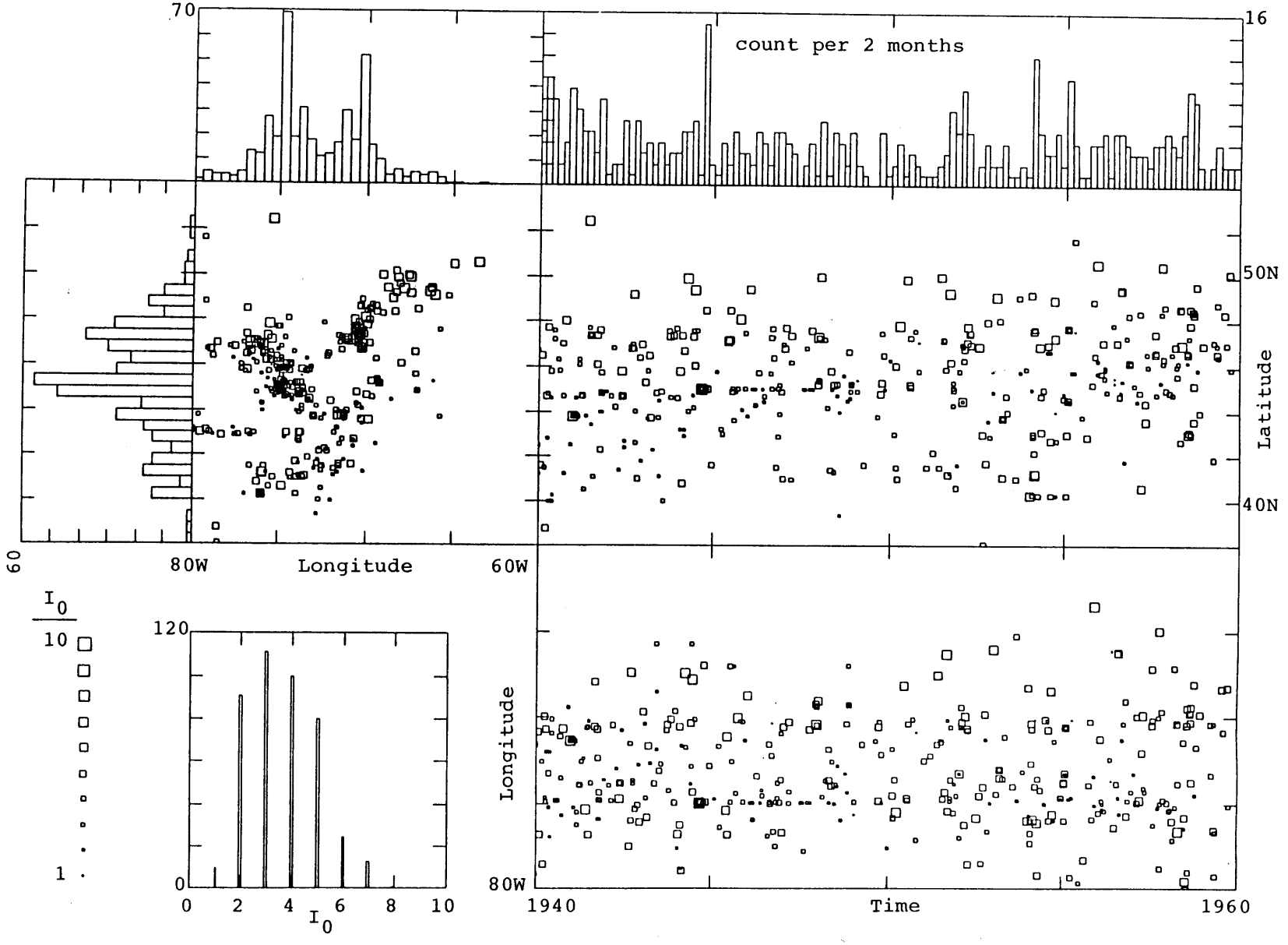


Figure 3.14d - 1. All events (1940-1960)

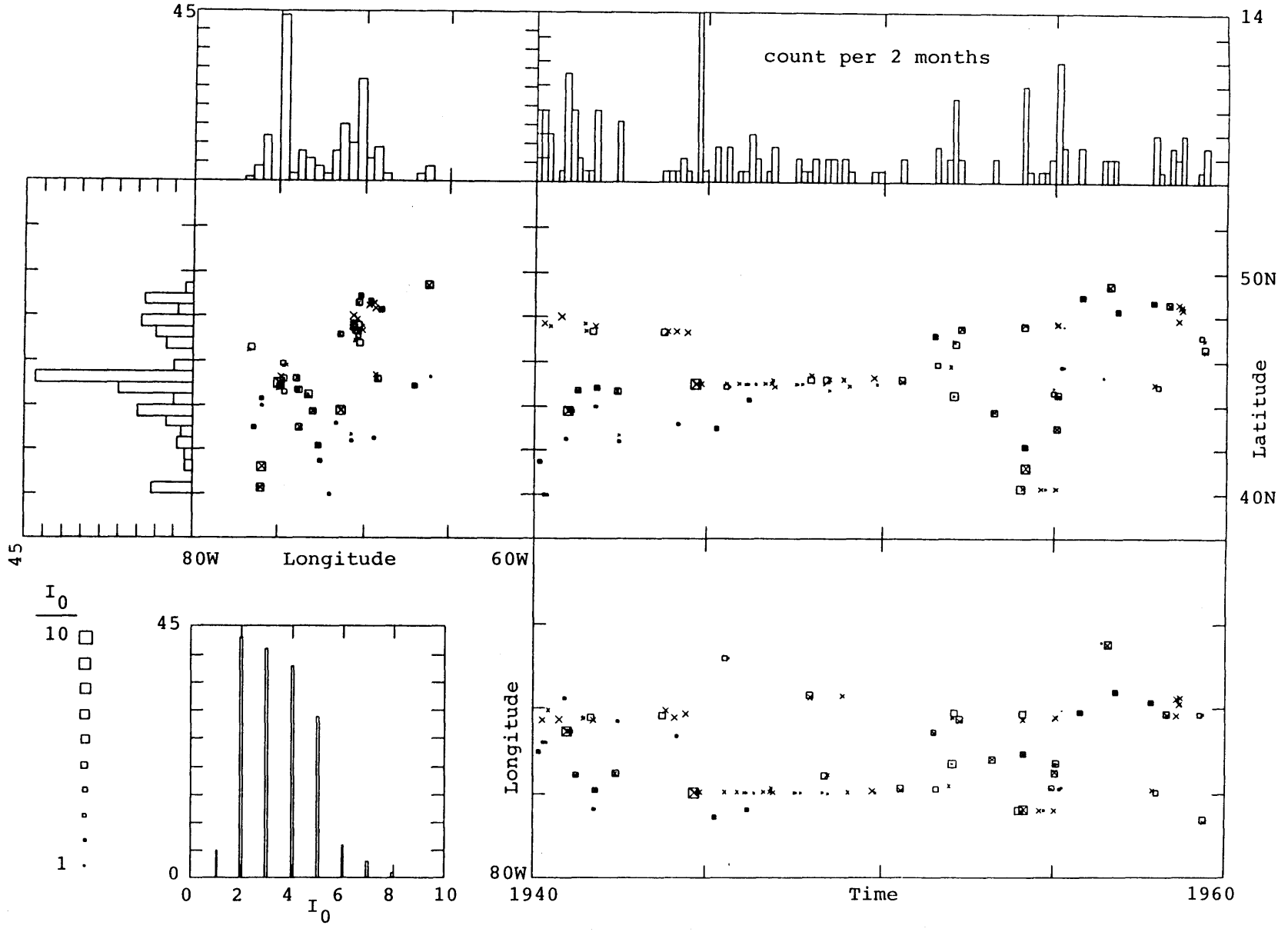


Figure 3.14d - 2. Clusters (1940-1960)

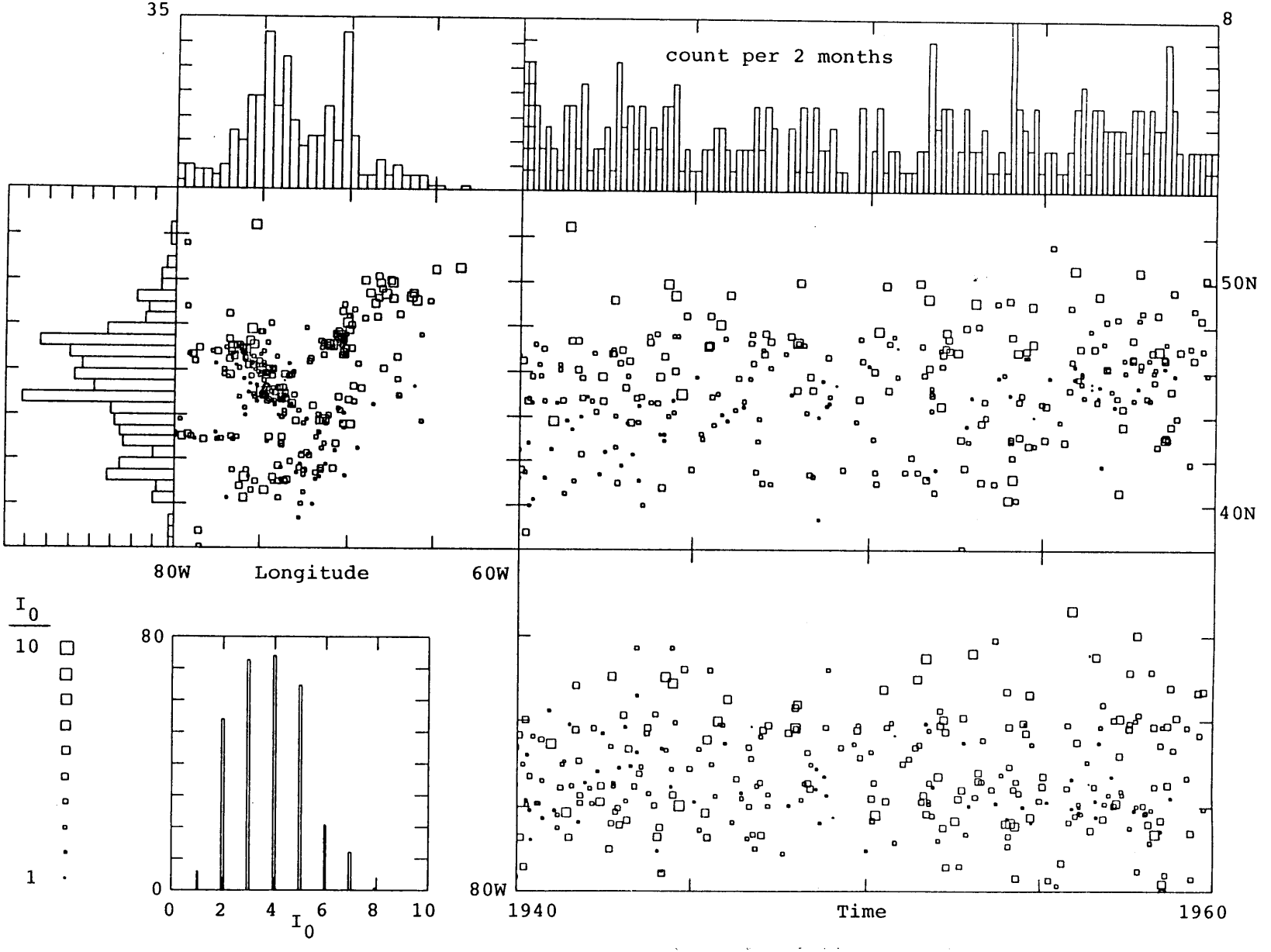


Figure 3.14d - 3. Main events (1940-1960)

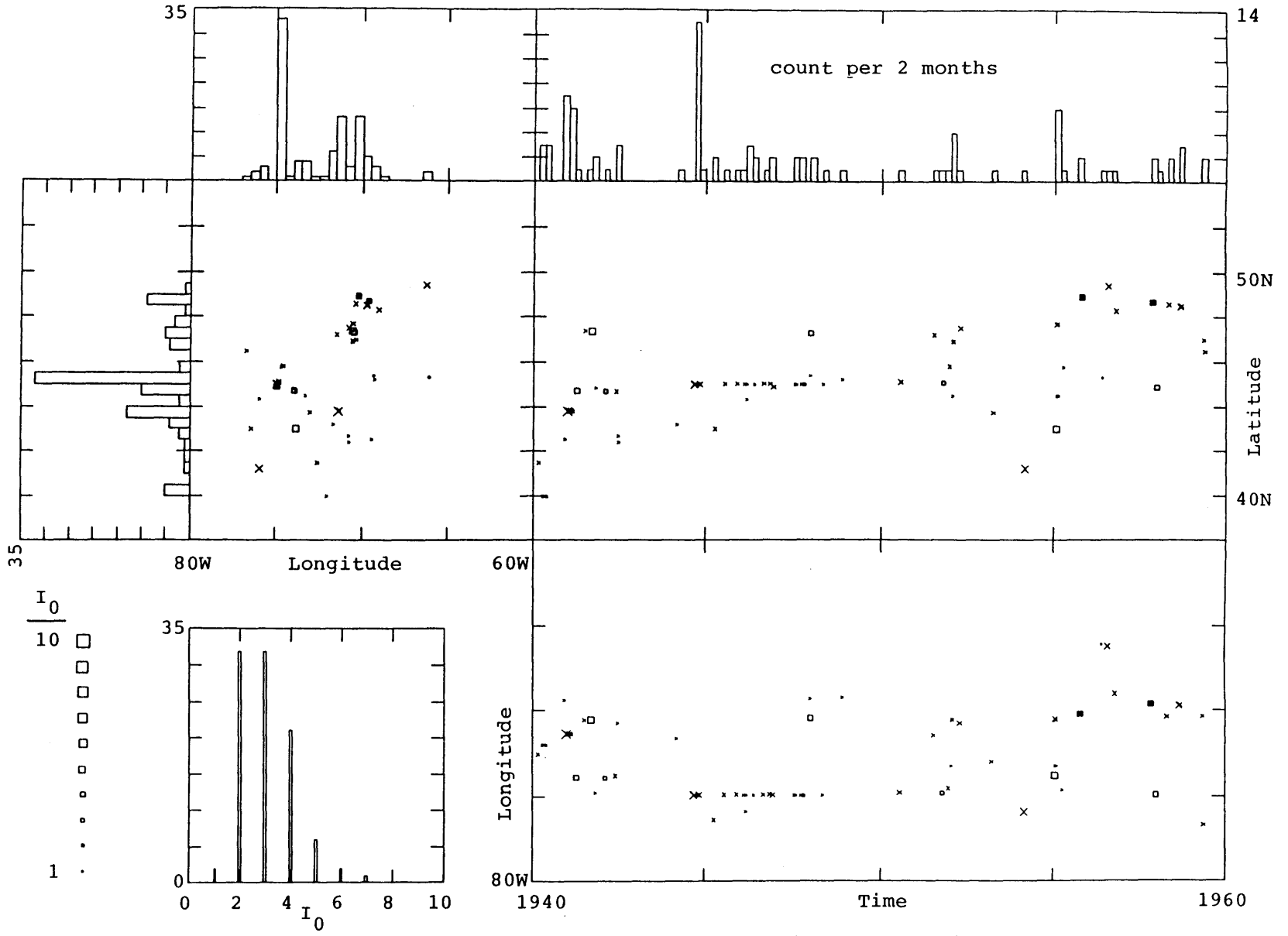


Figure 3.14d - 4. Judgemental aftershocks (1940-1960)

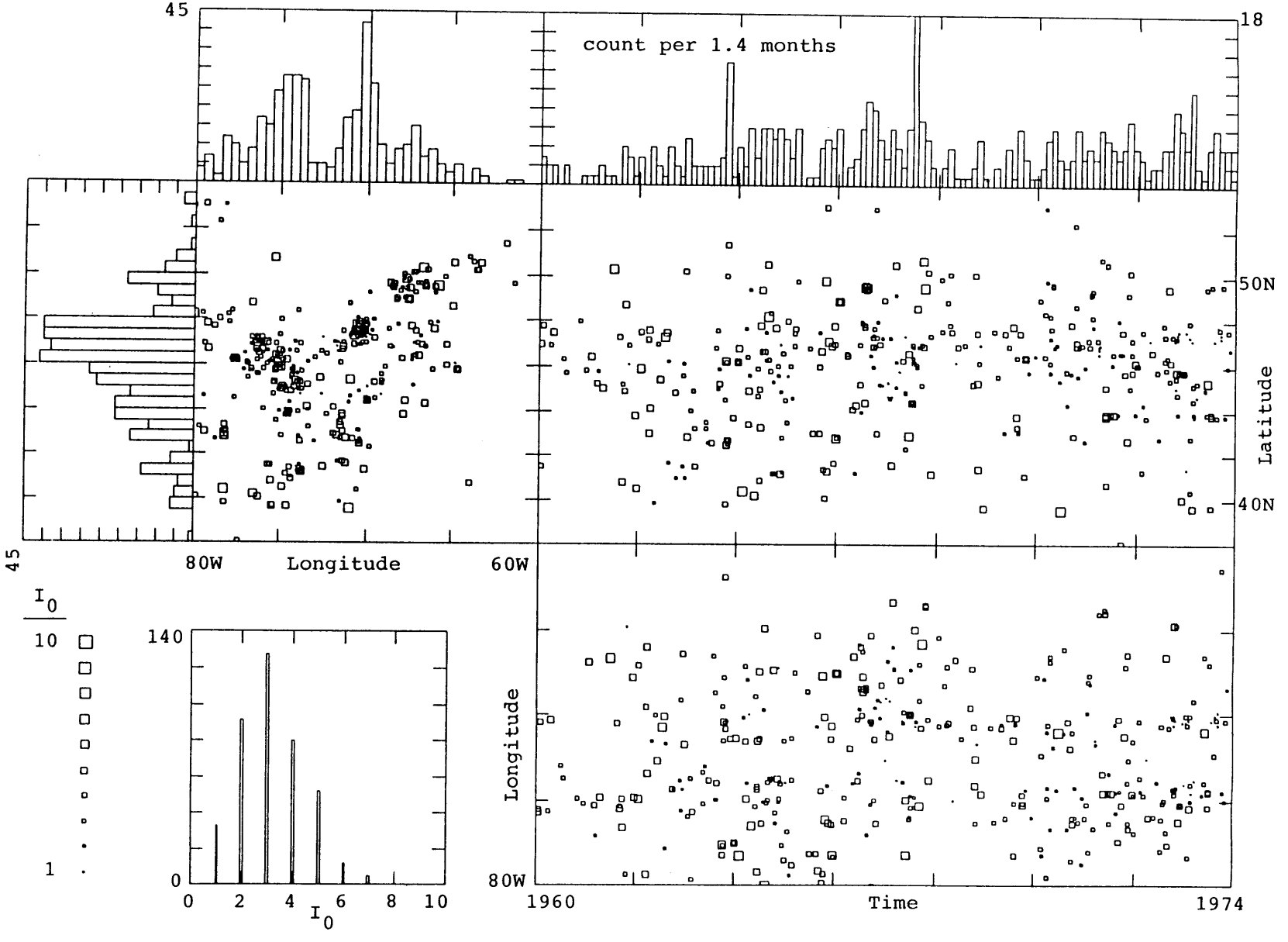


Figure 3.14e - 1. All events (1960-1974)

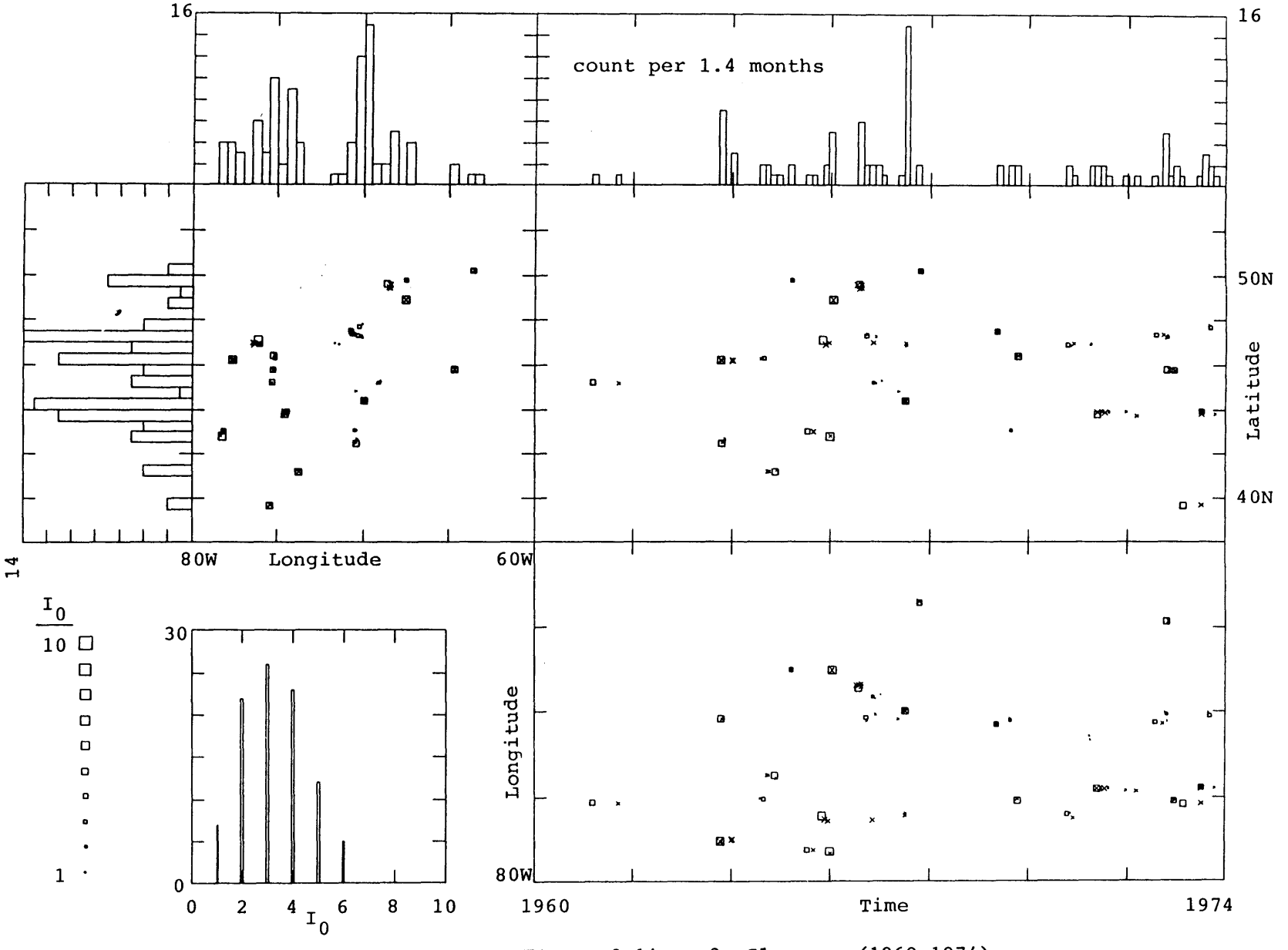


Figure 3.14e - 2. Clusters (1960-1974)

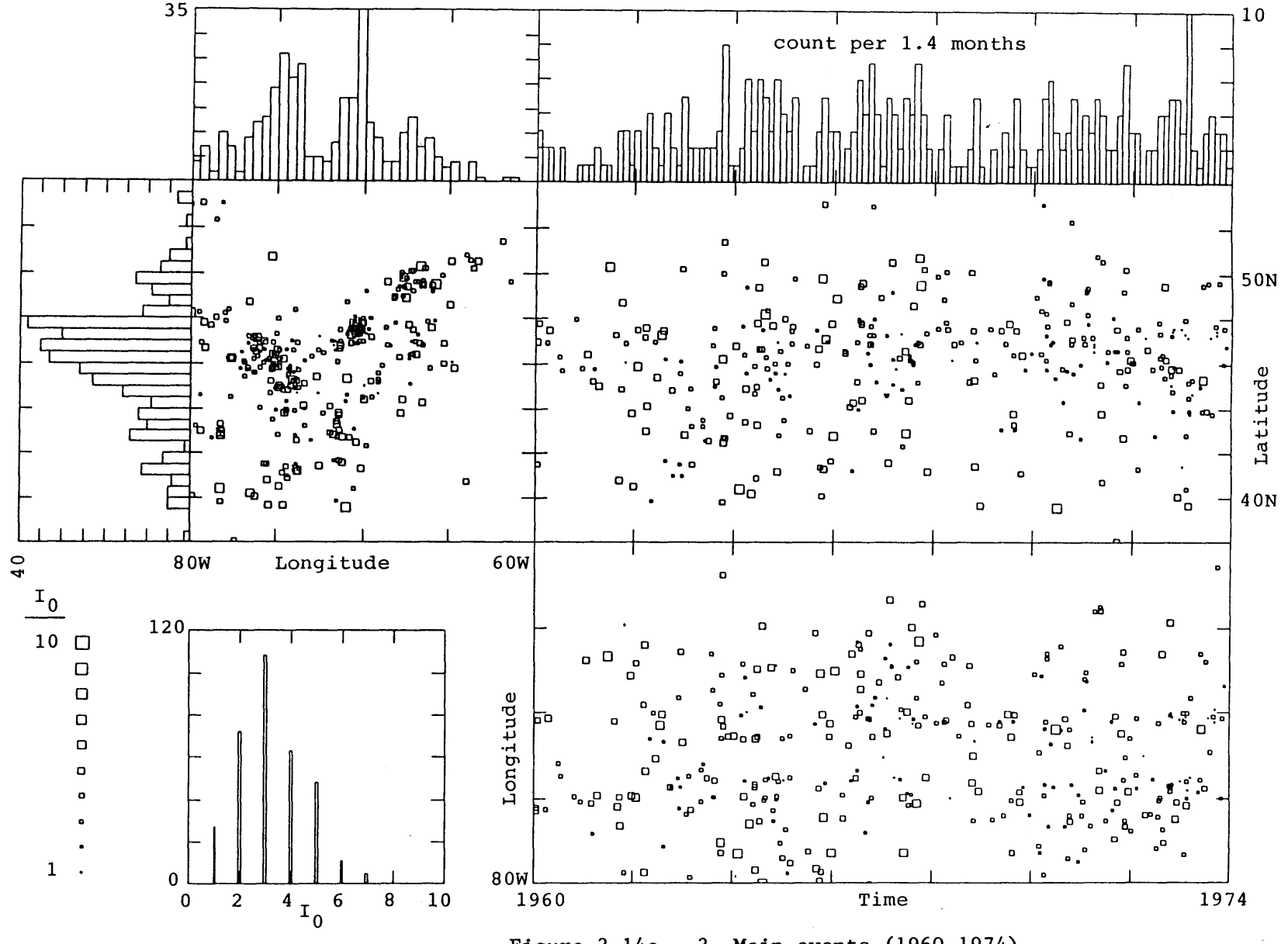


Figure 3.14e - 3. Main events (1960-1974)

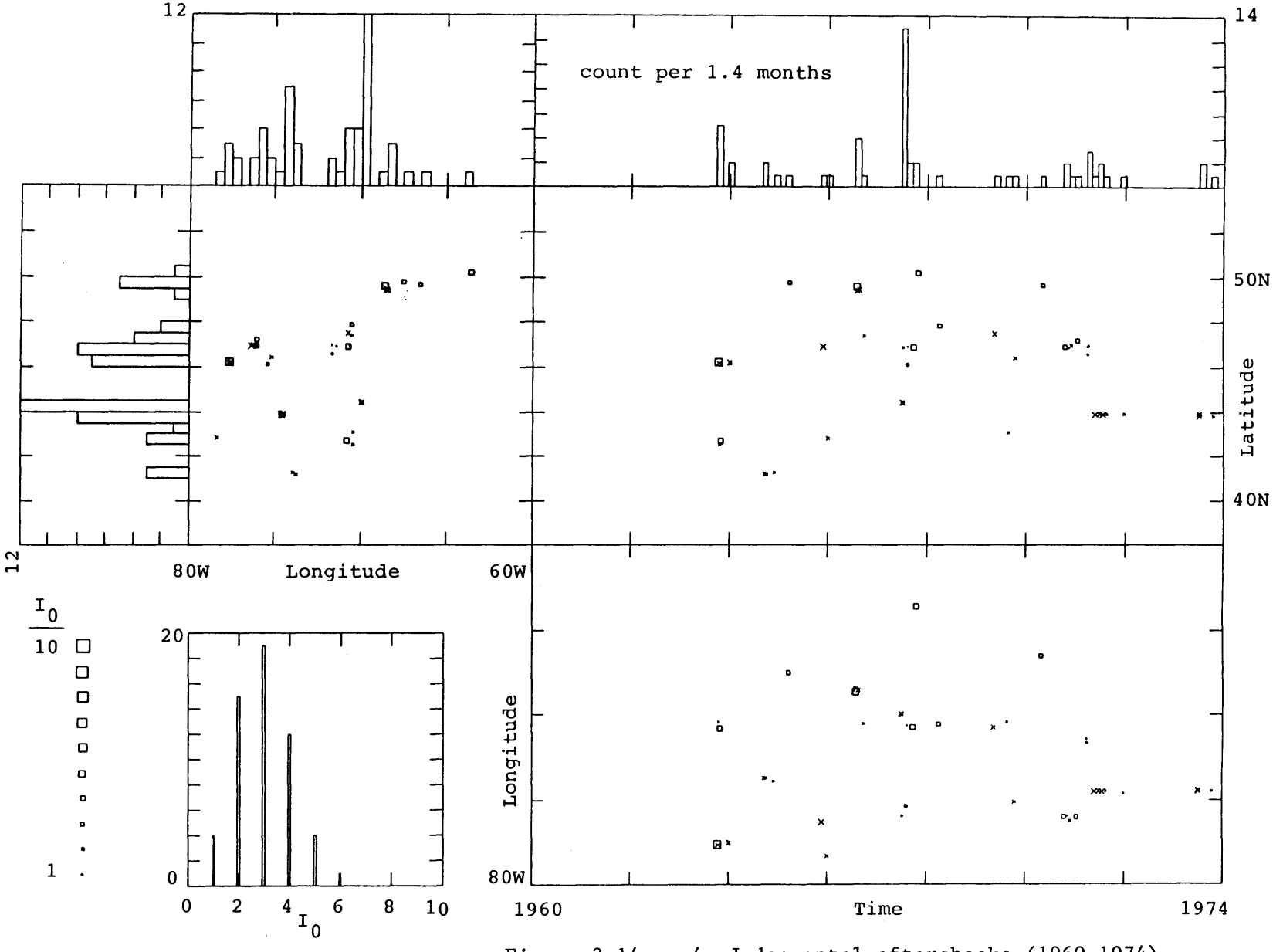


Figure 3.14e - 4. Judgemental aftershocks (1960-1974)

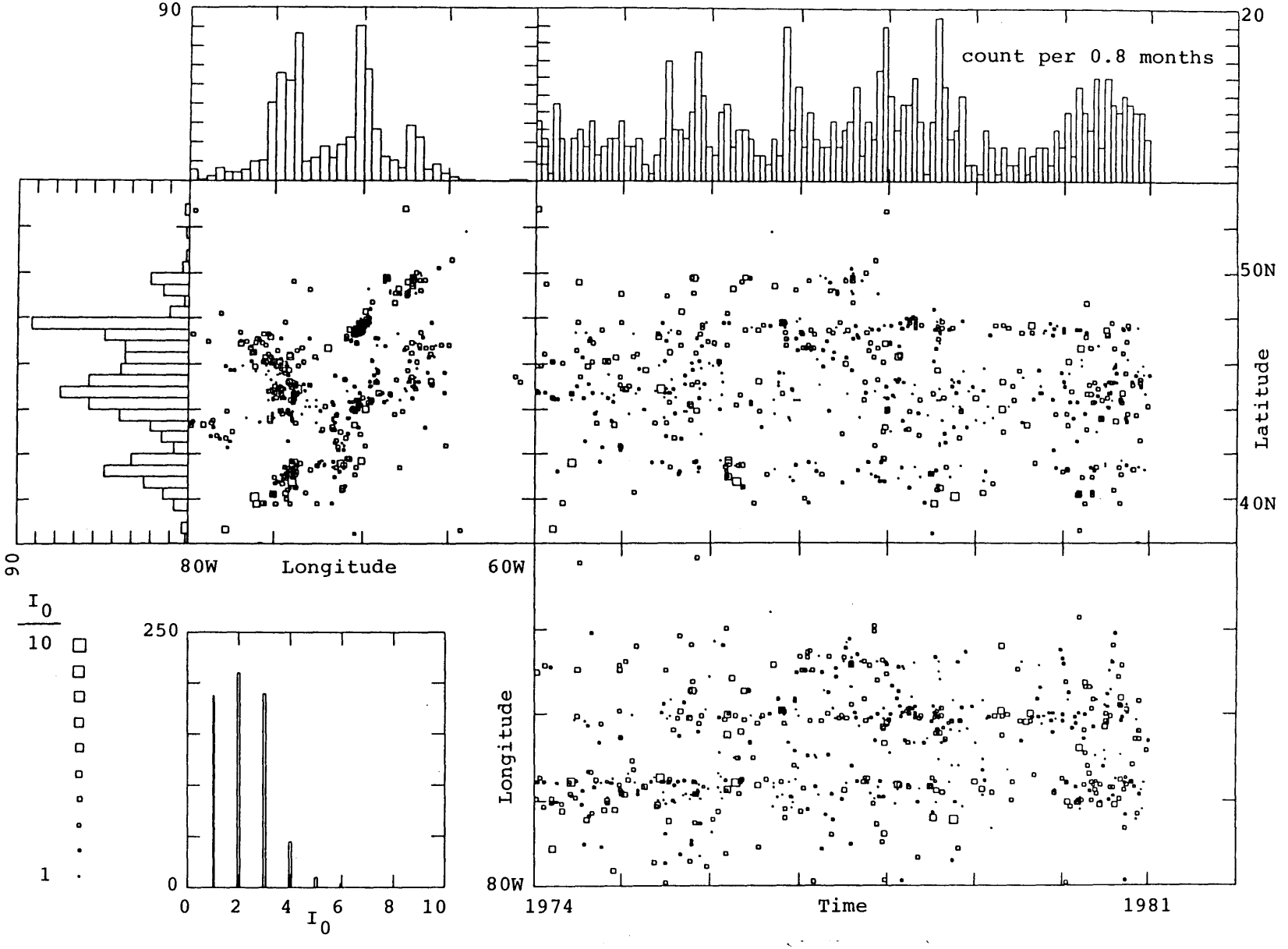


Figure 3.14f - 1. All events (1974-1981)

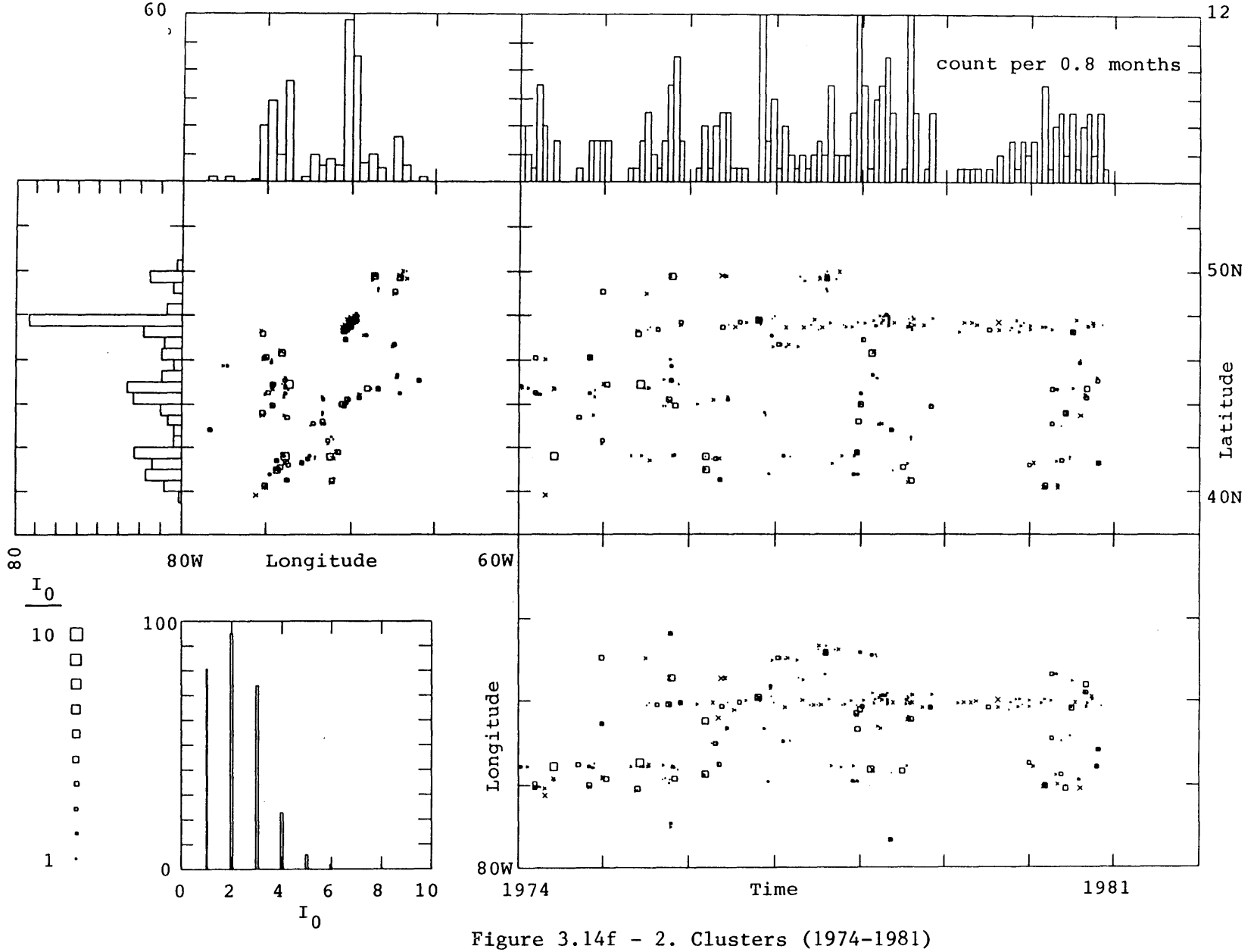


Figure 3.14f - 2. Clusters (1974-1981)

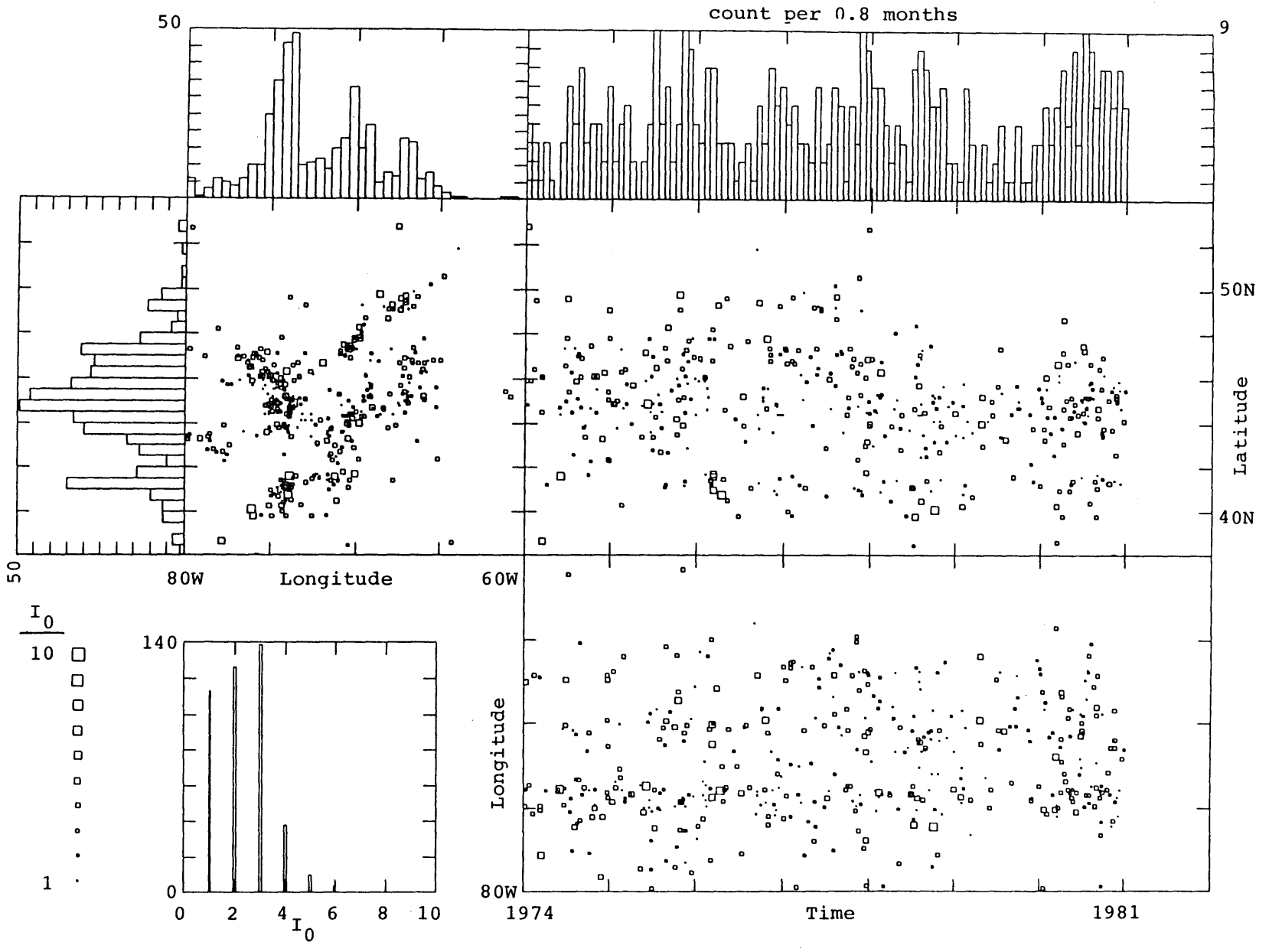


Figure 3.14f - 3. Main events (1974-1981)

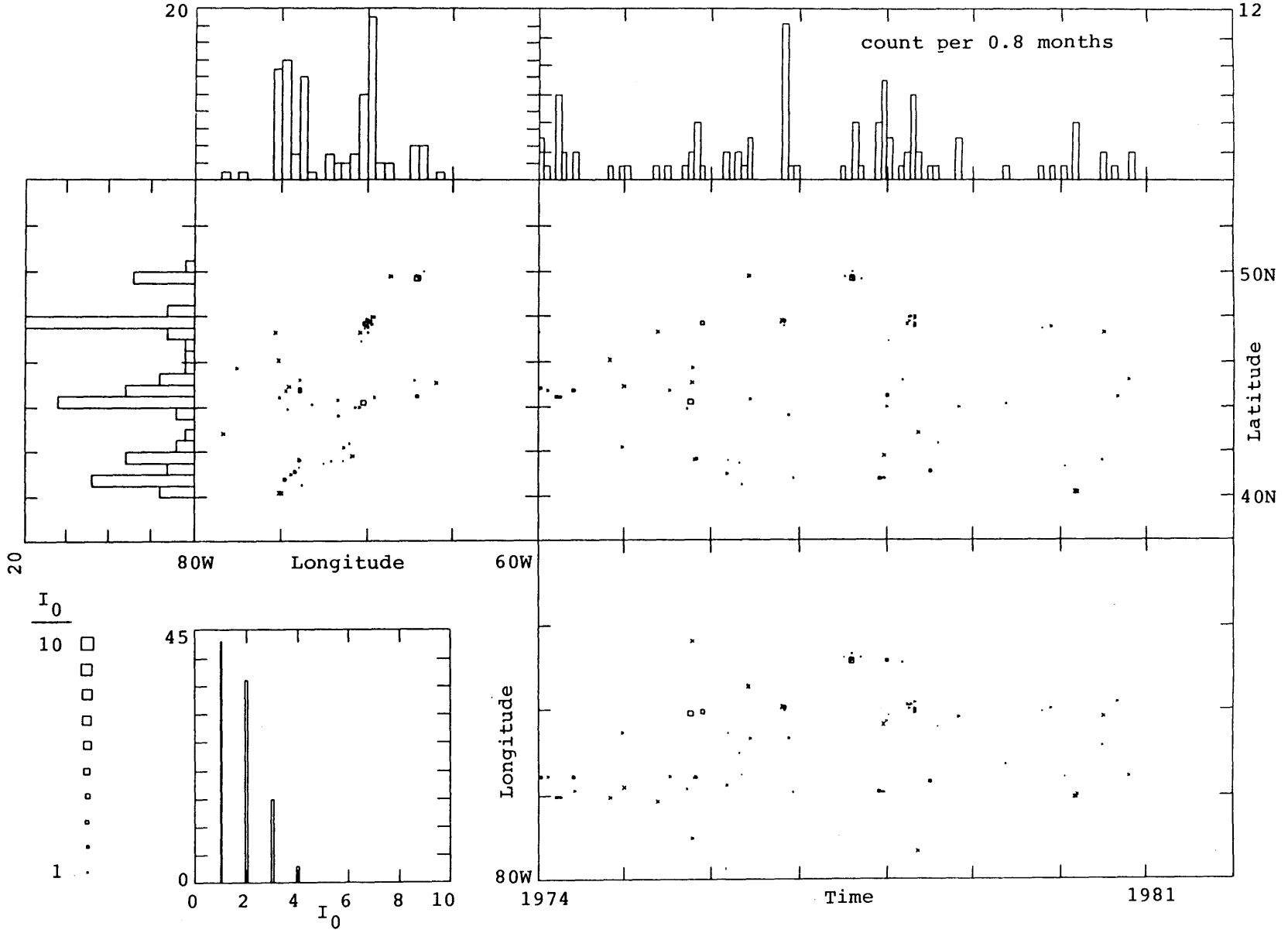
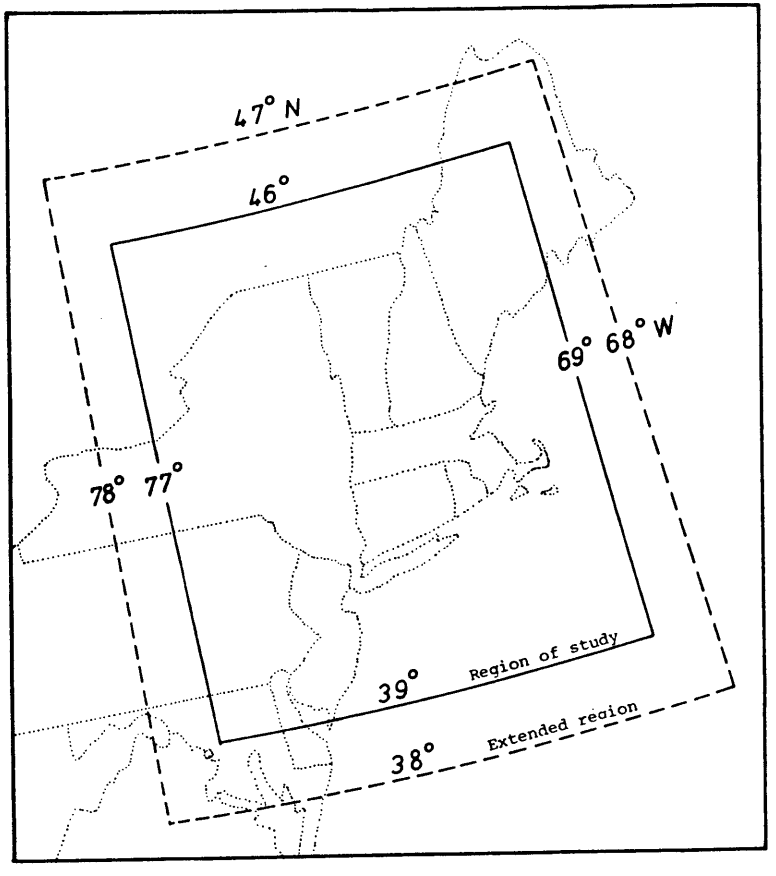
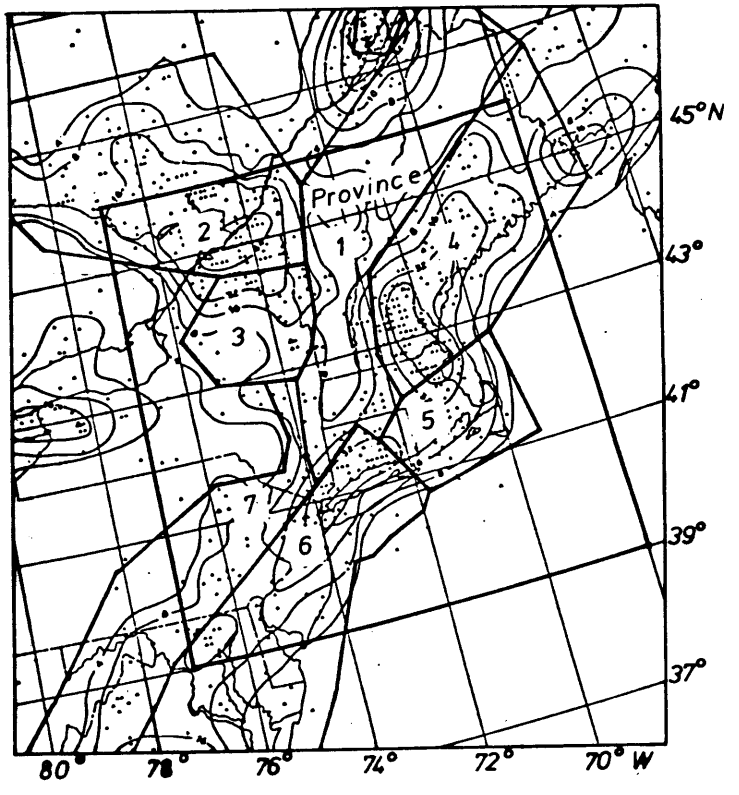


Figure 3.14f - 4. Judgemental aftershocks (1974-1981)



a) Models B and D



b) Seismogenic Provinces (from YAEC, 1981) used in Model A

Figure 4.1 - Region of Study

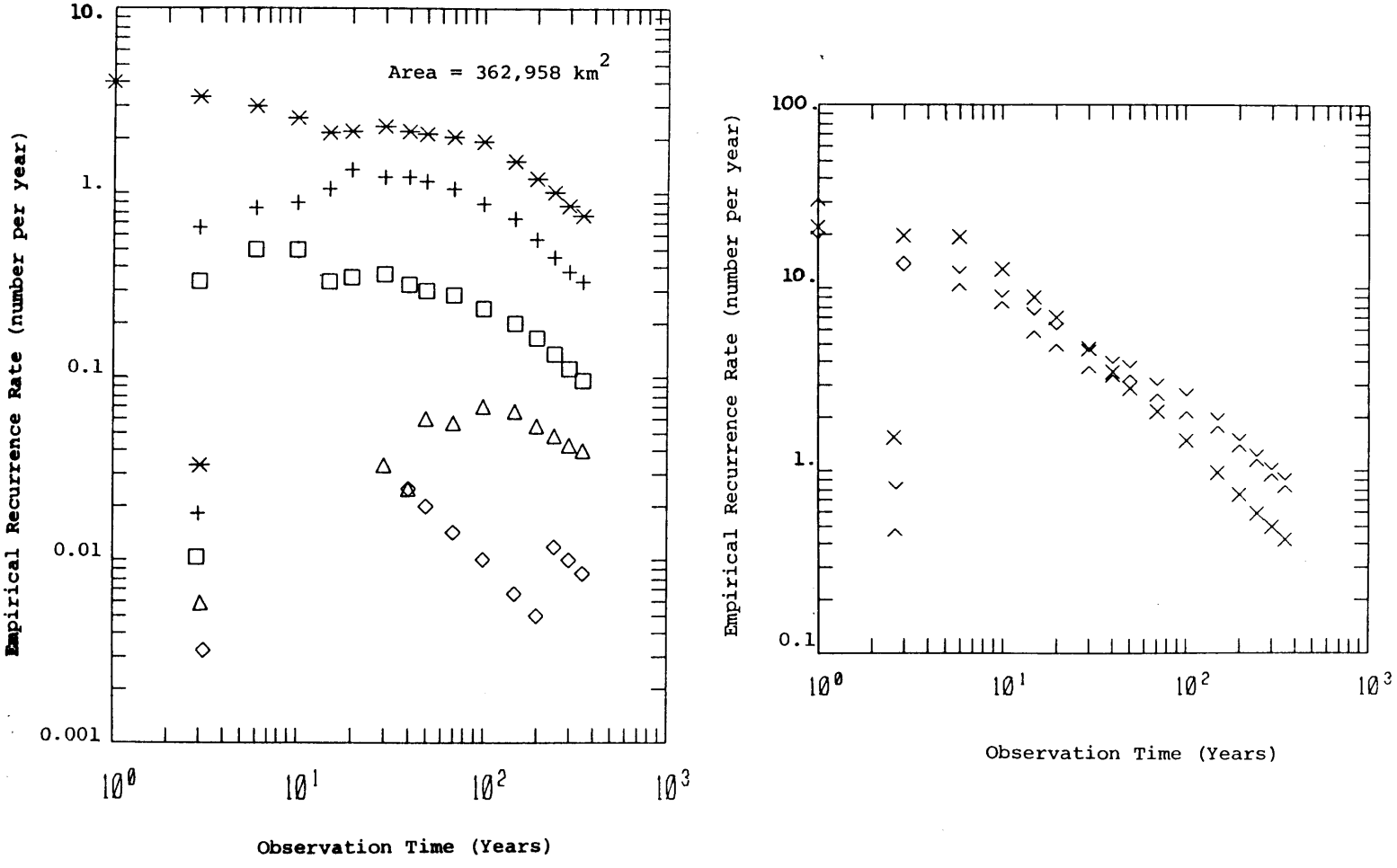


Figure 4.2a - Empirical recurrence rate versus observation time for all provinces

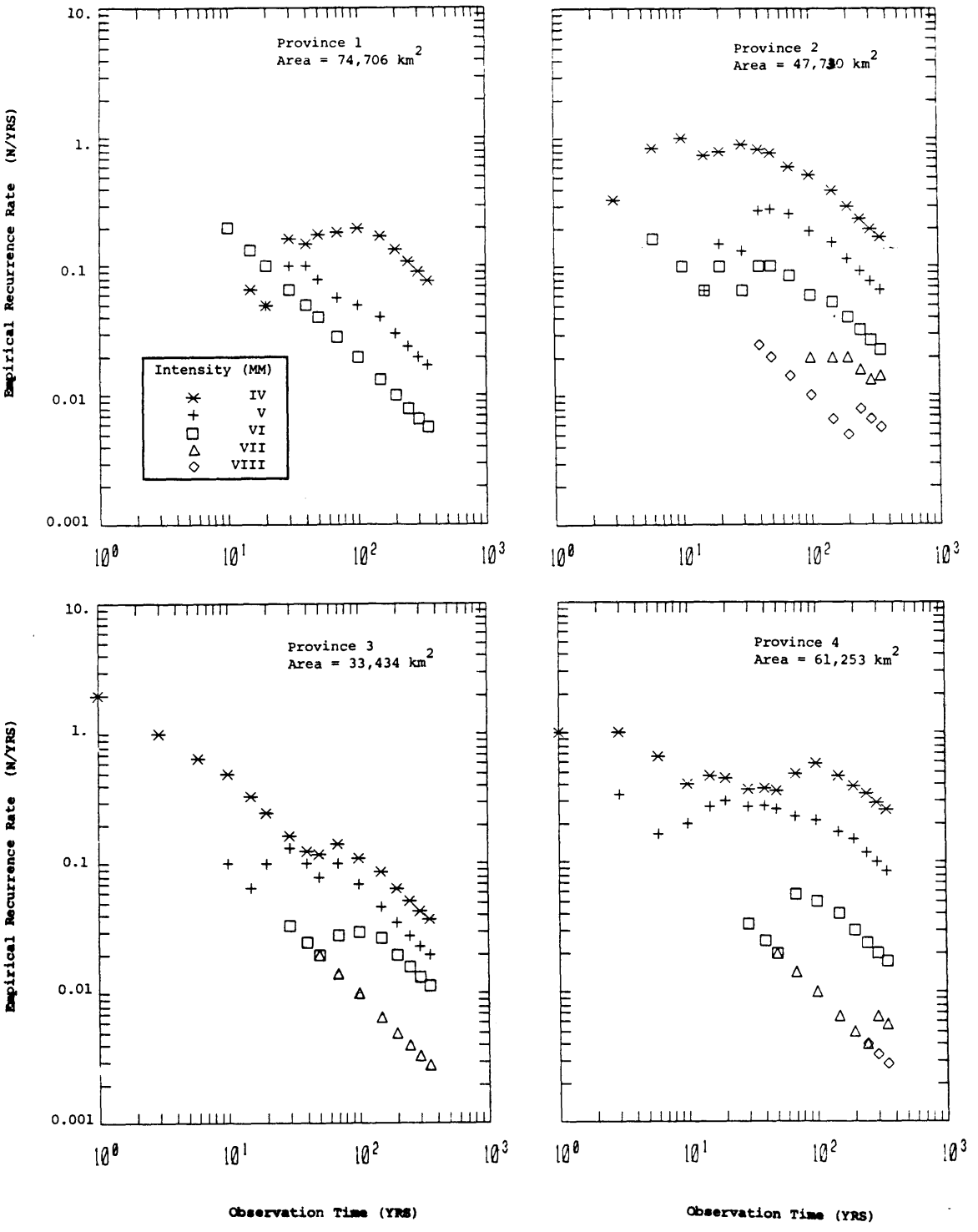


Figure 4.2b - Empirical recurrence rate versus observation time for individual provinces

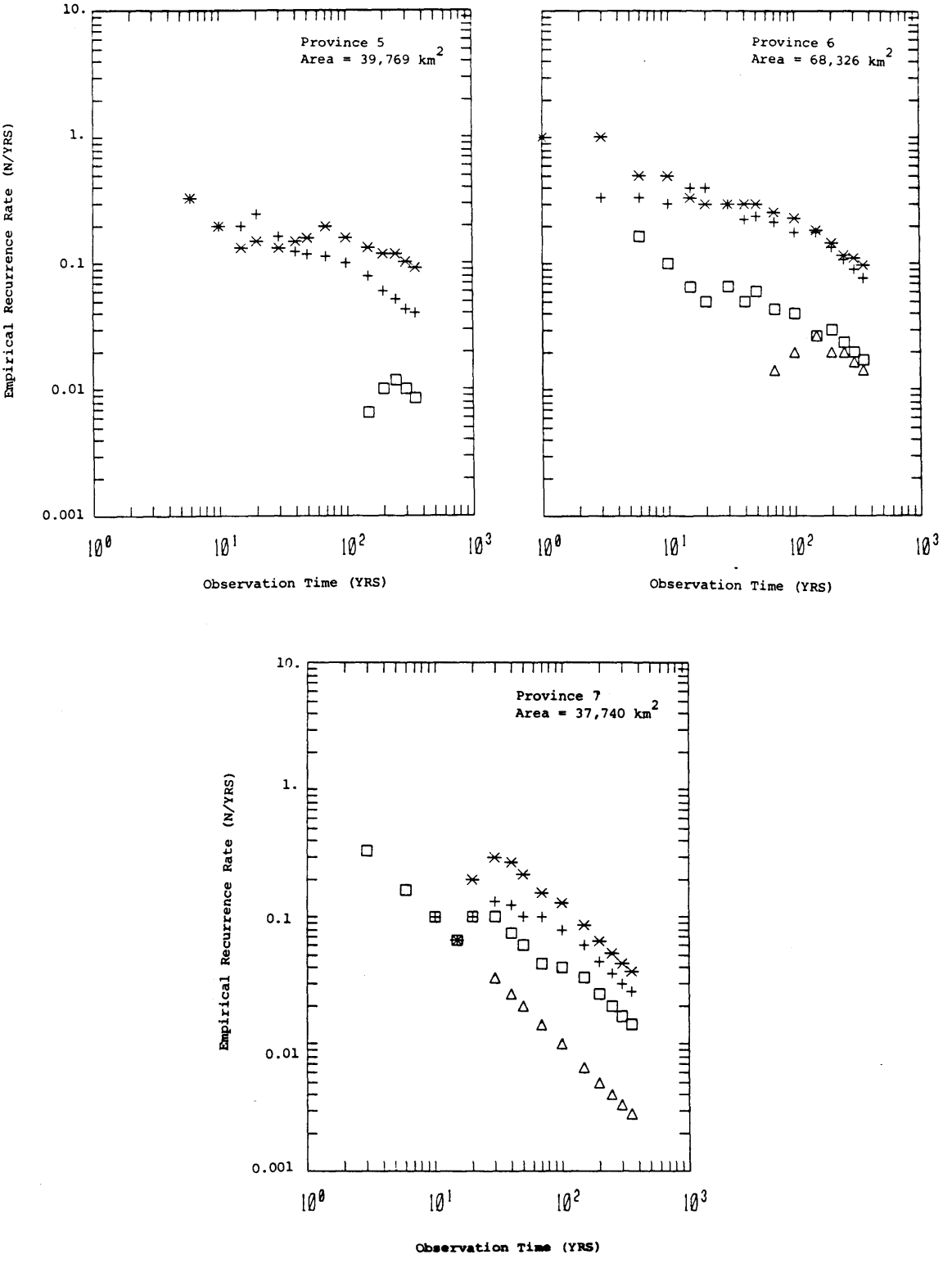


Figure 4.2b - (End)

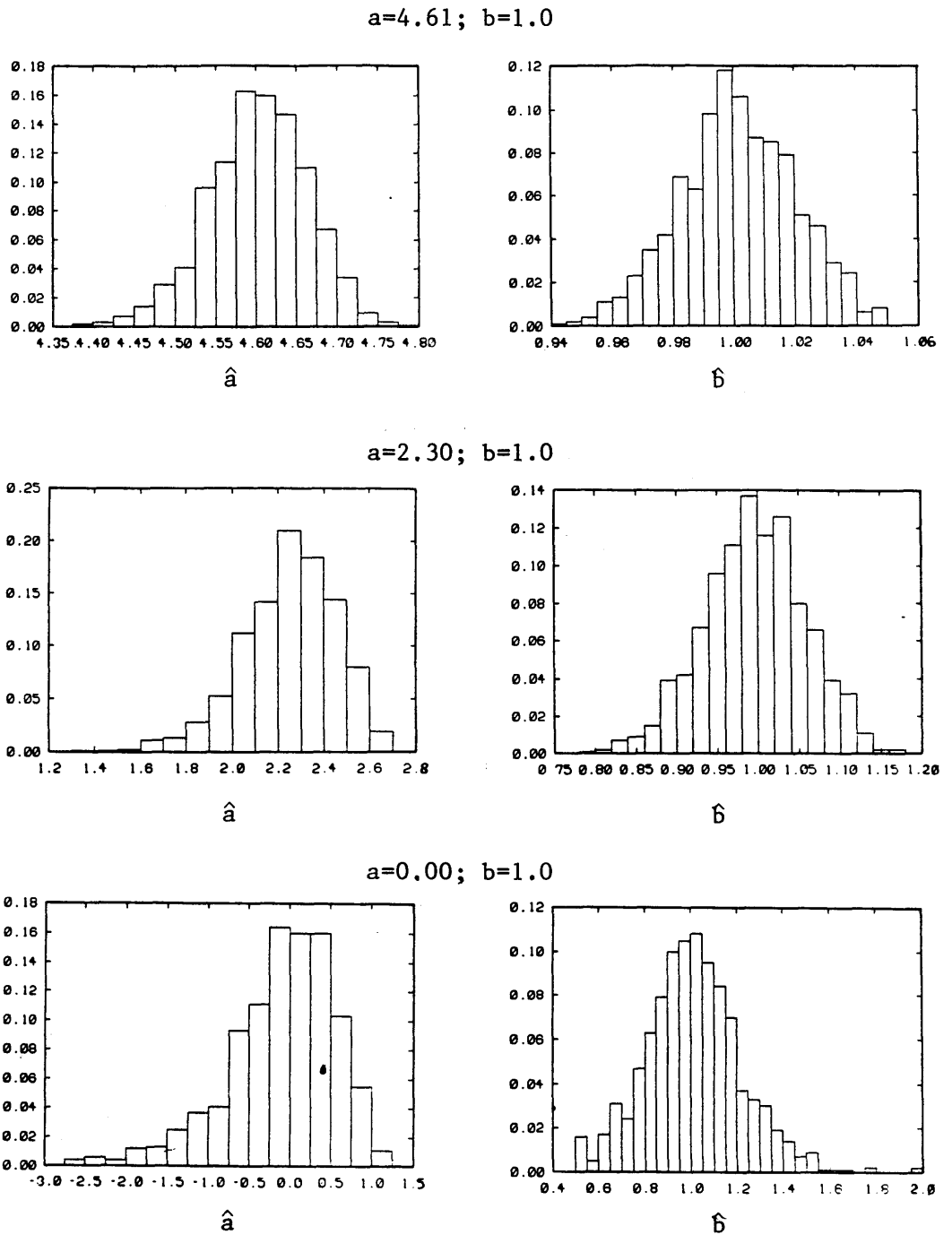


Figure 4.3a - Histograms of parameter estimates in Eq. 4.3 for unequal periods of observation (Case A)

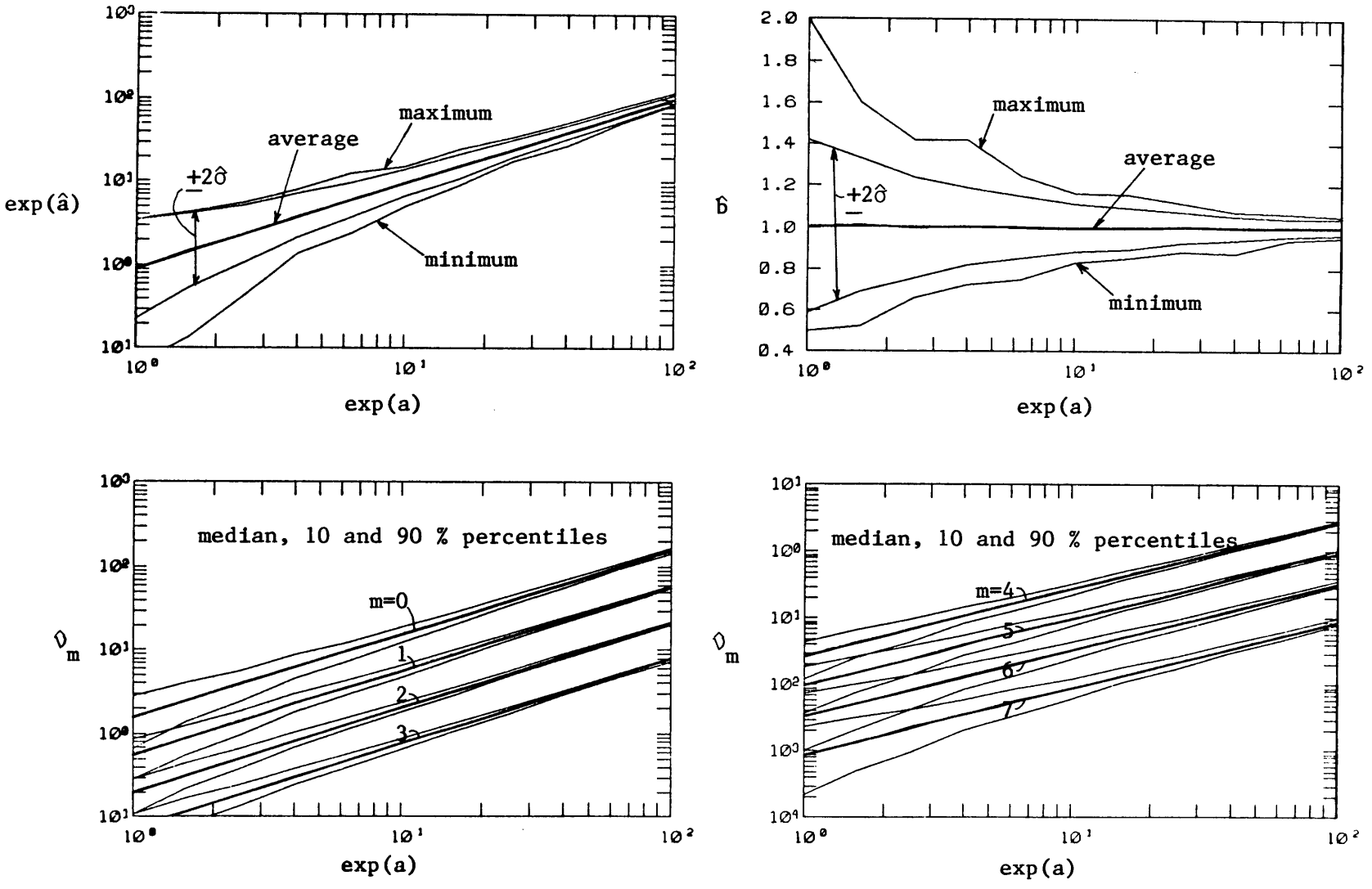


Figure 4.3b - Distribution of recurrence parameters and rate estimates (Case A) as a function of the true recurrence rate at m=0, exp(a)

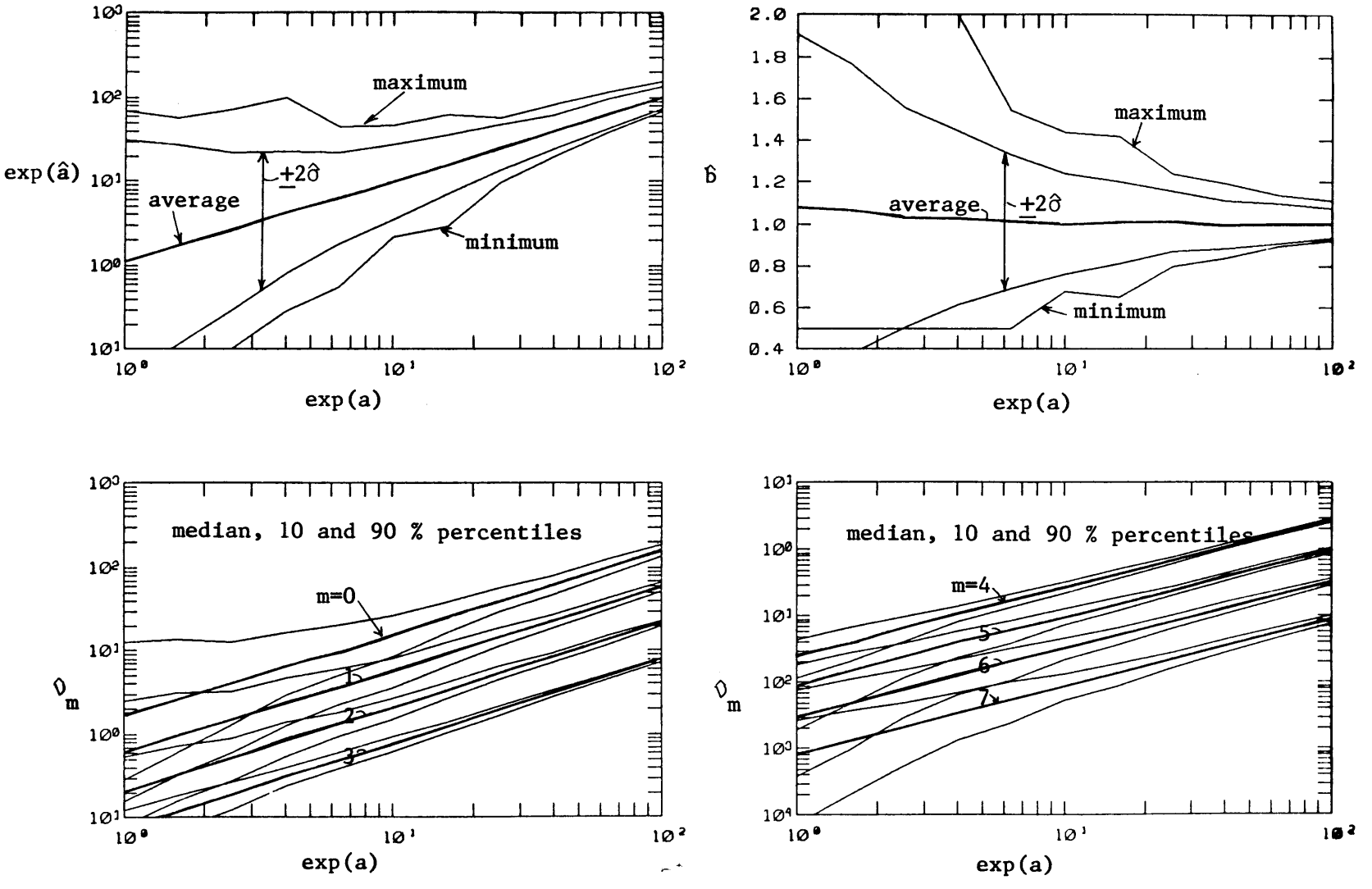


Figure 4.3c - Distribution of recurrence parameters and rate estimates (Case B) as a function of the true recurrence rate at m=0, exp(a)

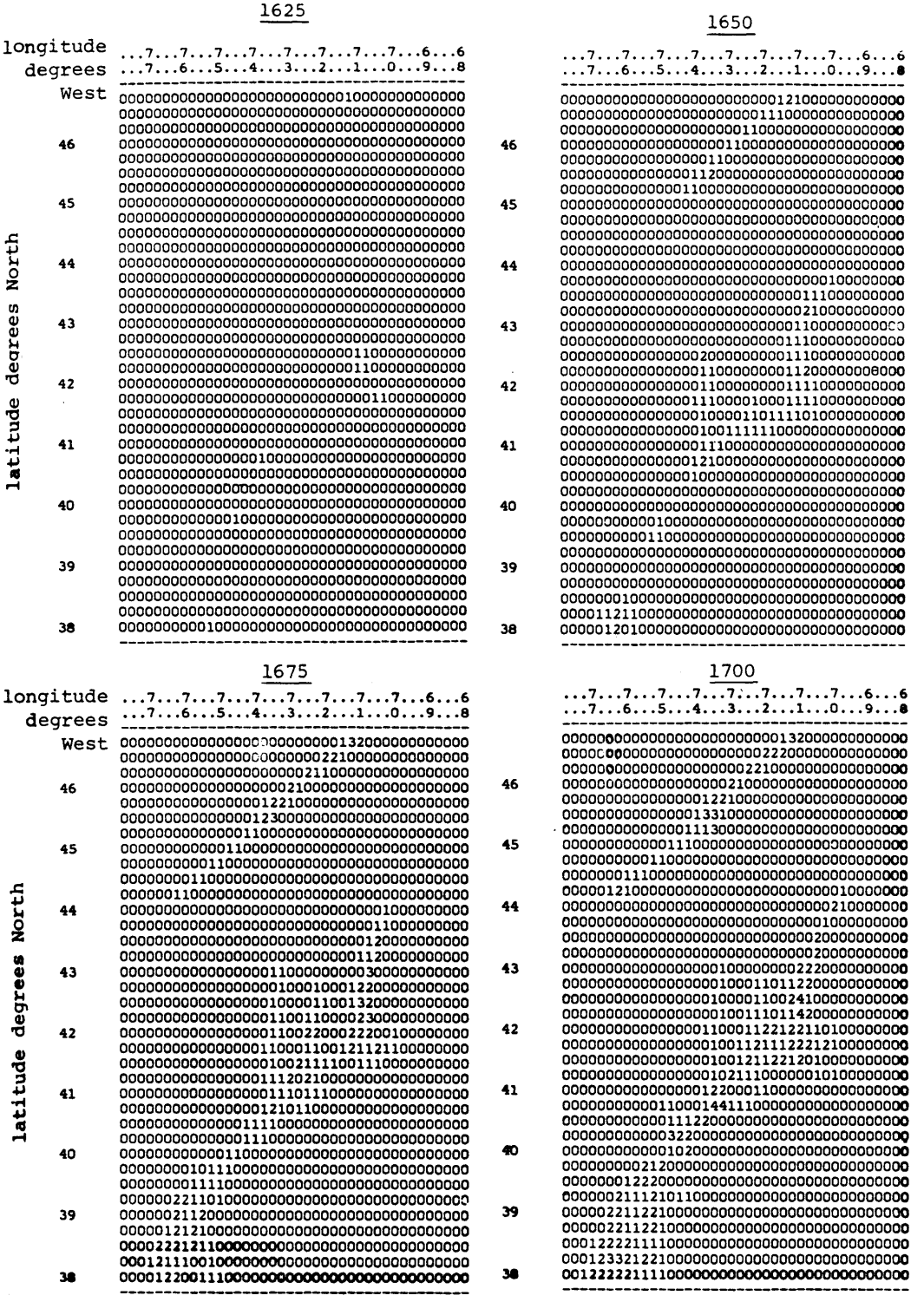


Figure 4.4 - Discretized population maps

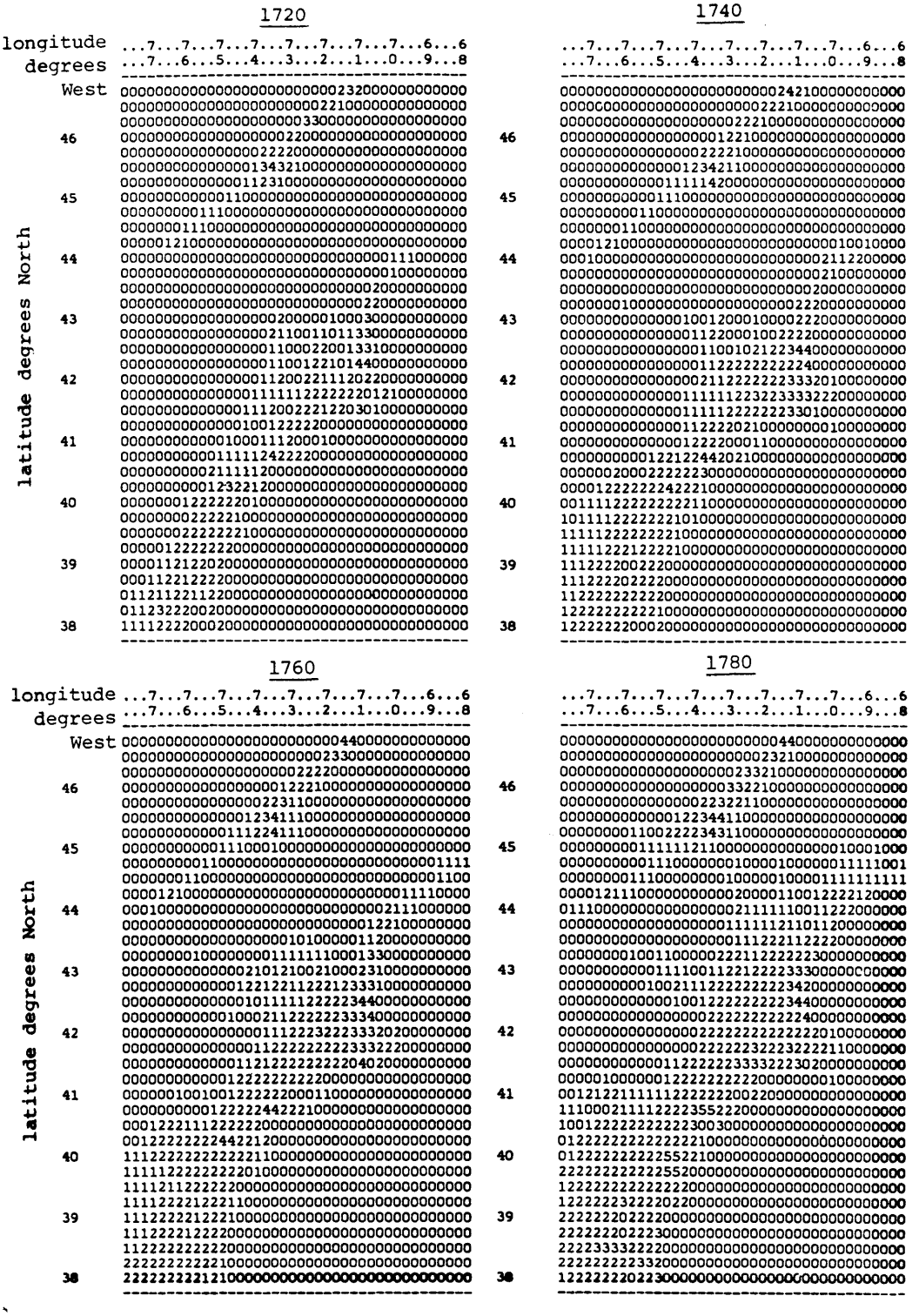


Figure 4.4 - (Continued)

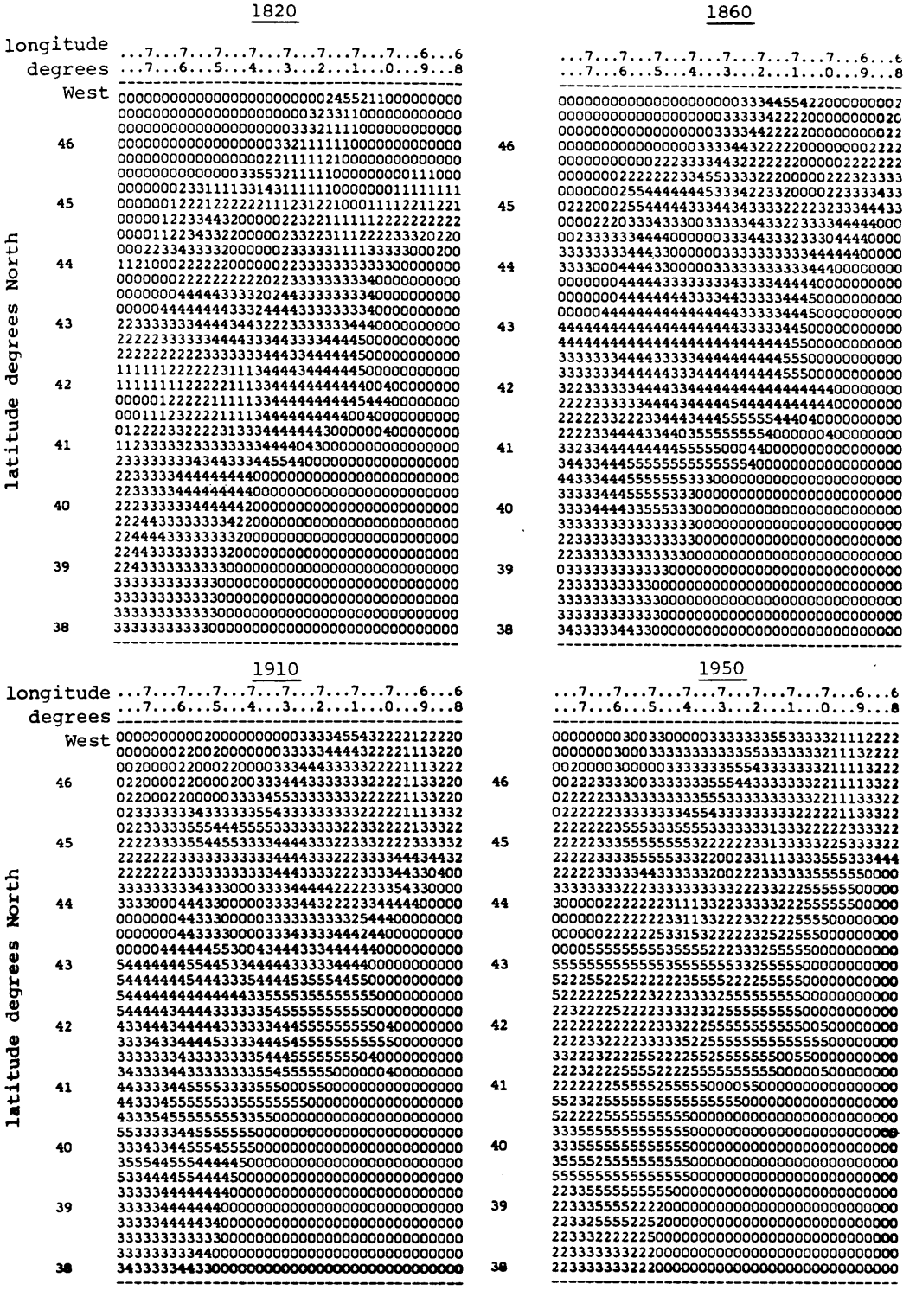


Figure 4.4 - (End)

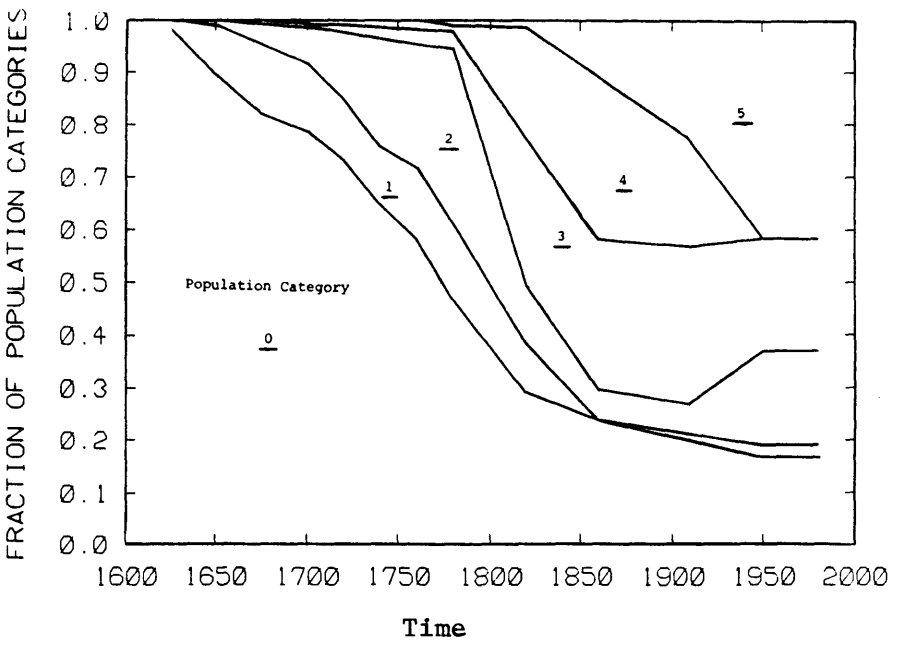


Figure 4.5 - Cumulative fraction of area associated with each population category as a function of time

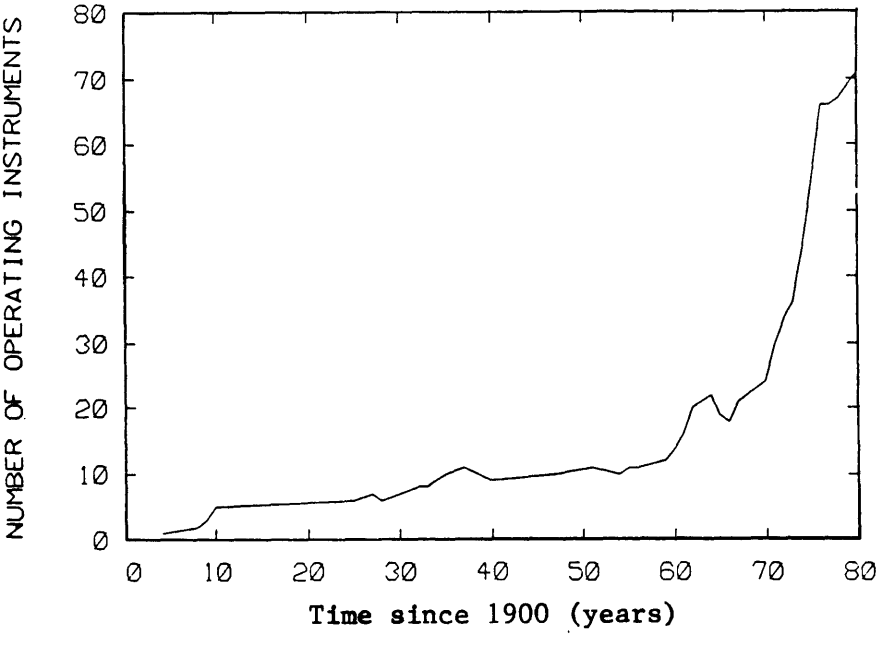


Figure 4.6 - Number of operating seismic instruments as a function of time

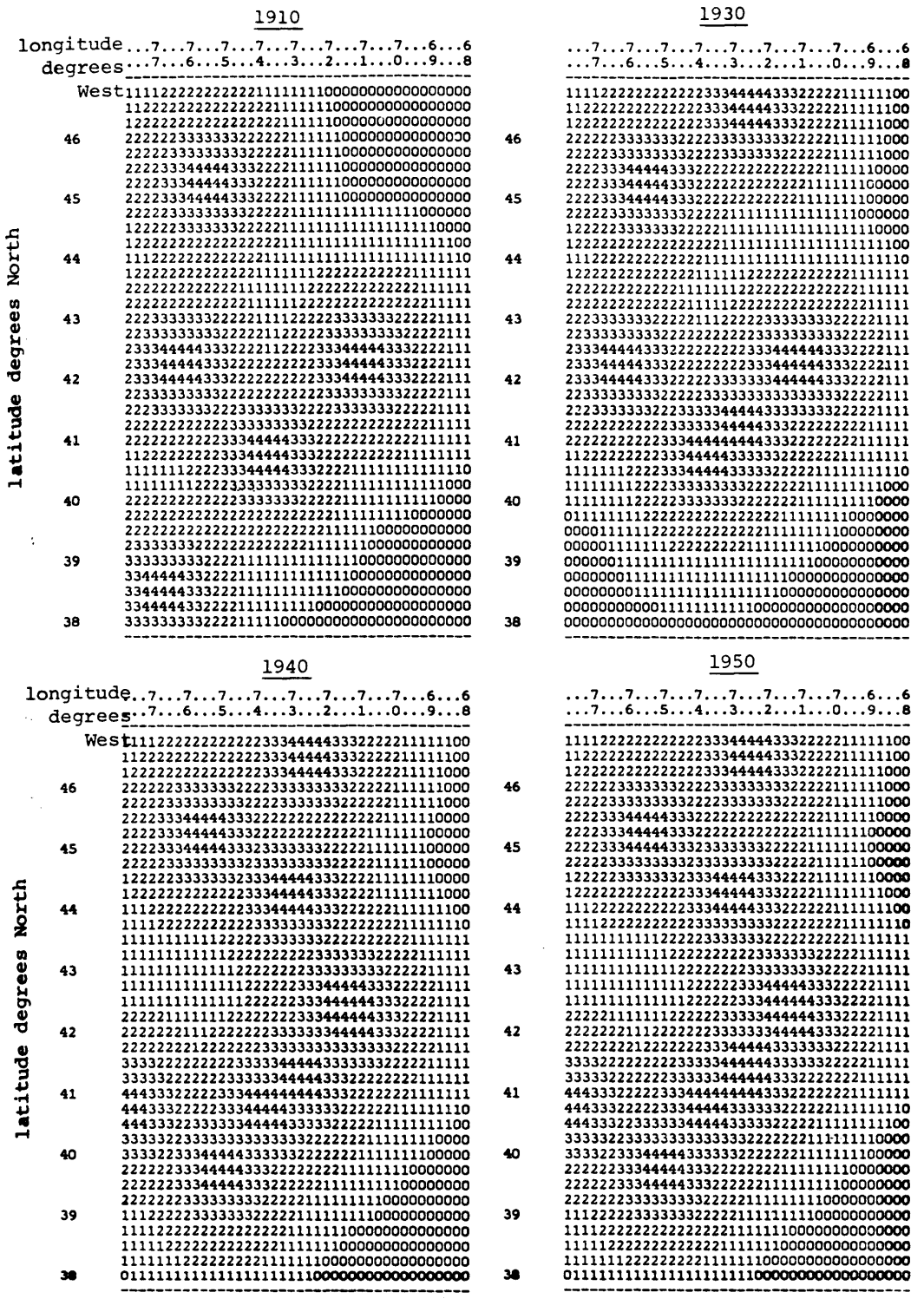


Figure 4.7 - Discretized instrumentation maps

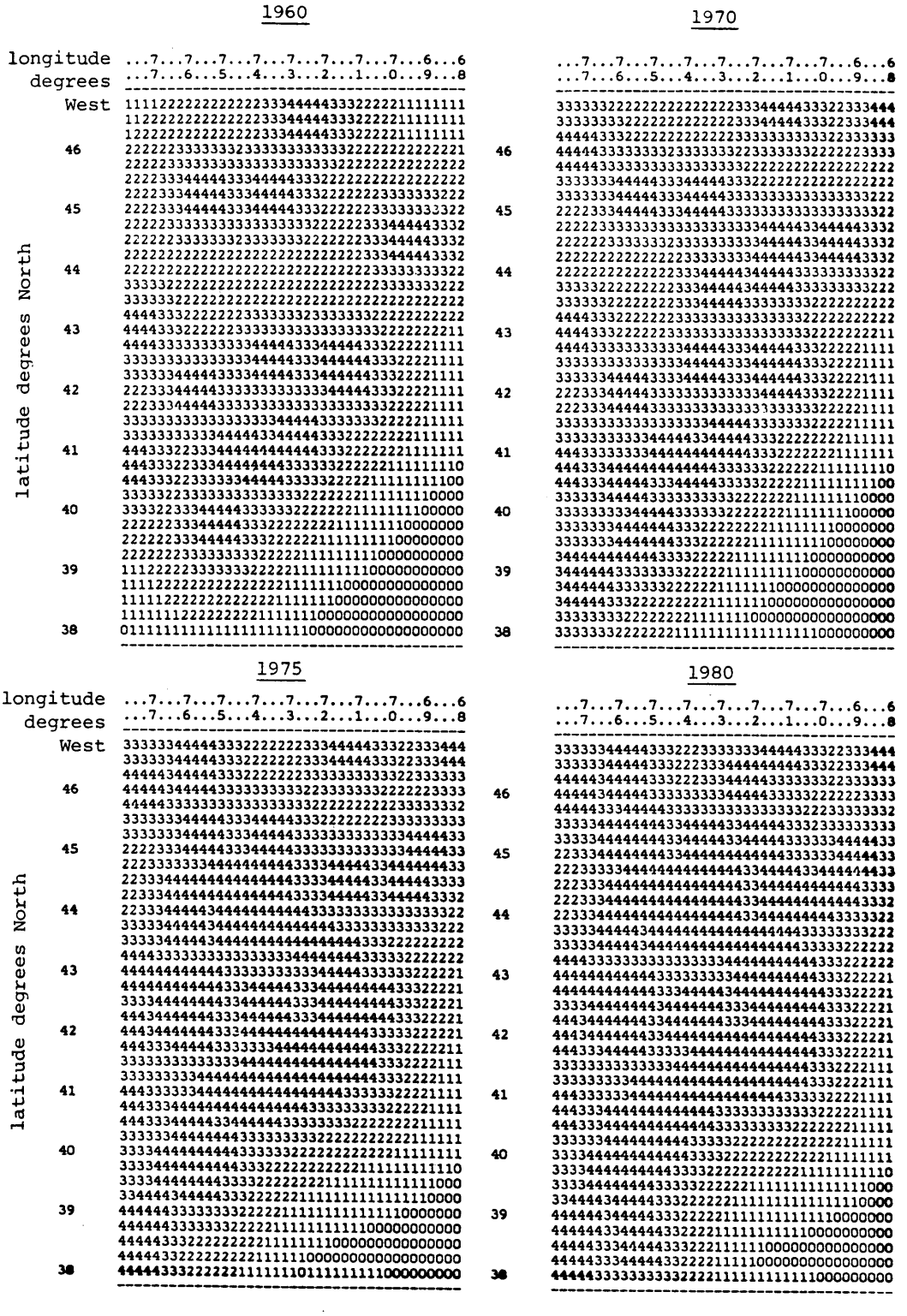
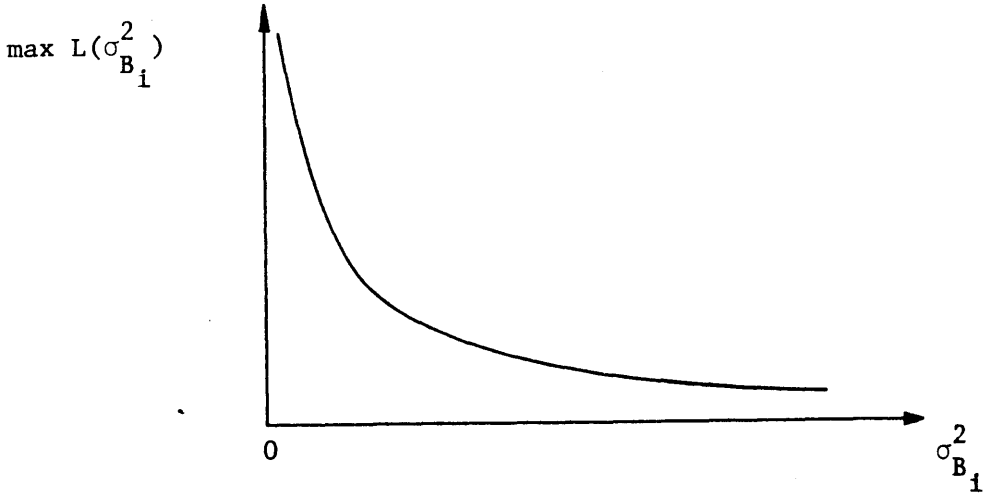
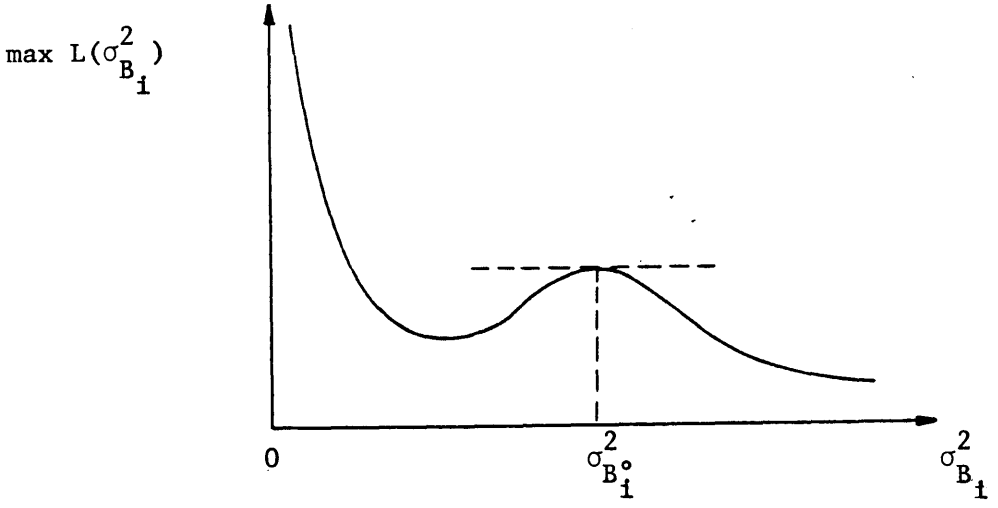


Figure 4.7 - (End)



a) No point of stationarity



b) With a stationary point

Figure 4.8 - Illustration of the loglikelihood function when the slope parameters $b_{\underline{x}}$ are iid $N(m_{B_i}, \sigma^2_{B_i})$

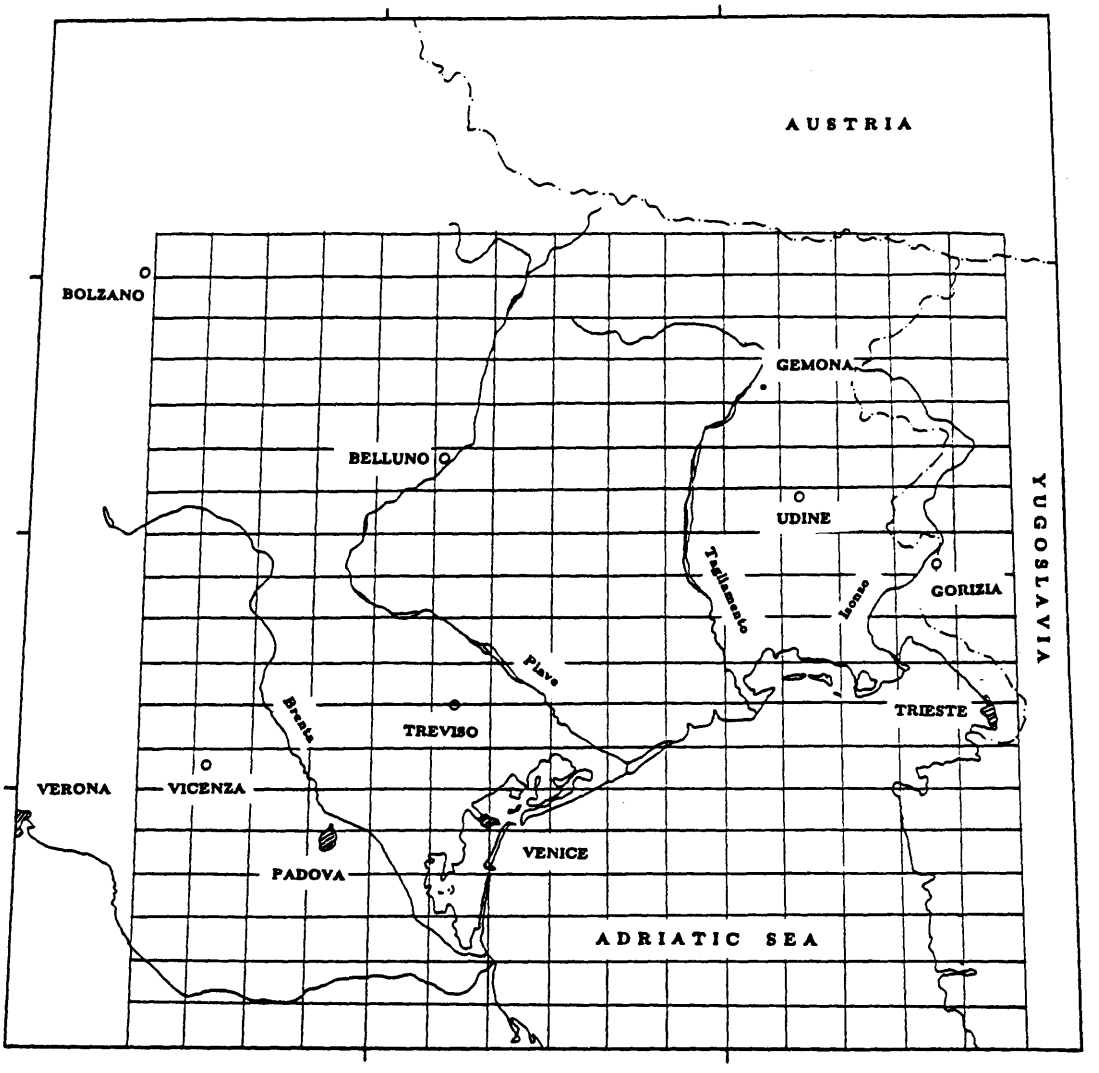


Figure 4.9 - Region of study for Model C

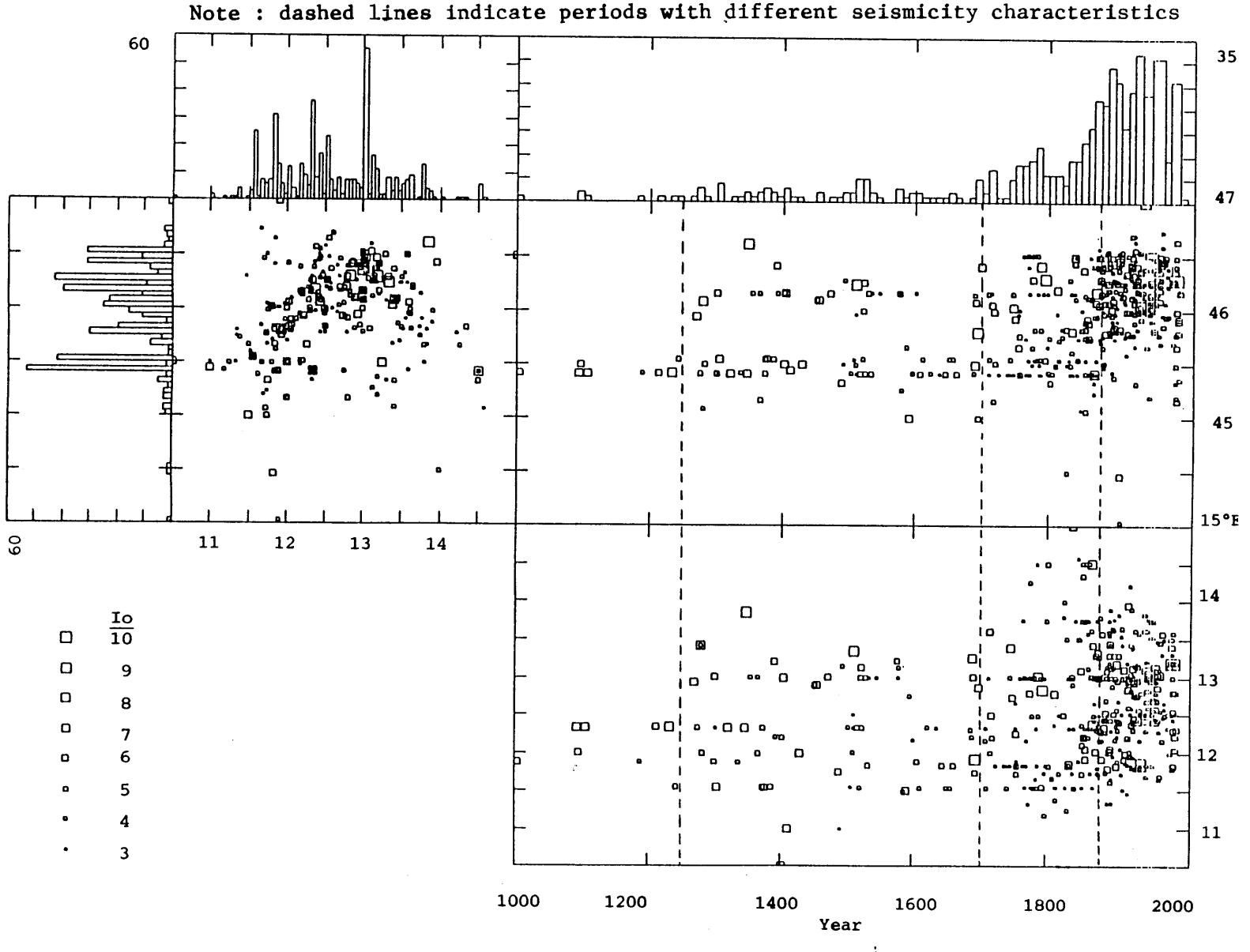


Figure 4.10a - Space-time distribution and histograms of all main events in the Friuli region

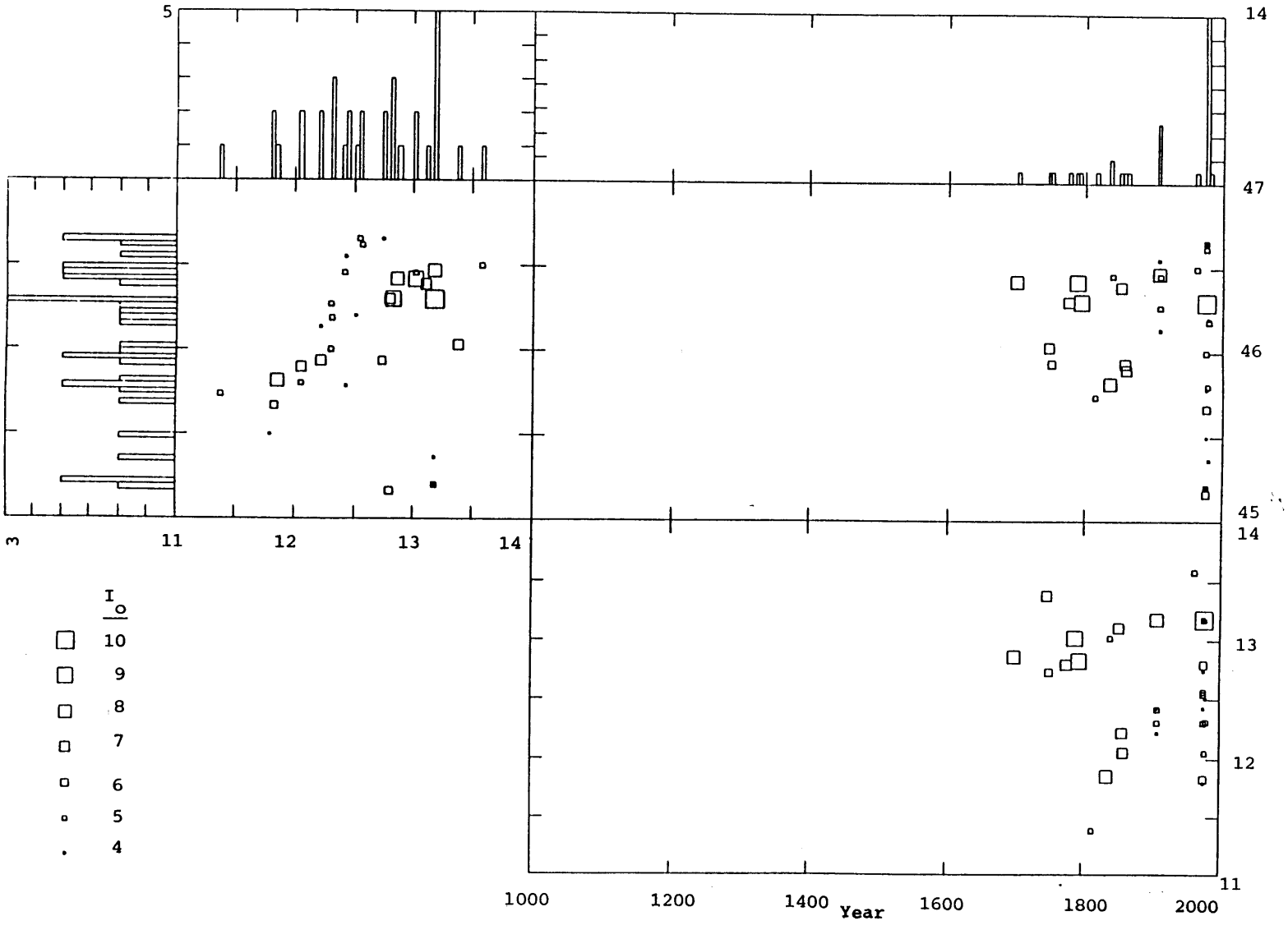


Figure 4.10b - Space-time distribution of earthquakes with i_L=1,2 and 6 (Table 4.3)

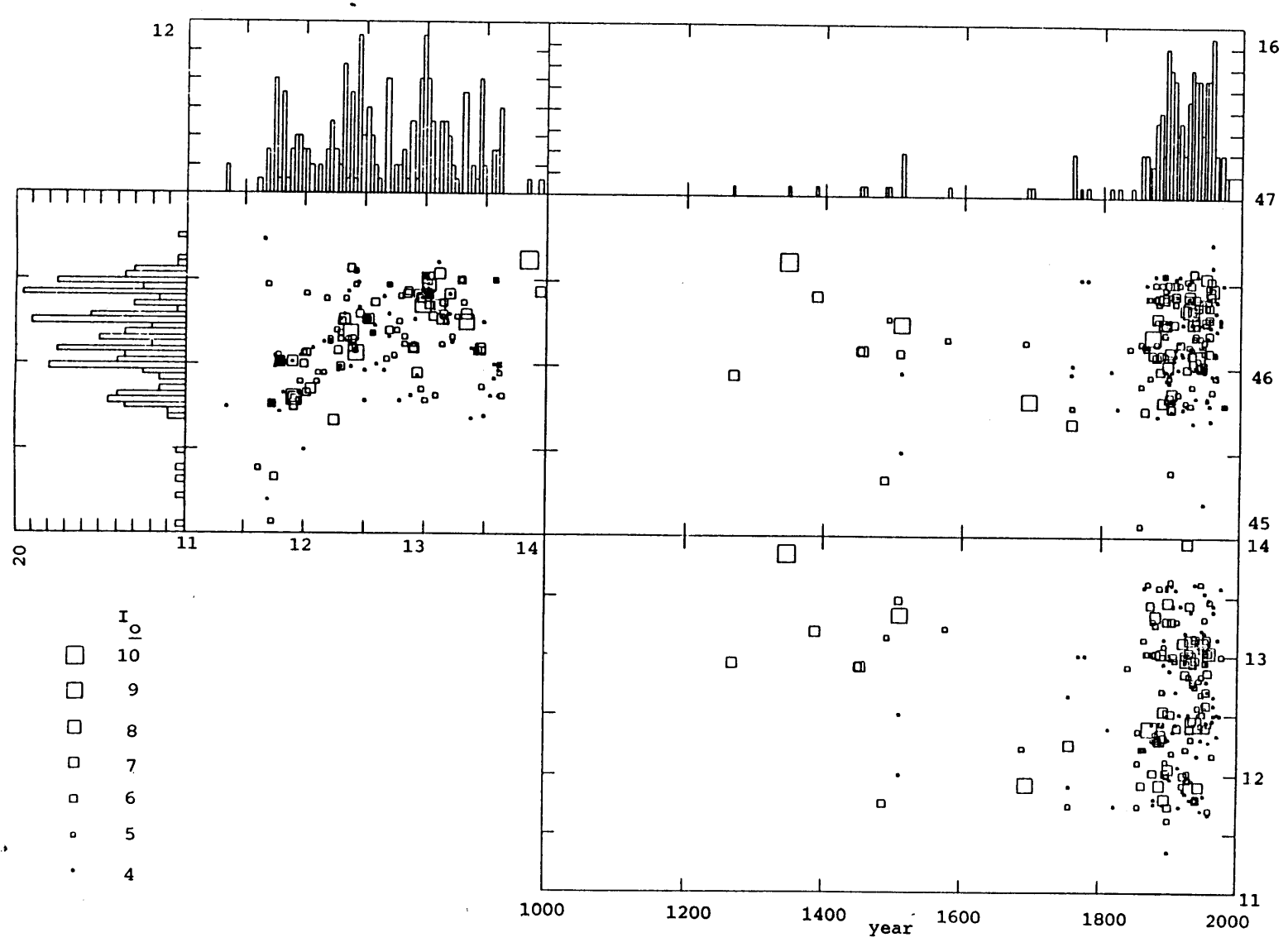


Figure 4.10c - Space-time distribution of earthquakes with i_L=3 (Table 4.3)

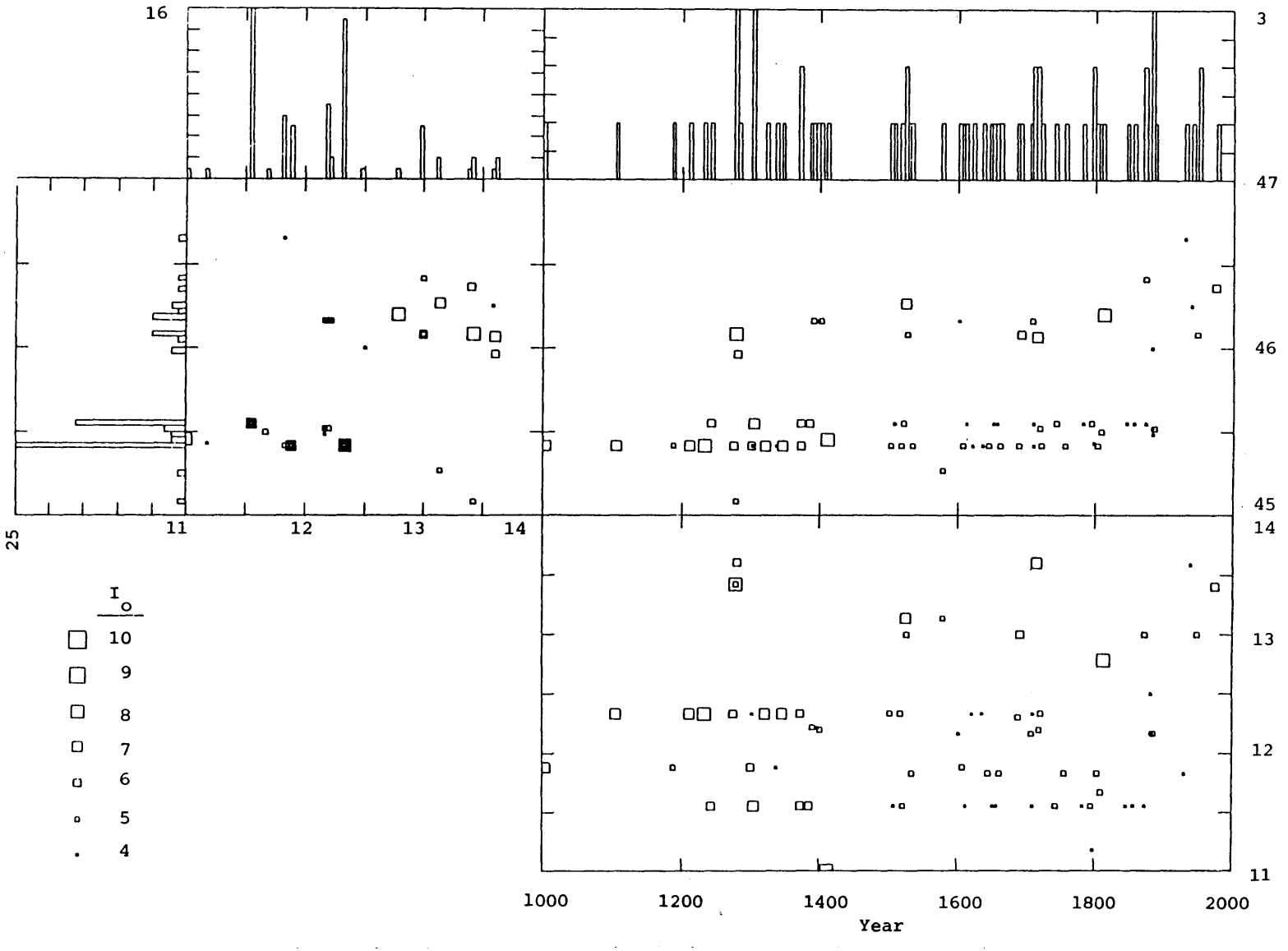


Figure 4.10d - Space-time distribution of earthquakes with $i_L = 4$ (Table 4.3)

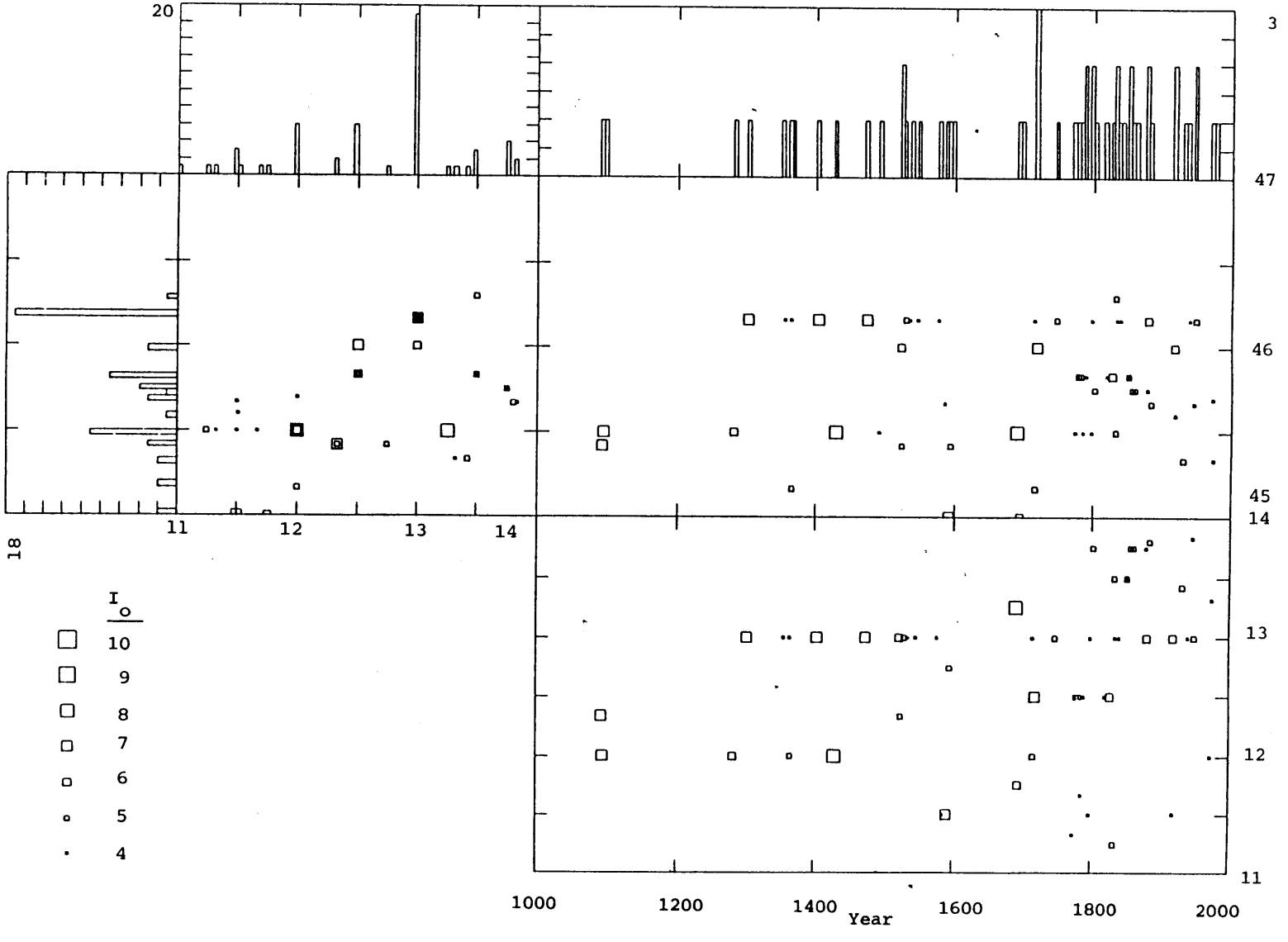
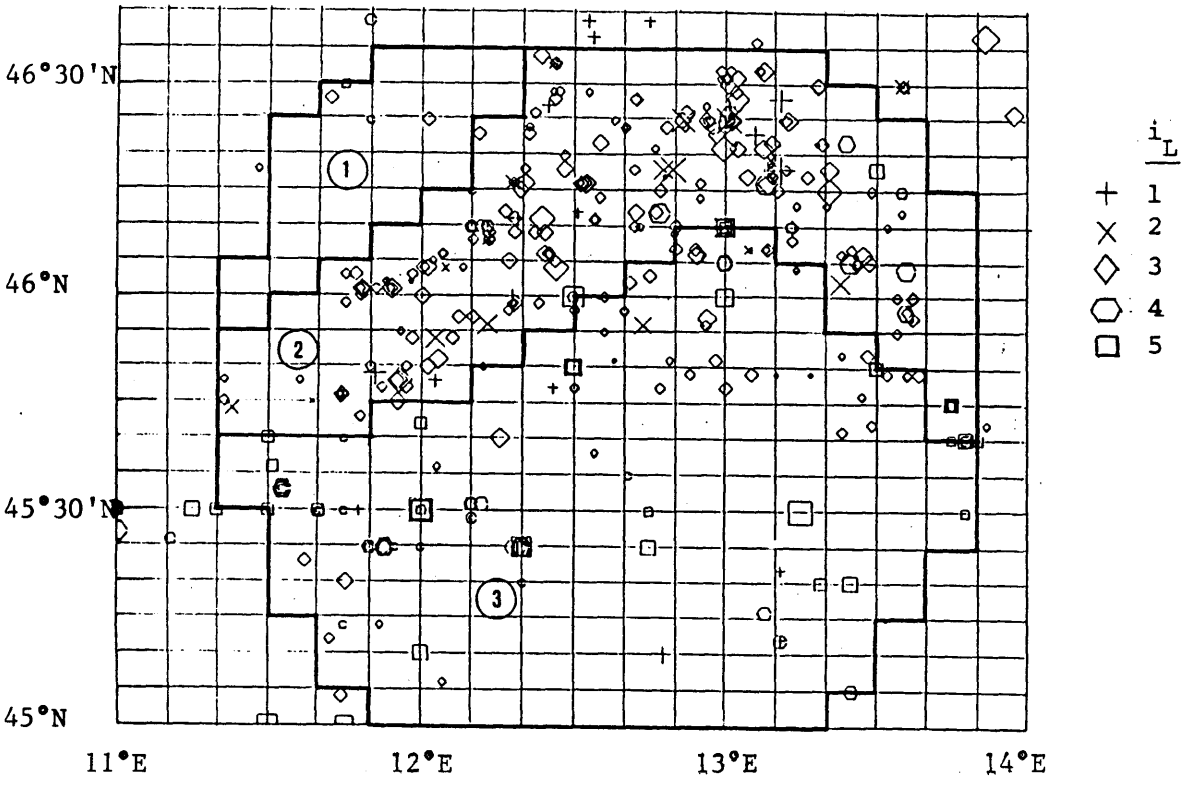


Figure 4.10e - Space-time distribution of earthquakes with $i_L = 5$ (Table 4.3)



Note : i_L=1 includes instrumentally located earthquakes (see Table 4.3)

Figure 4.10f - Spatial distribution of earthquakes classified according to i_L

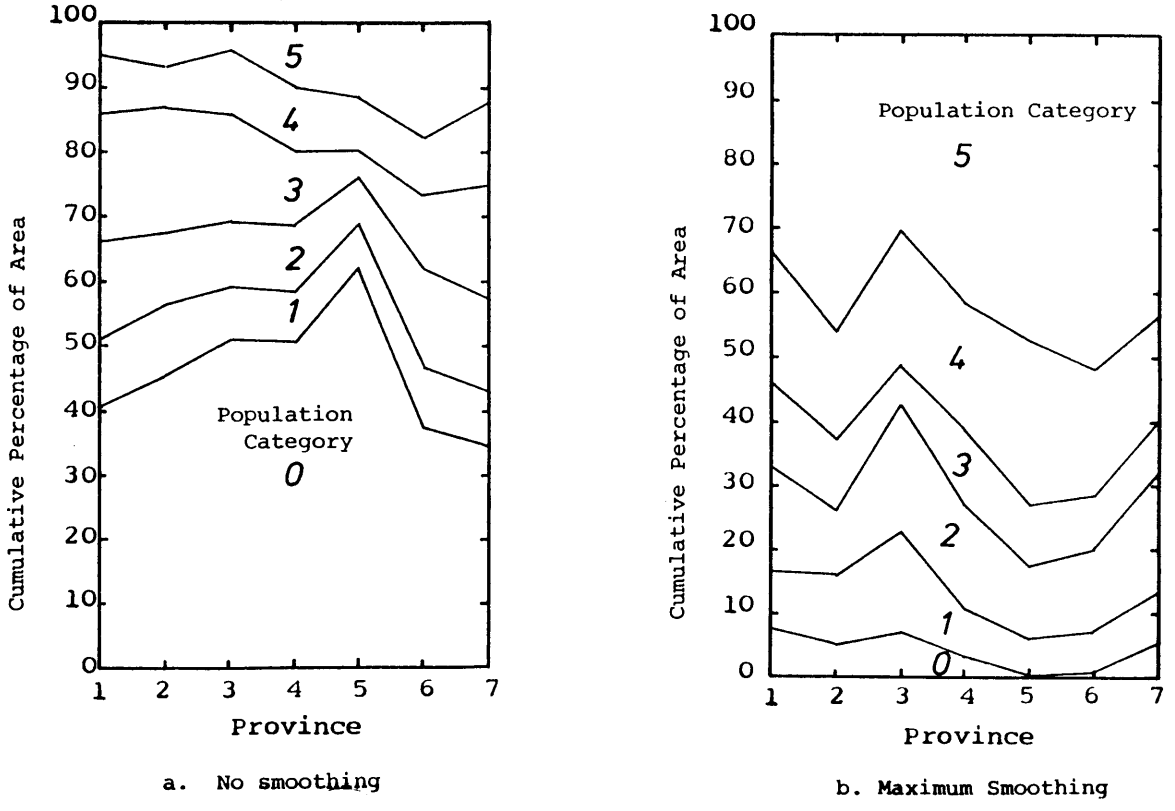
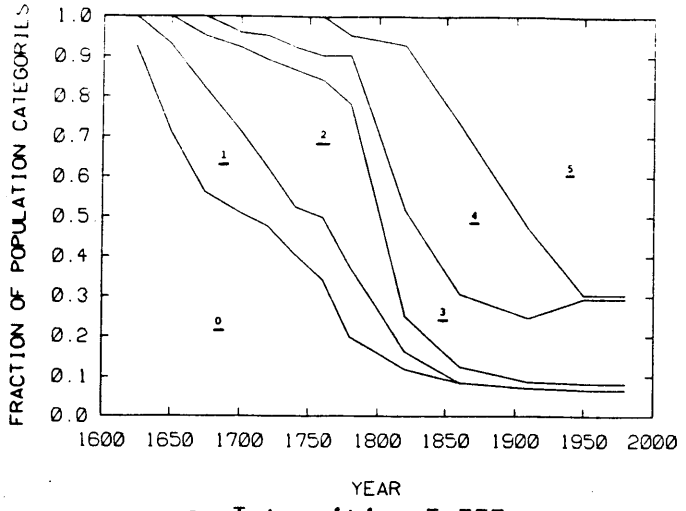
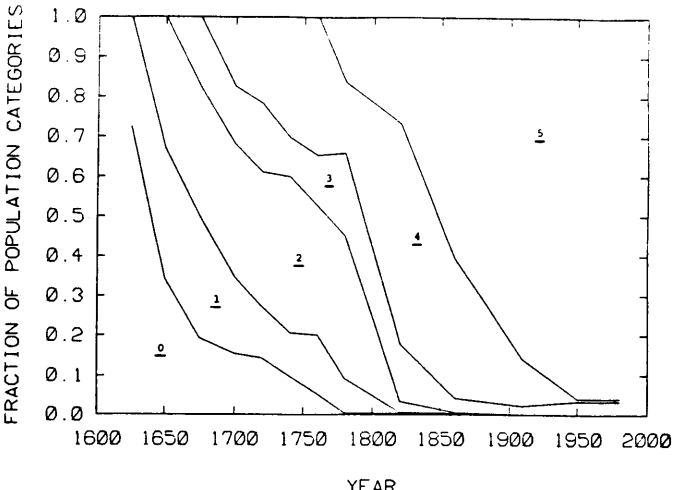


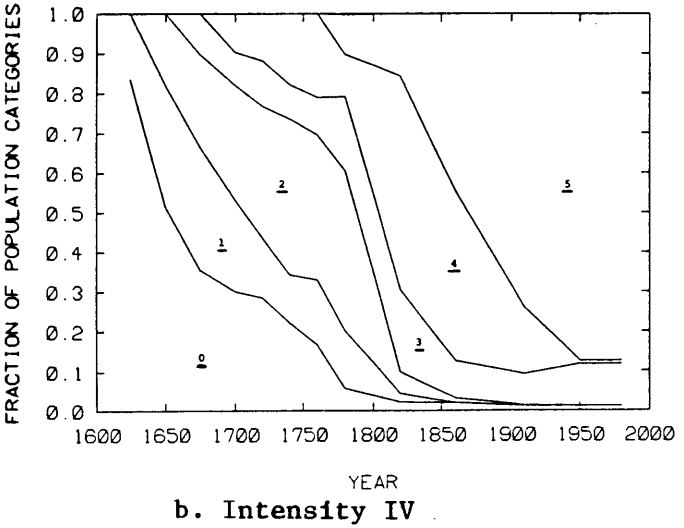
Figure 4.11 - Cumulative percentage of province area, averaged in time, associated with different population densities



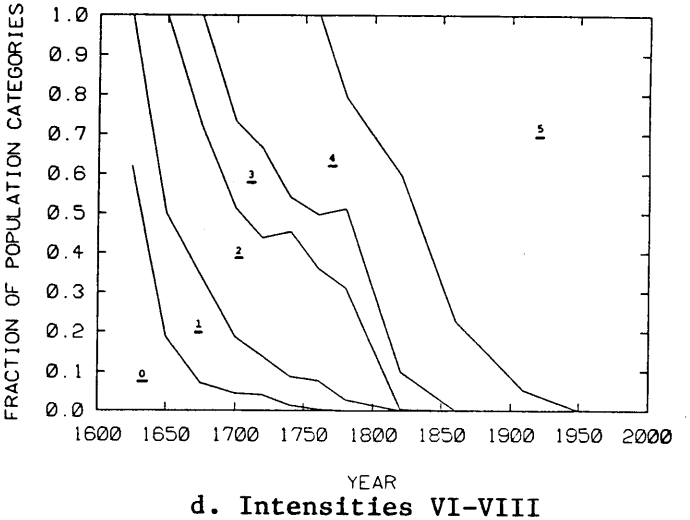
a. Intensities I-III



c. Intensity V



b. Intensity IV



d. Intensities VI-VIII

Figure 4.12 - Cumulative fraction of total area associated with each population category as a function of time for smoothed population (Case B)

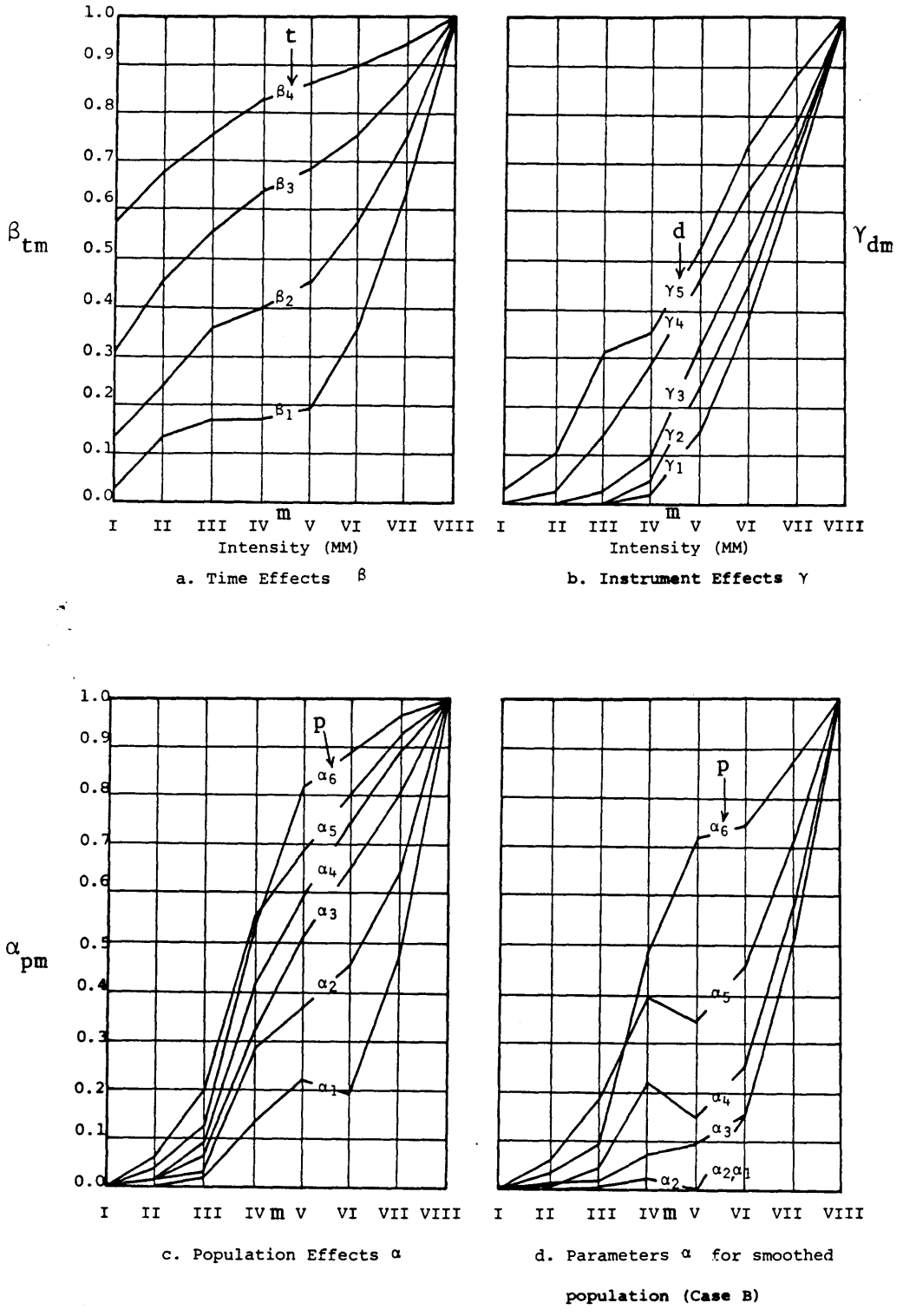


Figure 4.13 - Incompleteness parameters (Model A, Case 1 with independent b_k parameters)

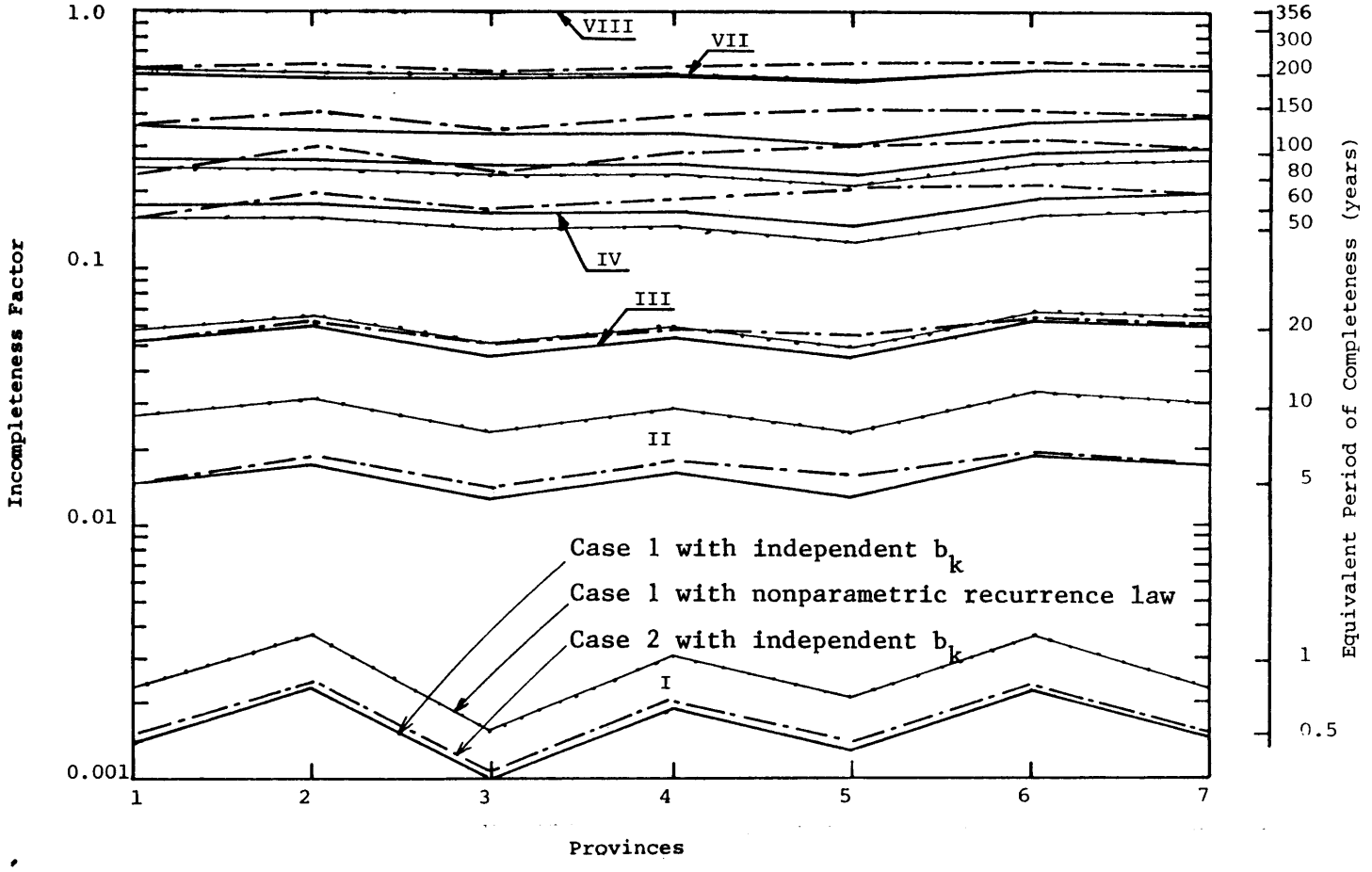
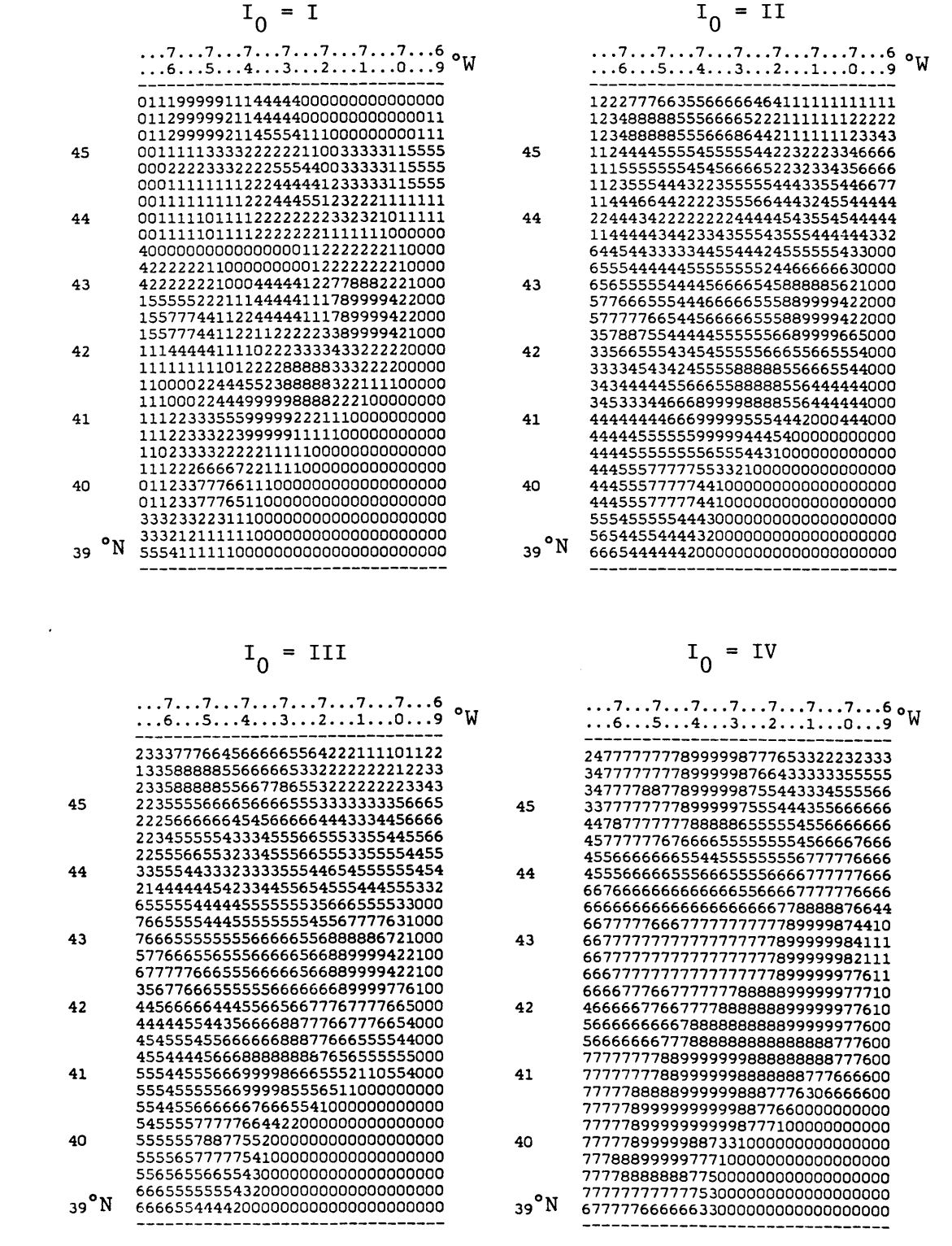


Figure 4.14 - Incompleteness factor and equivalent periods of completeness for each province and earthquake size



Note : to obtain actual values of the equivalent periods of completeness use Equation 4.126

Figure 4.15 - Relative values of the equivalent period of completeness for different intensities (Case 2 with independent b_k parameters)

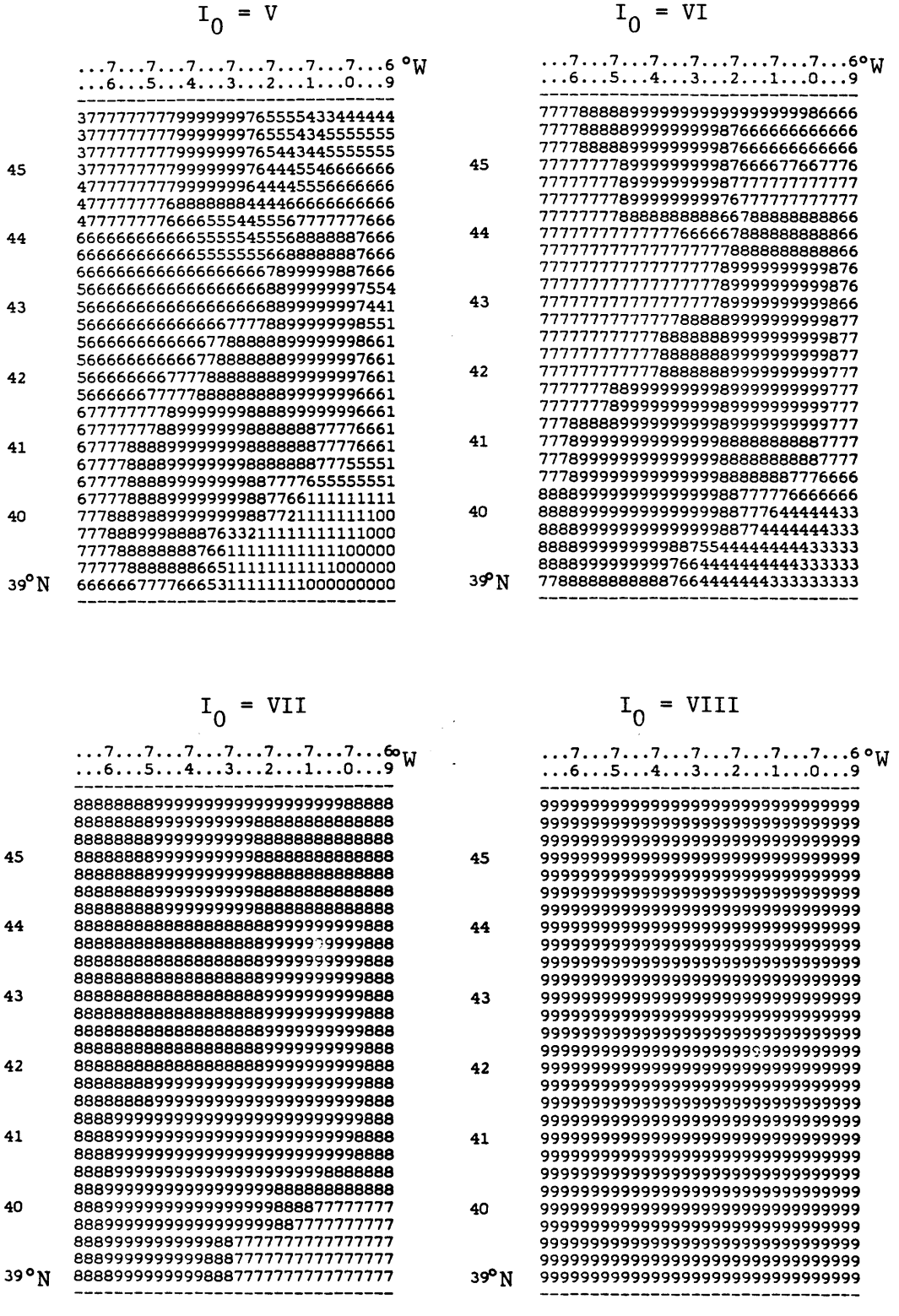


Figure 4.15 - (End)

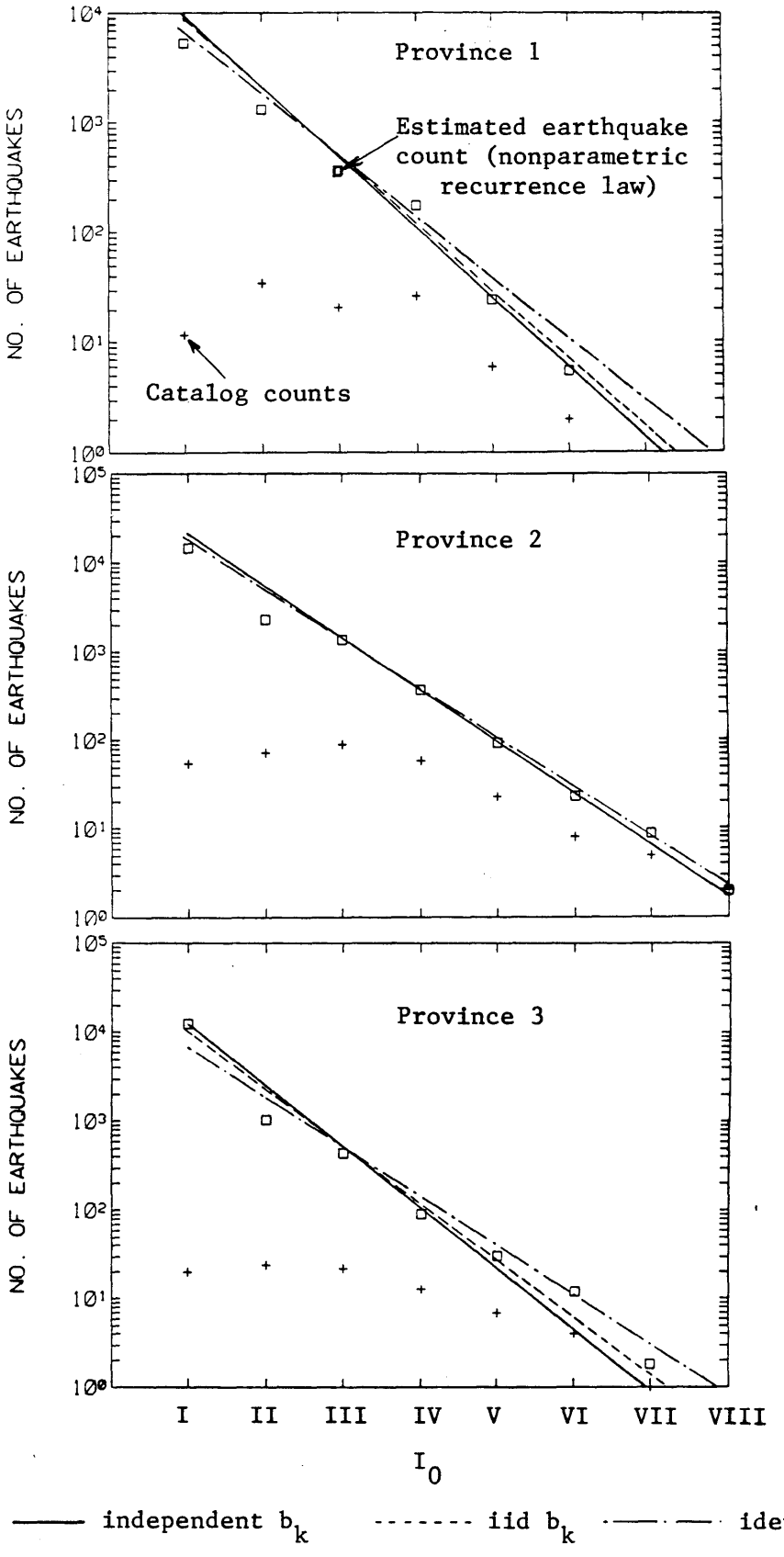


Figure 4.16 - Fitted exponential recurrence rates for different assumptions on the b_k parameters (Case 1)

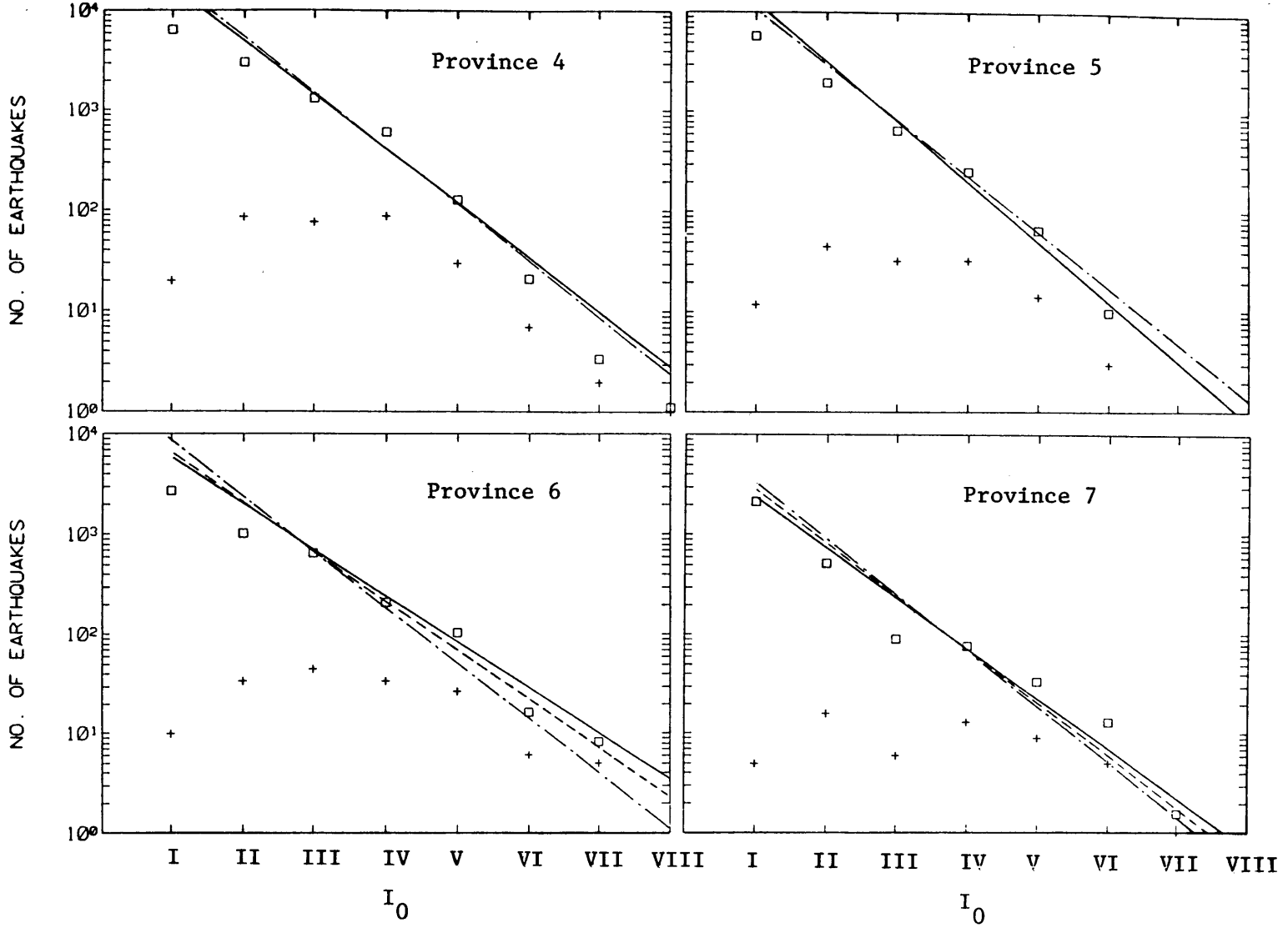


Figure 4.16 - (End)

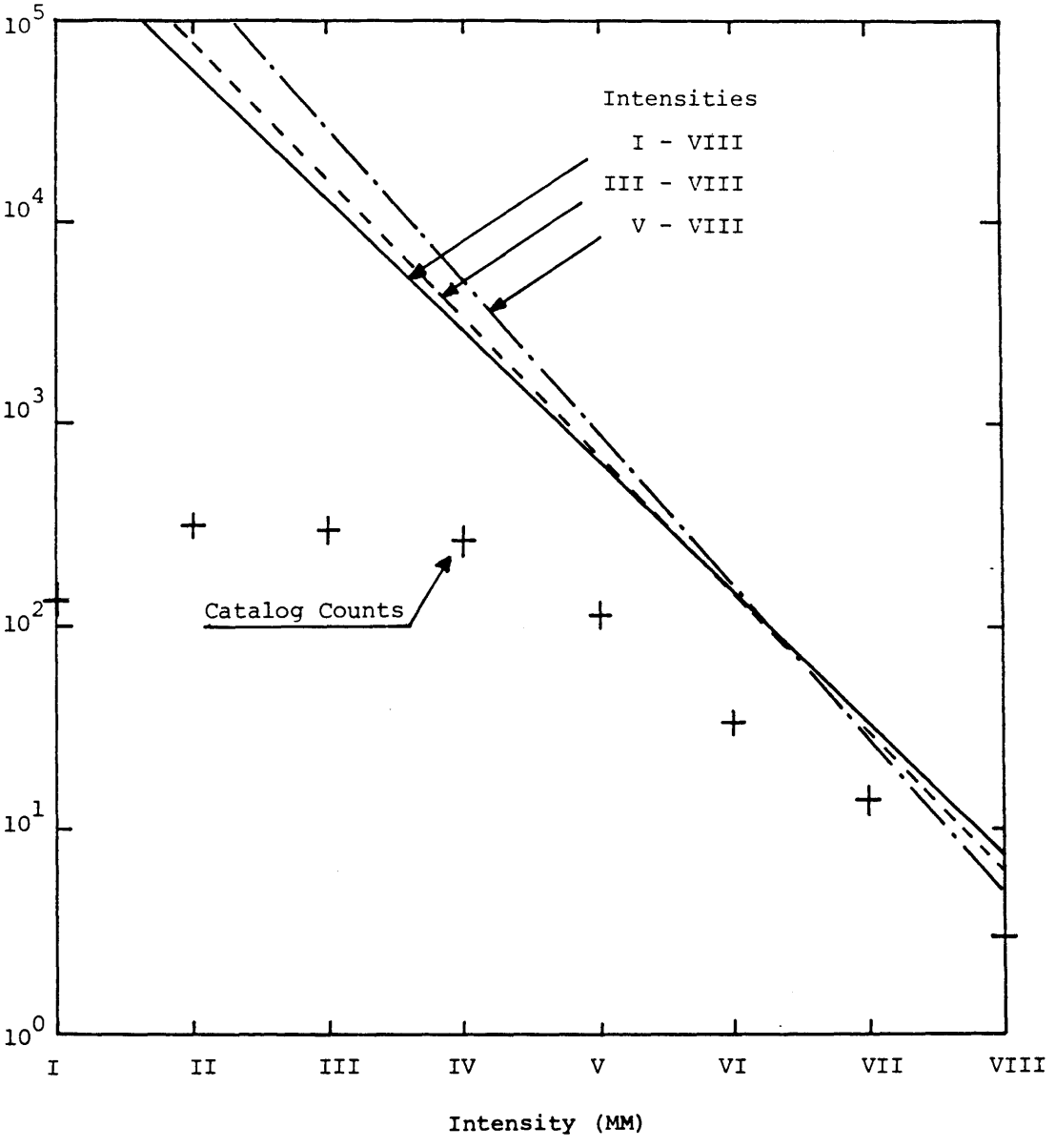
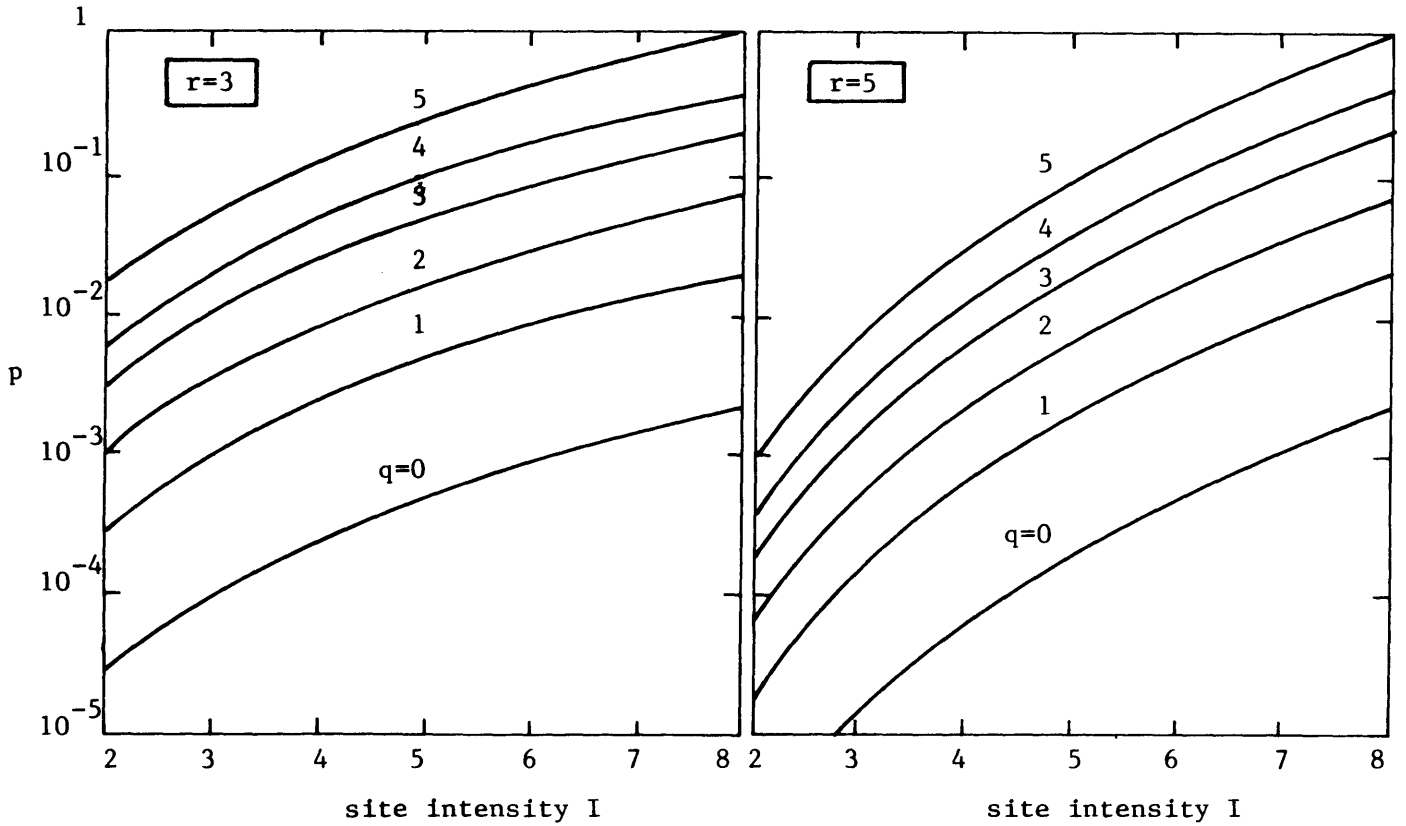


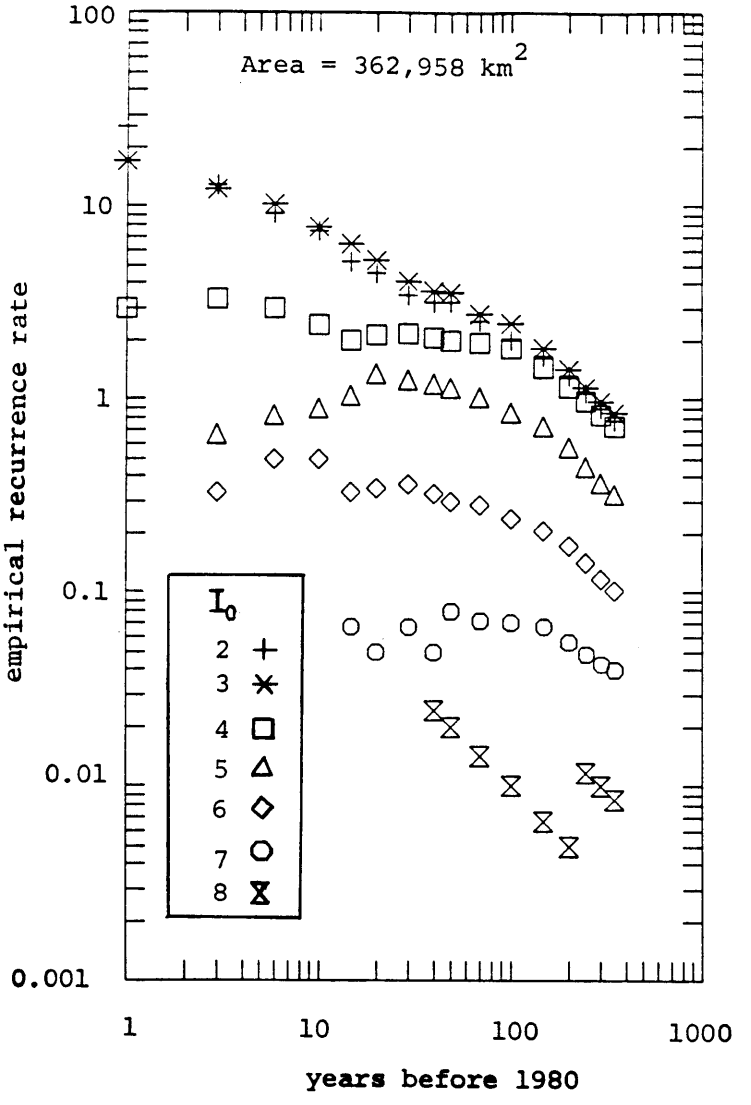
Figure 4.17 - Fitted exponential recurrence rates, summed over all provinces, using different lower bounds of I_0 (Case 2, identical b_k)



Note - $p \propto q\ I^r$, where q is a nominal population density (see Table 4.10)

- p is normalized to 1 for q=5 and I=8 in both figures
- in application, p corresponds to $\int_{\underline{x}} q(\underline{x}) I^r(\underline{x}) d\underline{x}$ and is discretized as in Table 4.11

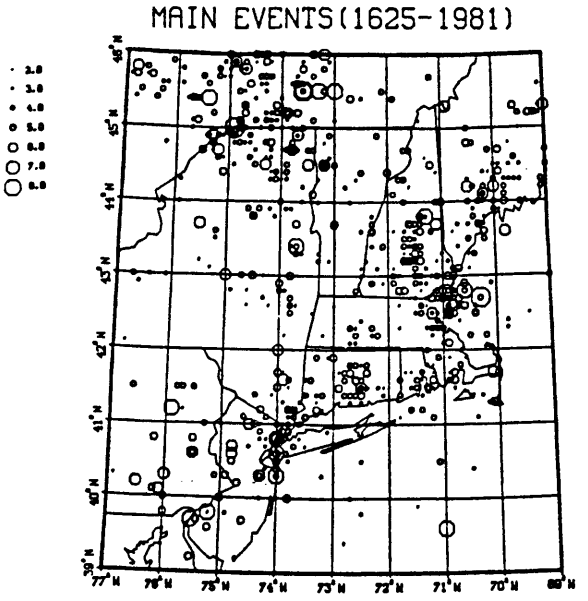
Figure 4.18 - Equivalent population p for r=3 and r=5



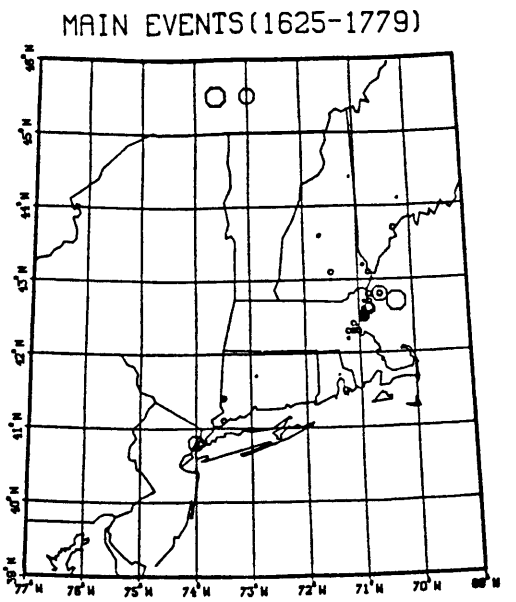
Note : based on Chiburis catalog with dependent events, identified in the Base Case of Chapter 3, removed

Figure 4.19 - Empirical recurrence rate versus observation time for the region in Fig. 4.1a

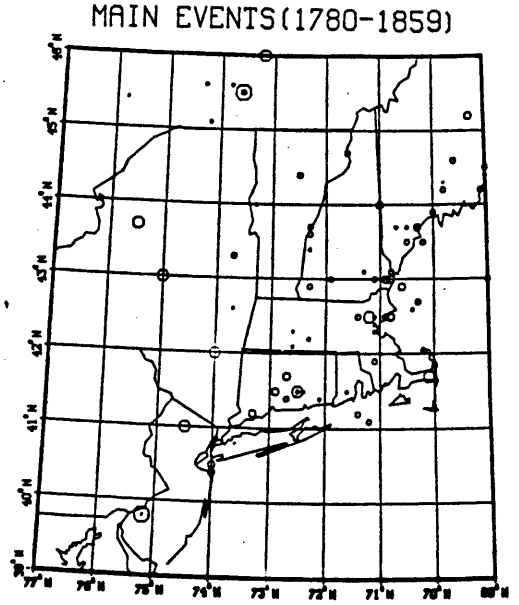
(a)



(b)



(c)



(d)

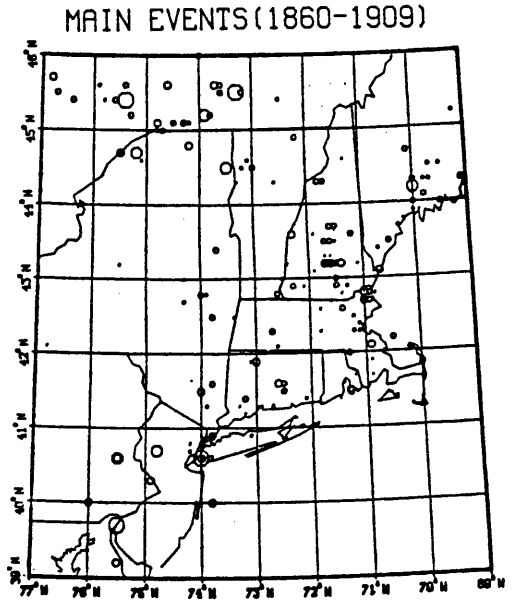


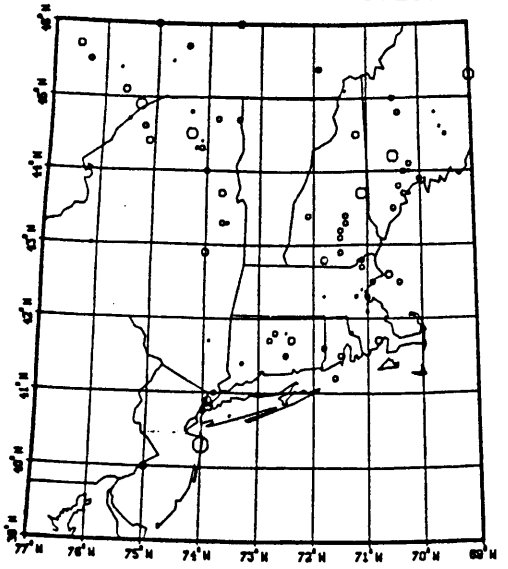
Figure 4.20 - Historical occurrence of main events,

(a) from 1625 to 1981

(b-h) over selected time intervals

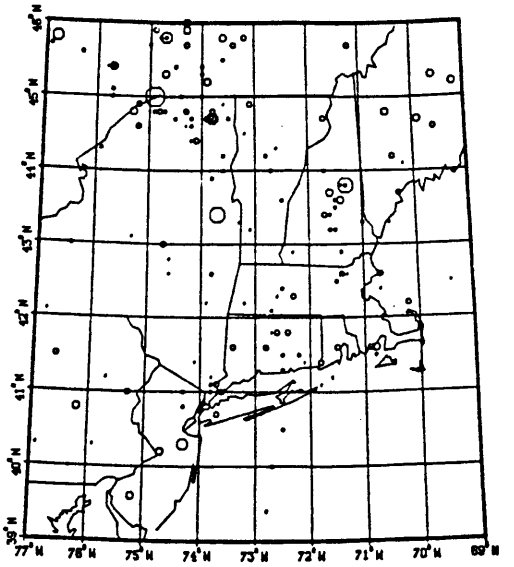
(e)

MAIN EVENTS(1910-1929)



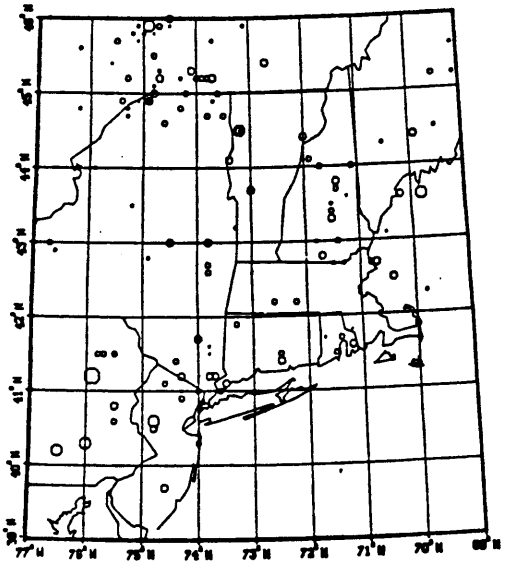
(f)

MAIN EVENTS(1930-1949)



(g)

MAIN EVENTS(1950-1969)



(h)

MAIN EVENTS(1970-1981)

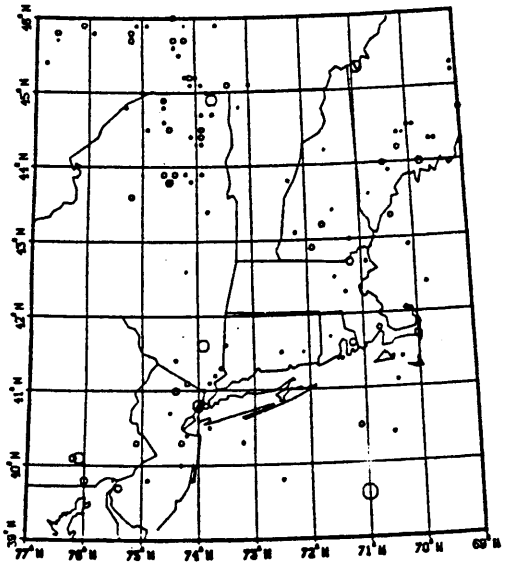
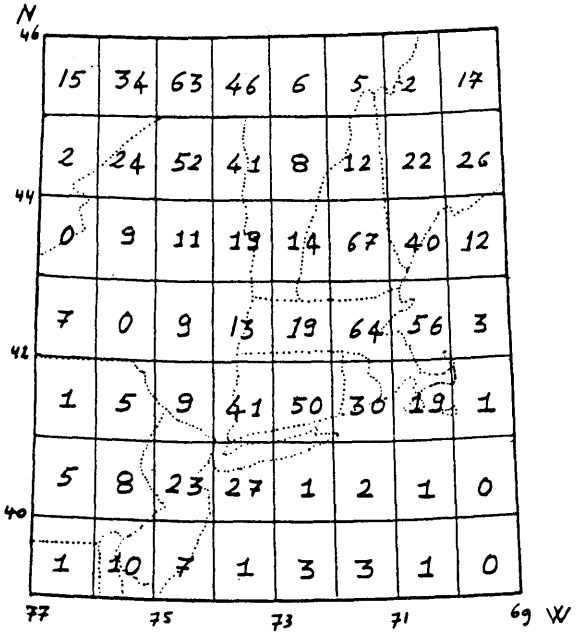


Figure 4.20 - (End)

(a)



(b)

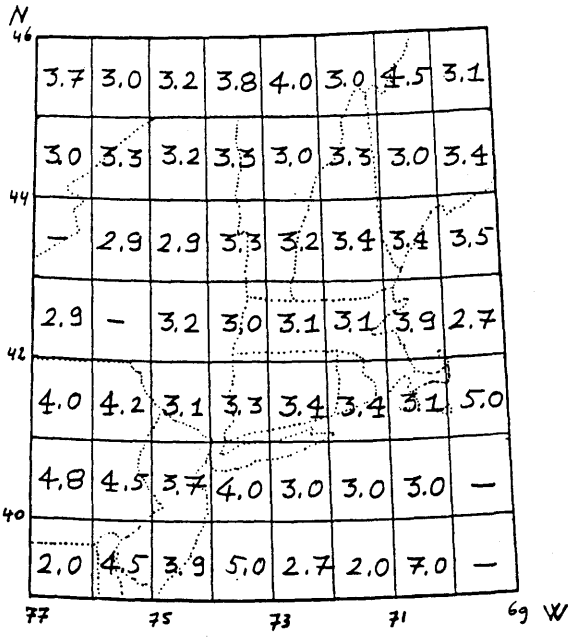


Figure 4.21 - Catalog counts at locations $\underline{x}$

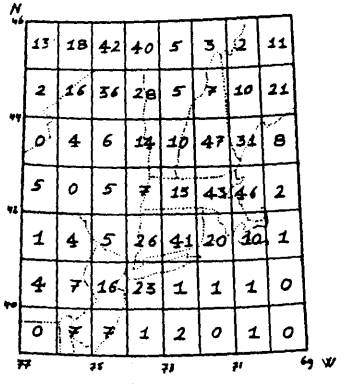
(a) Total counts of earthquakes with $I_0 > 1$

(b) Average epicentral intensity

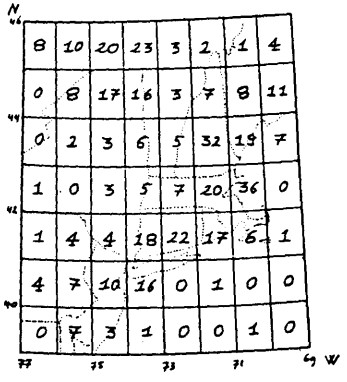
(c) Cumulative counts for $I_0 > 2, \ldots, 7$

(c)

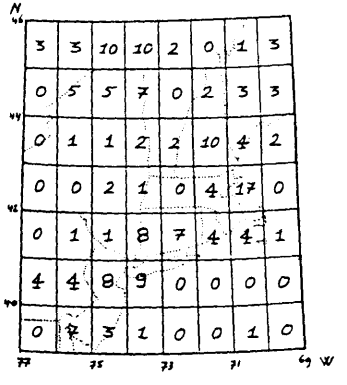
$I_o \geq 3$



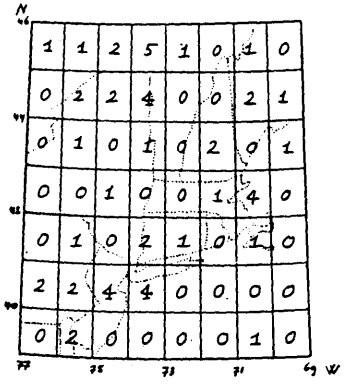
$I_o \geq 4$



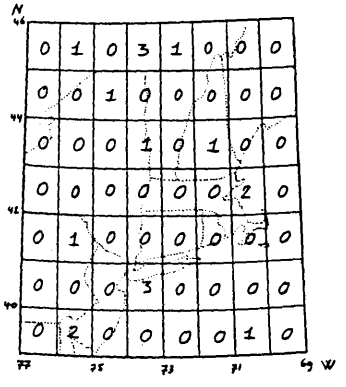
$I_o \geq 5$



$I_o \geq 6$



$I_o \geq 7$



$I_o \geq 8$

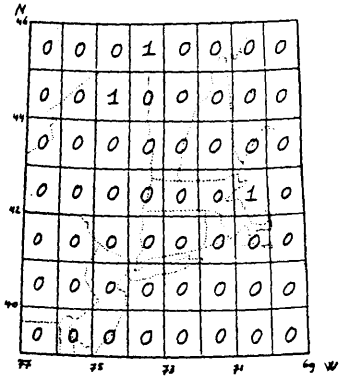


Figure 4.21 - (End)

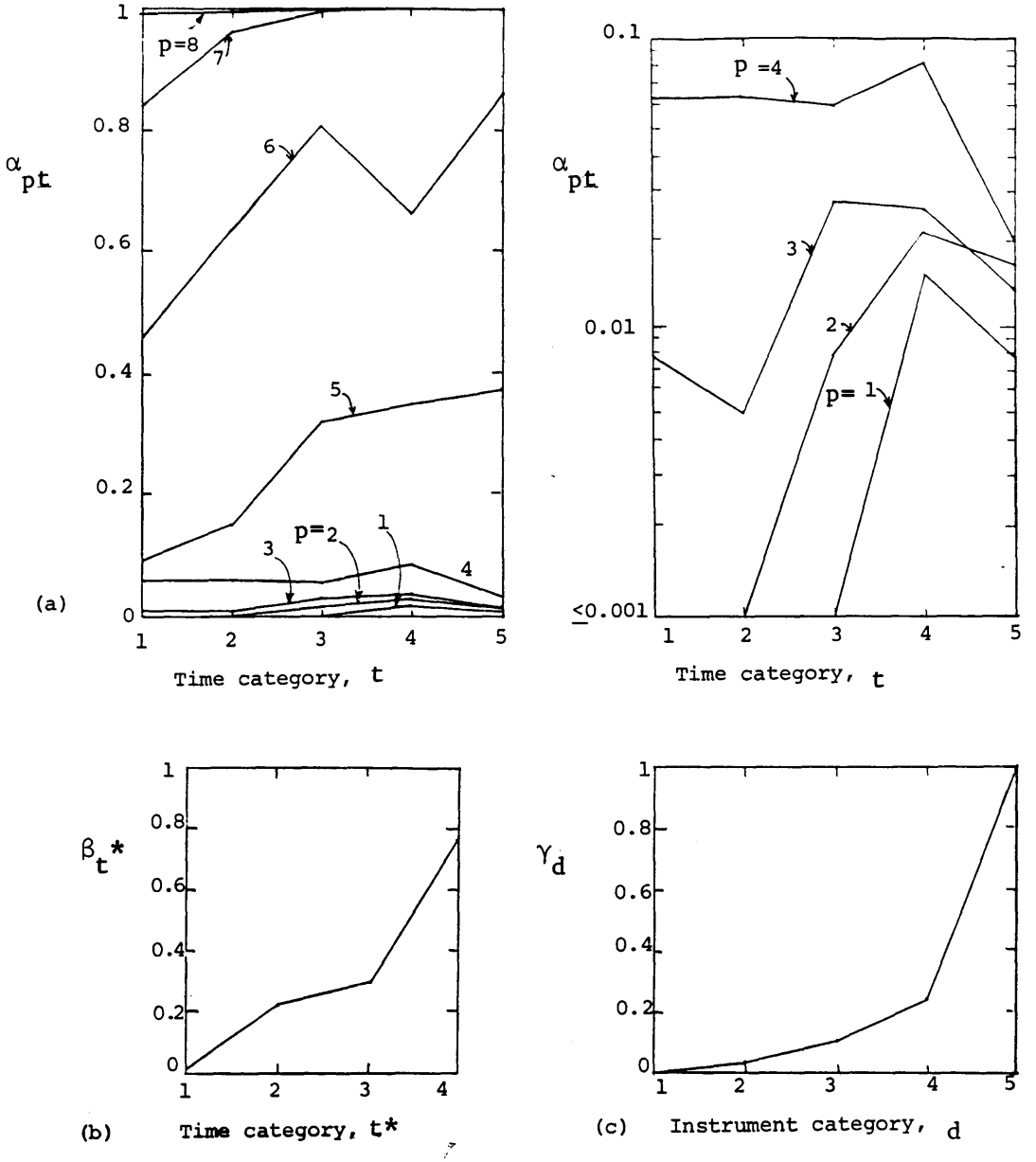


Figure 4.22 - Incompleteness parameters estimated in the base case analysis (r=5)

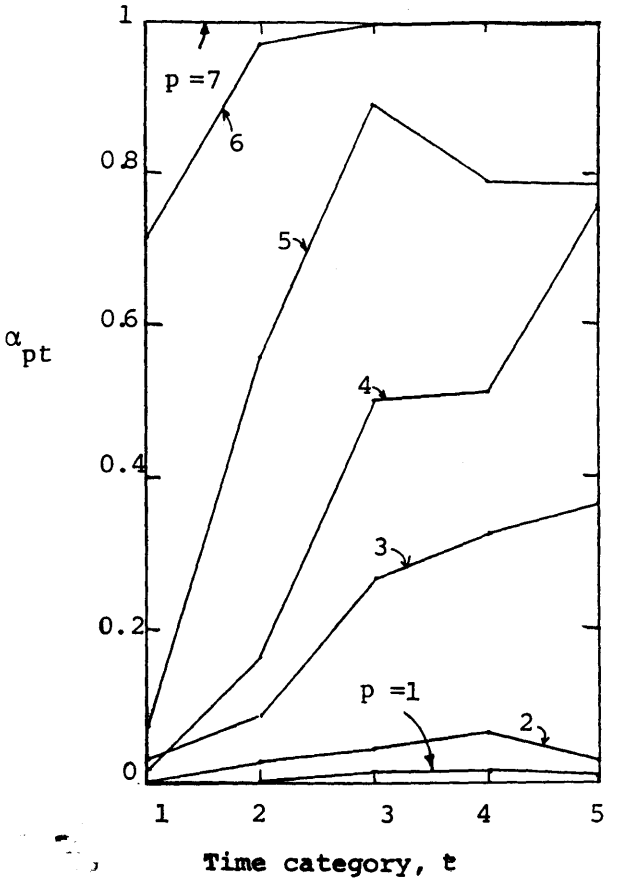


Figure 4.23 - Estimates of the population effects α_{pt} for r=∞ (Case 2)

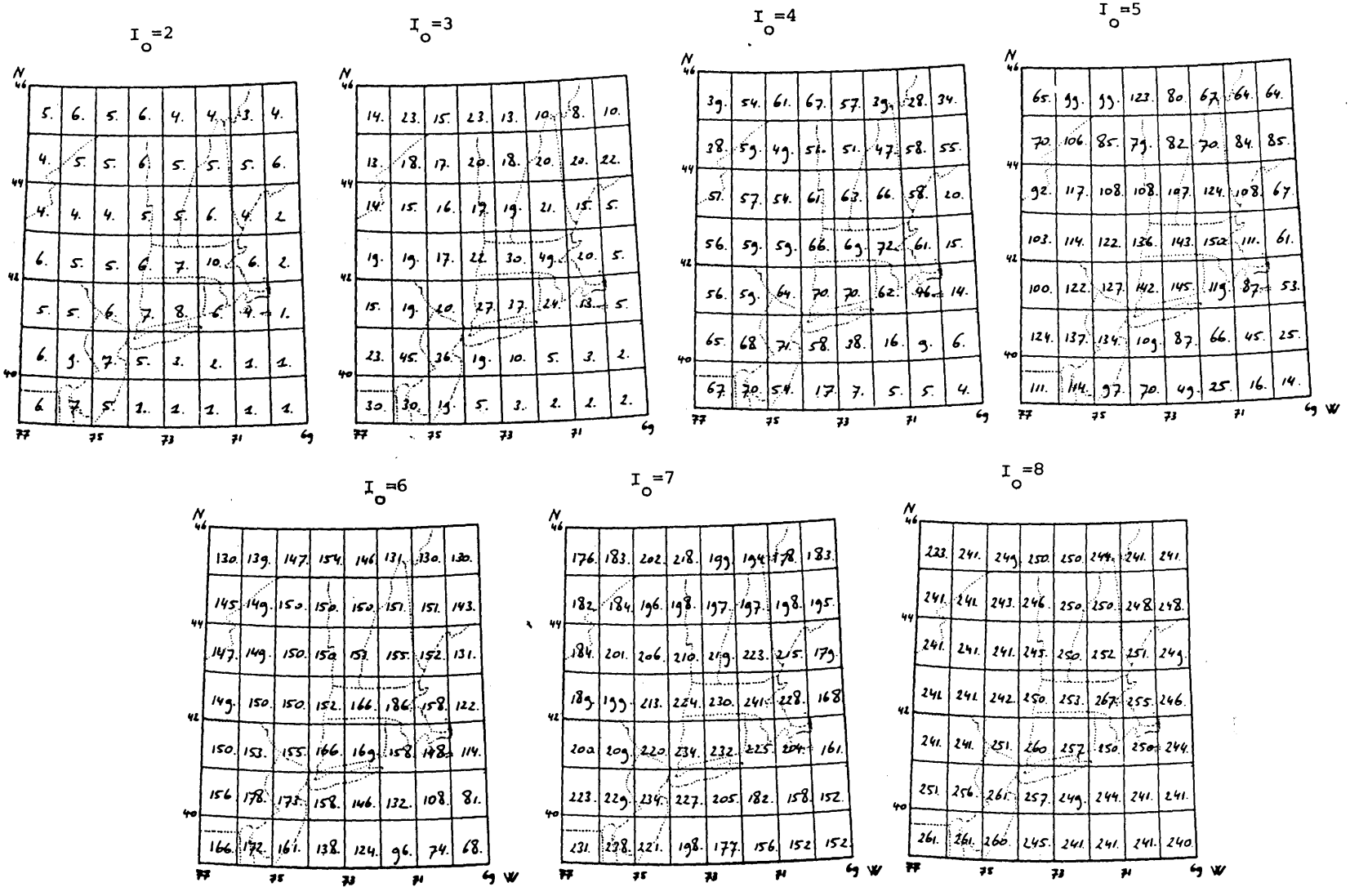


Figure 4.24 - Equivalent period of completeness (base case analysis)

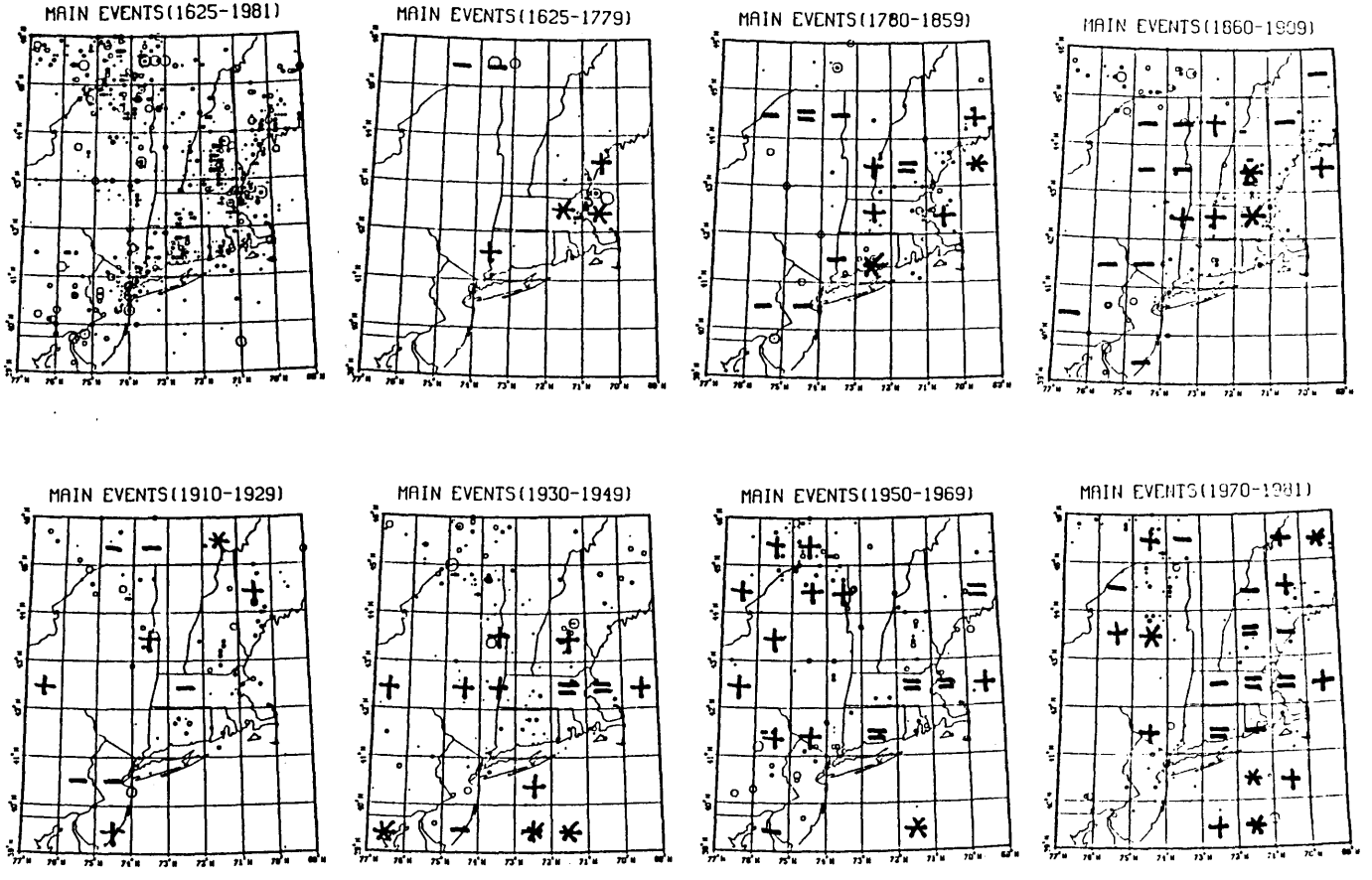


Figure 4.25a - Space-time pattern of "significant" deviations for the base case analysis (δ=0)

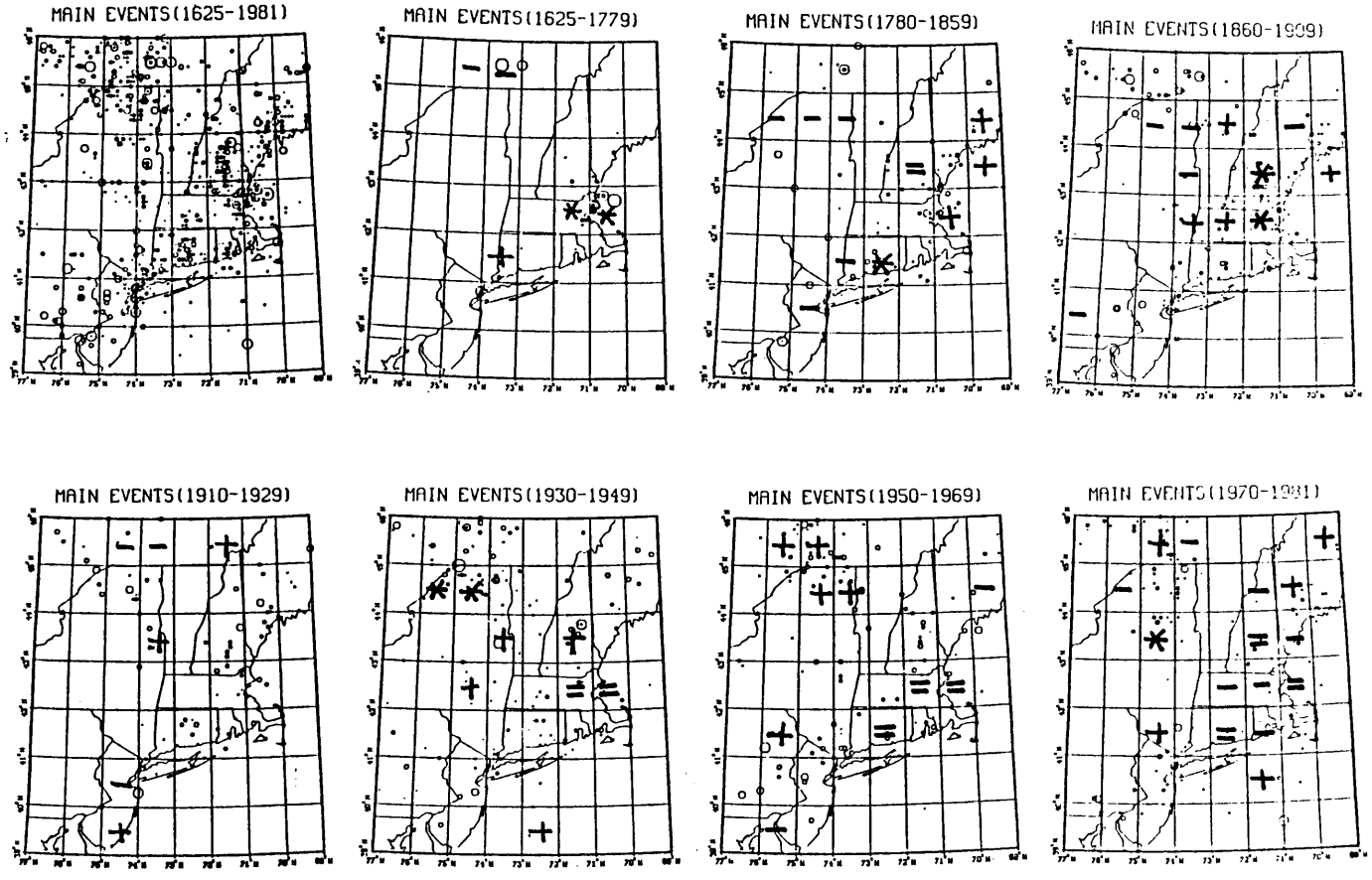


Figure 4.25b - Space-time pattern of "significant" deviations for the base case analysis (δ=1)

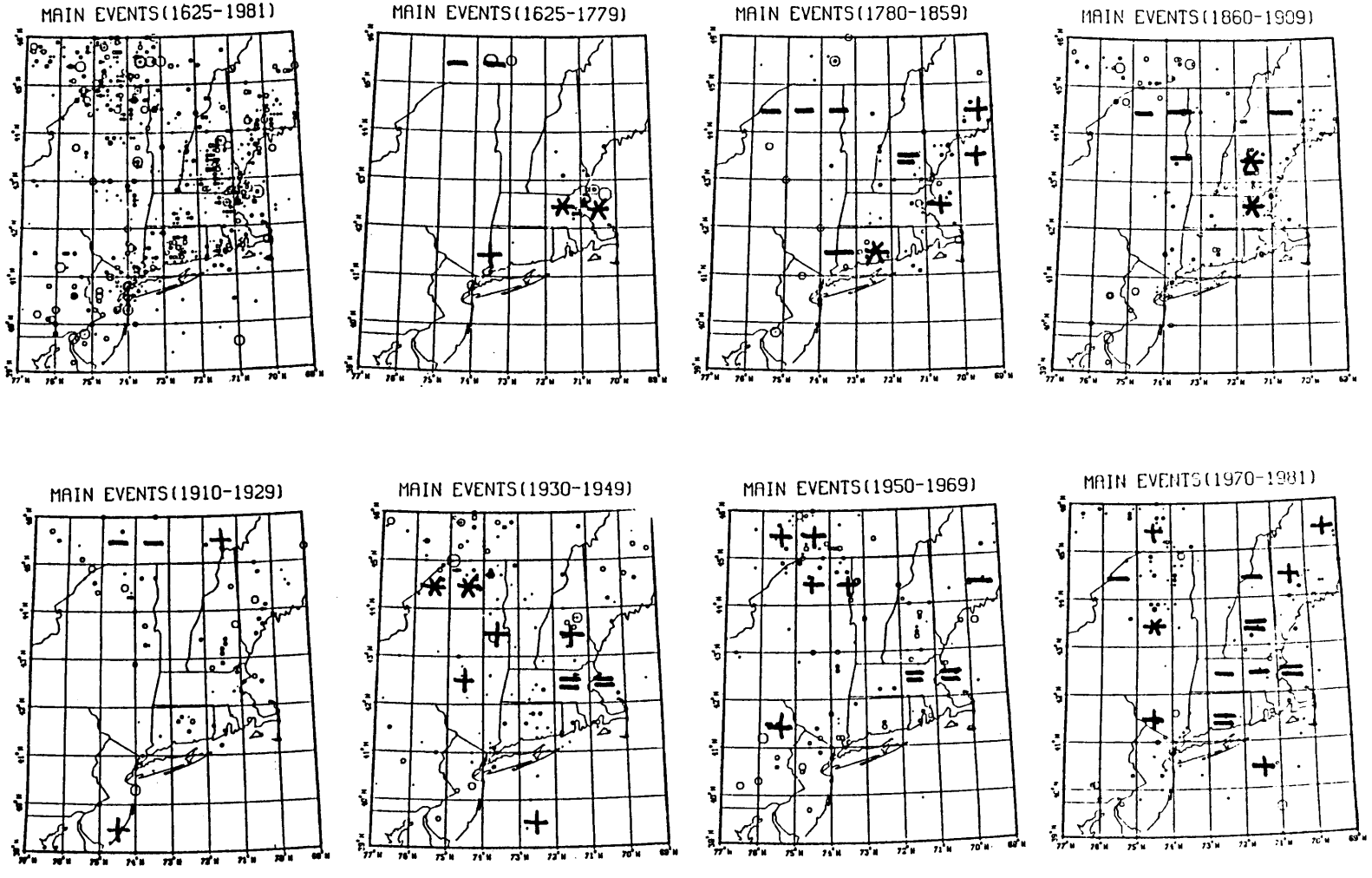


Figure 4.26a - Space-time pattern of "significant" deviations for Case 5 (δ=1)

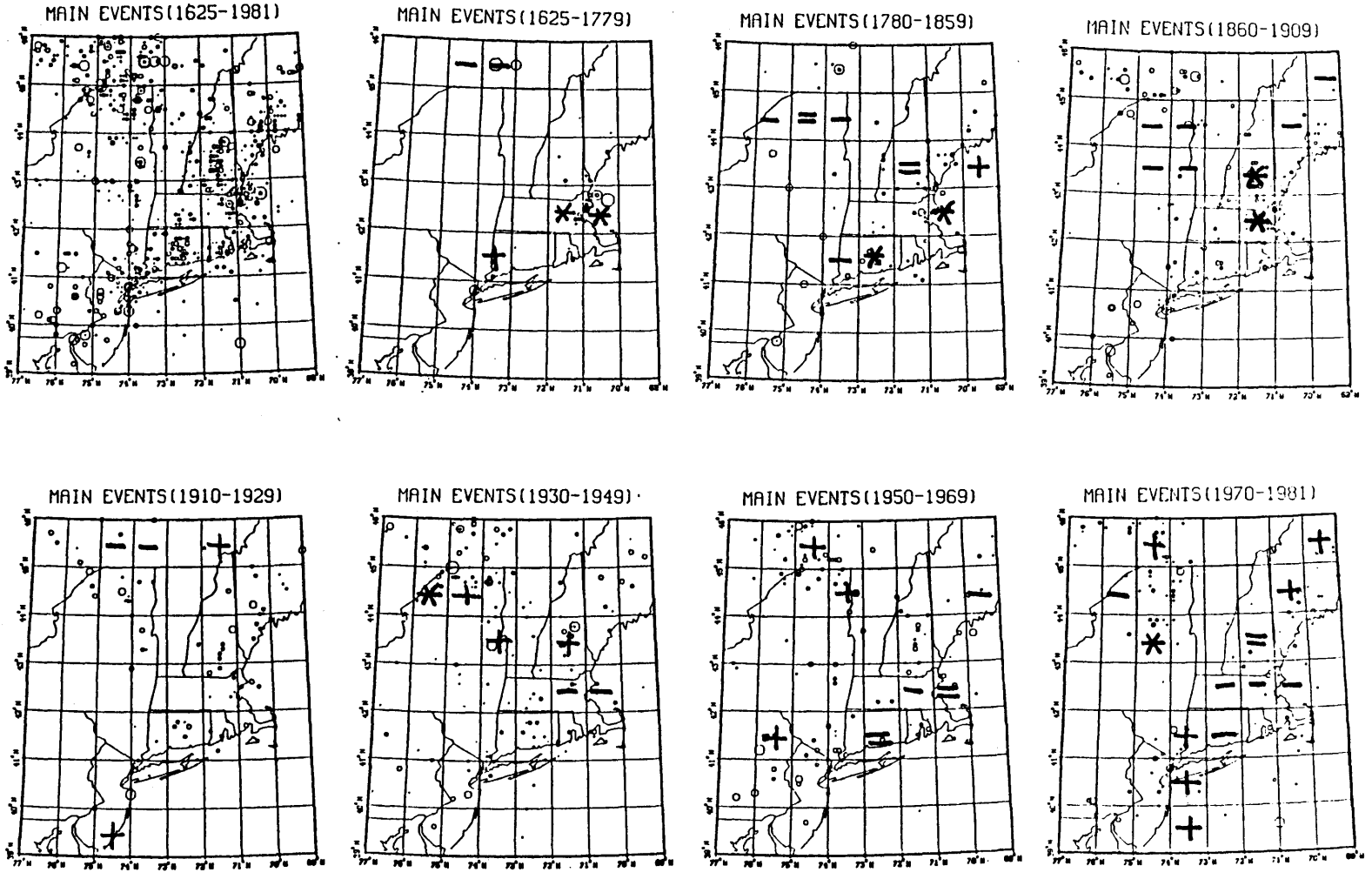


Figure 4.26b - Space-time pattern of "significant" deviations for Case 6 (δ=1)

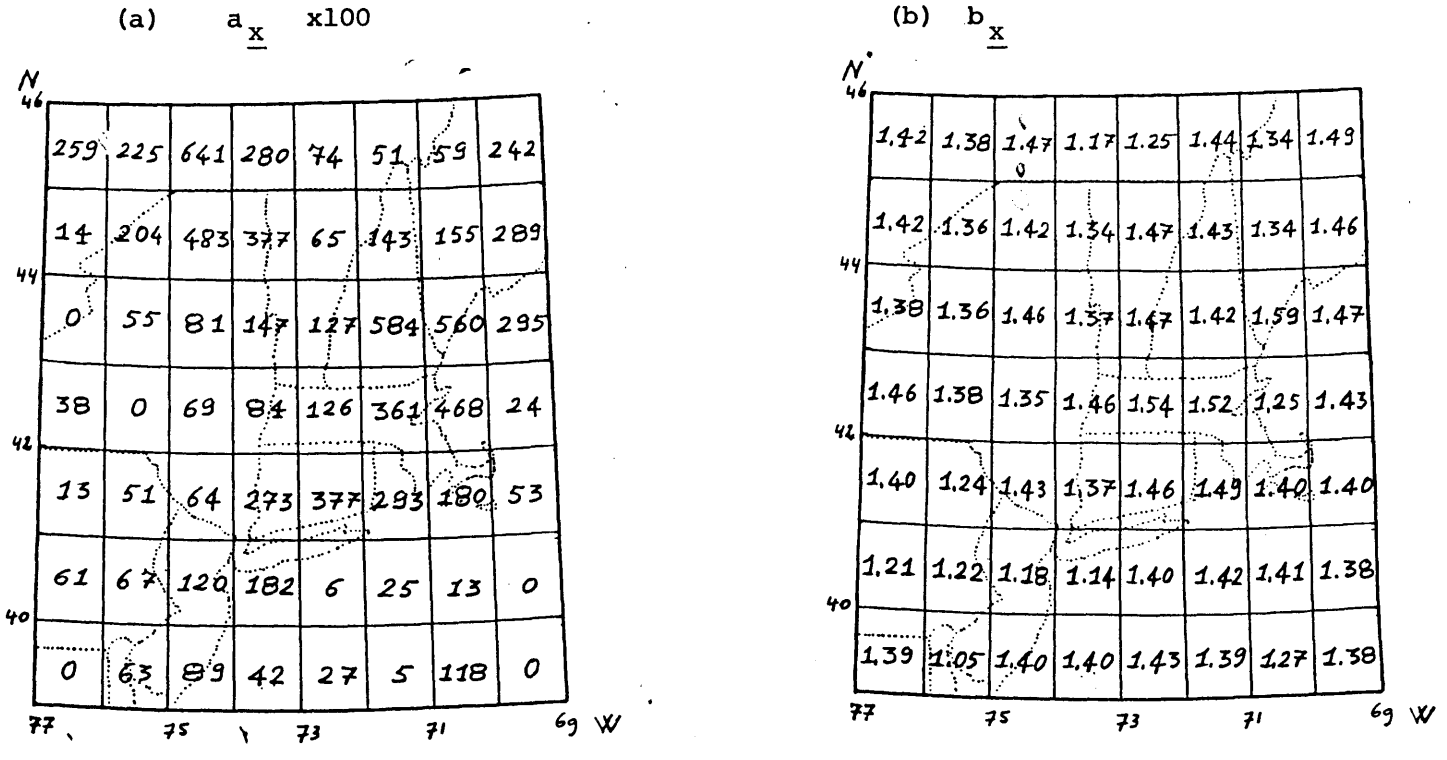


Figure 4.27 - Recurrence parameter estimates from base case analysis

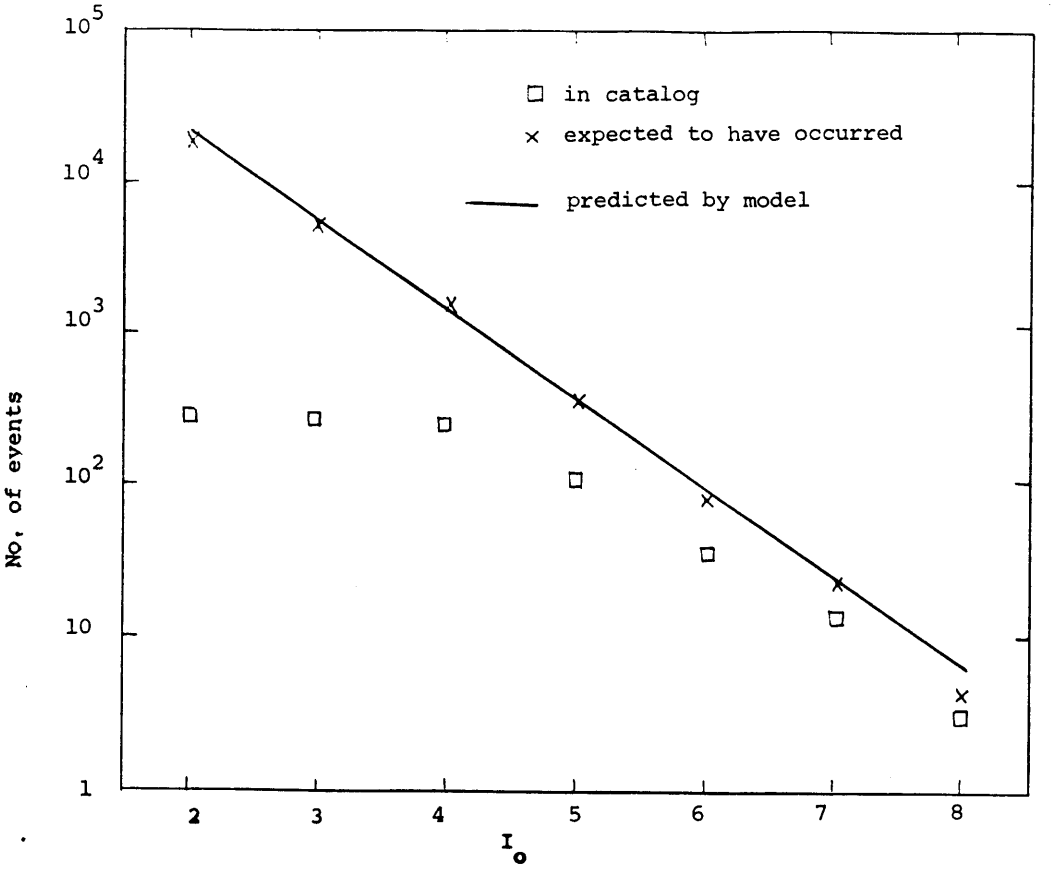


Figure 4.28 - Earthquake counts and expected counts for the entire region (base case analysis)

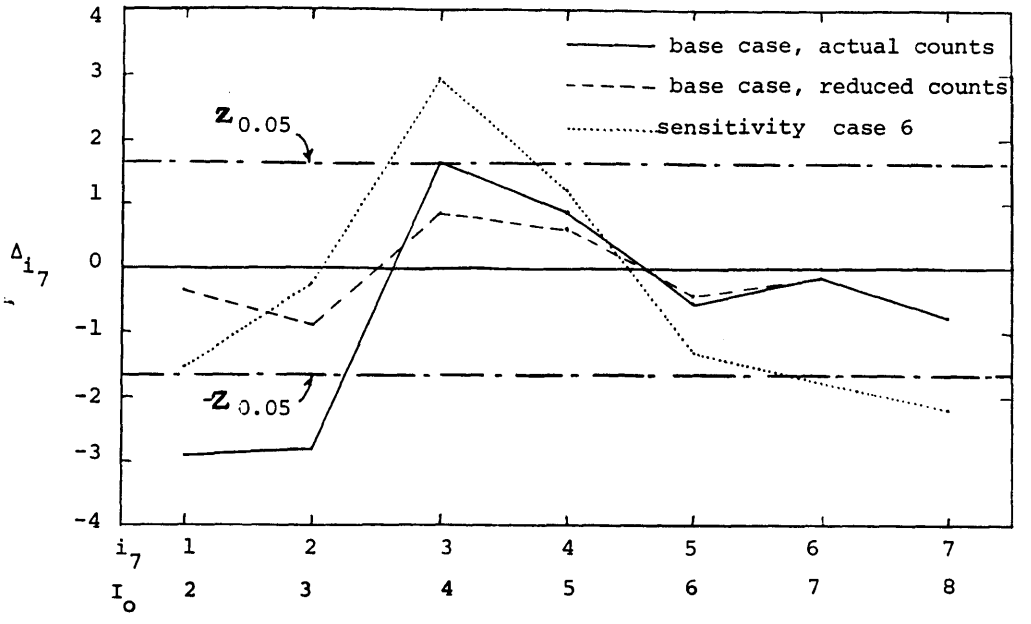
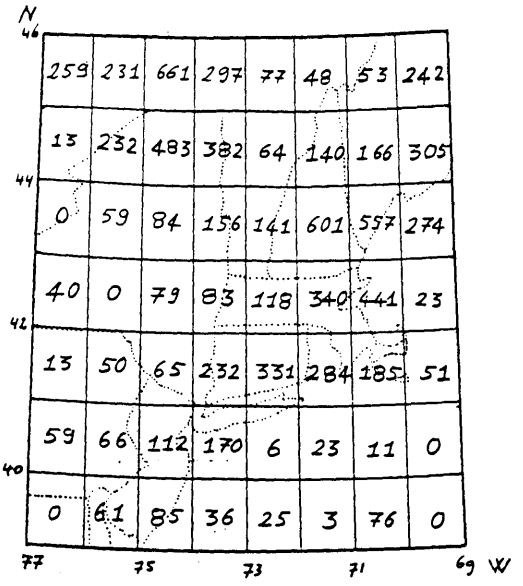
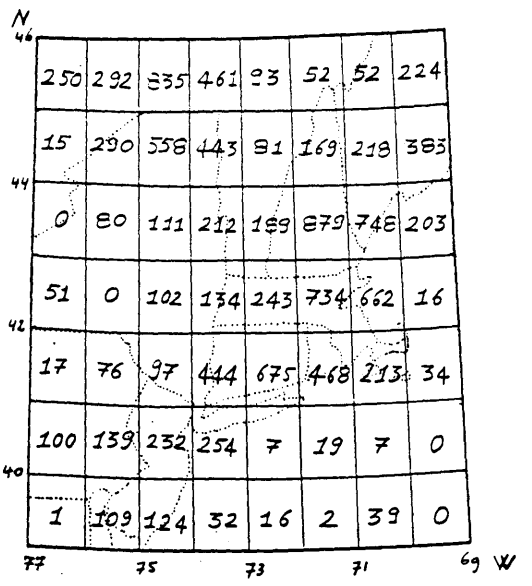


Figure 4.29 - Standardized residuals of expected observed counts for different analysis cases

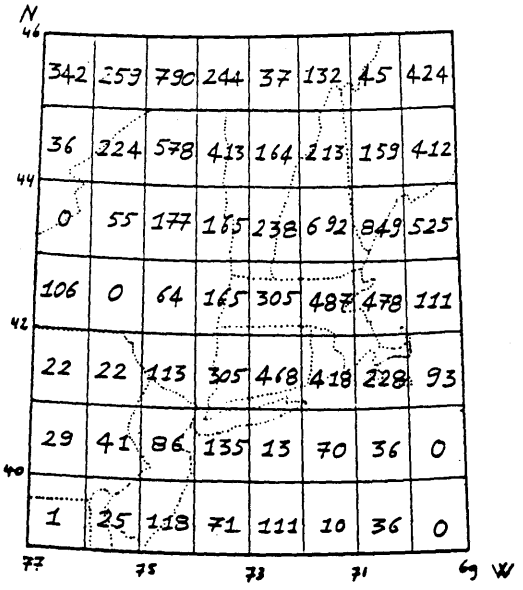
Case 1



Case 2



Case 3



Case 4

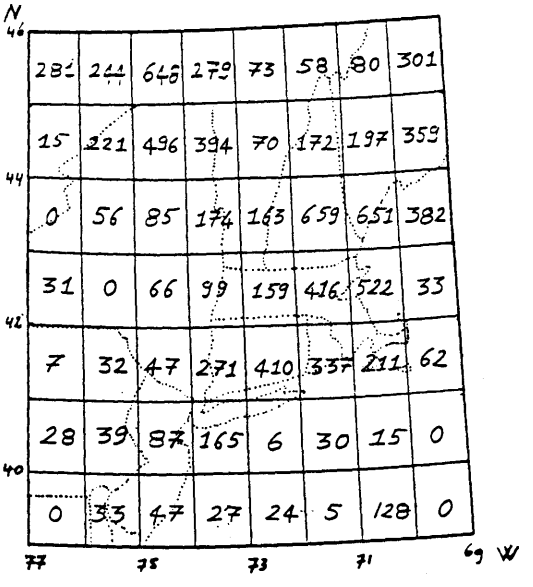
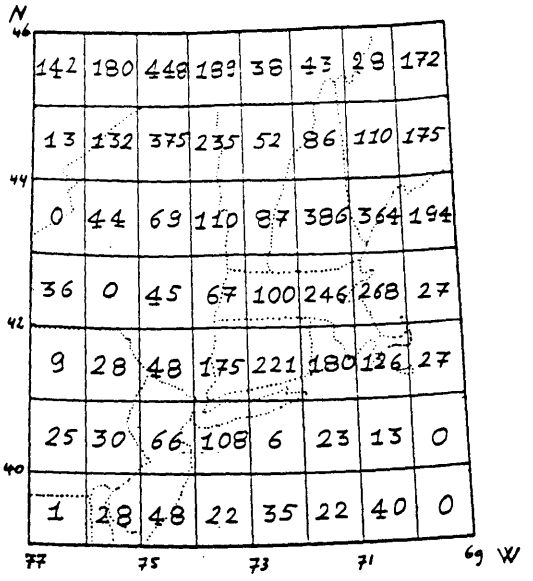
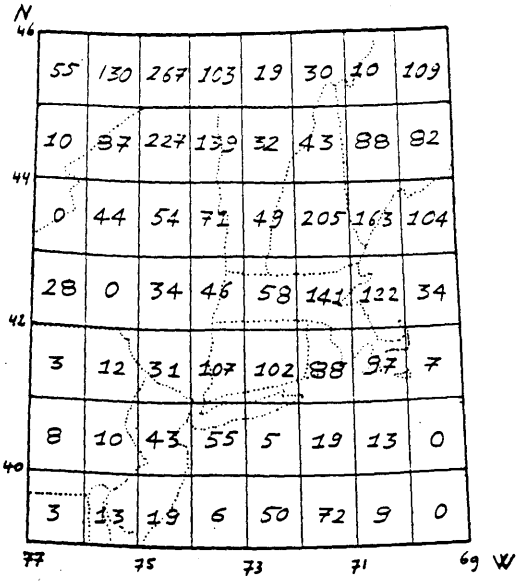


Figure 4.30 - Expected earthquake counts (in 100 years and per unit equatorial degree cell) at I_0=2 for different sensitivity cases

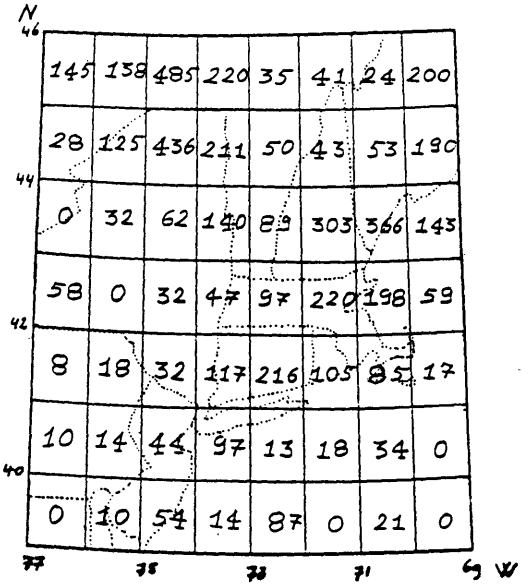
Case 5



Case 6



Case 7



Case 8

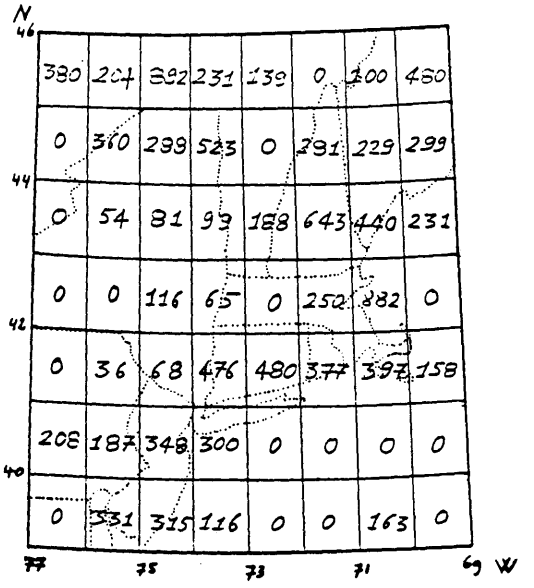
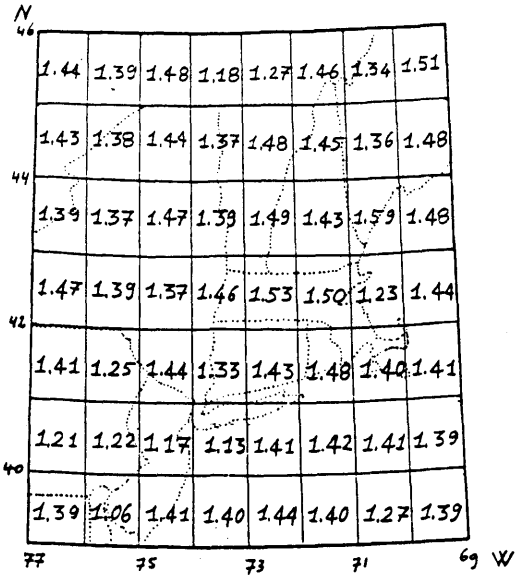
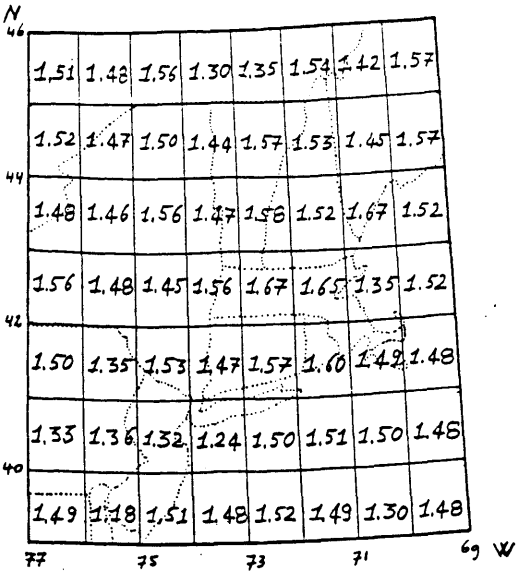


Figure 4.30 - (End)

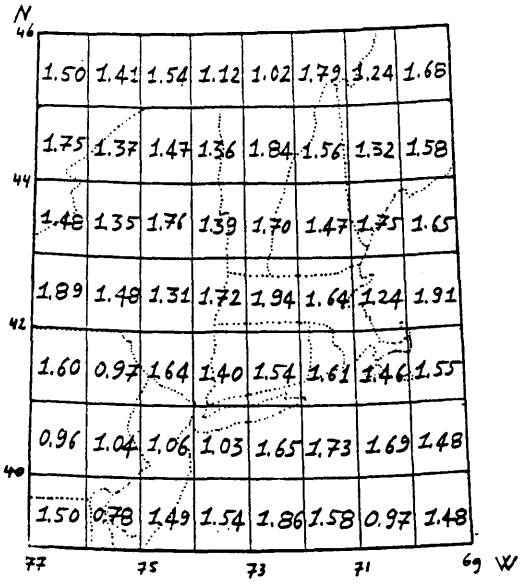
Case 1



Case 2



Case 3



Case 4

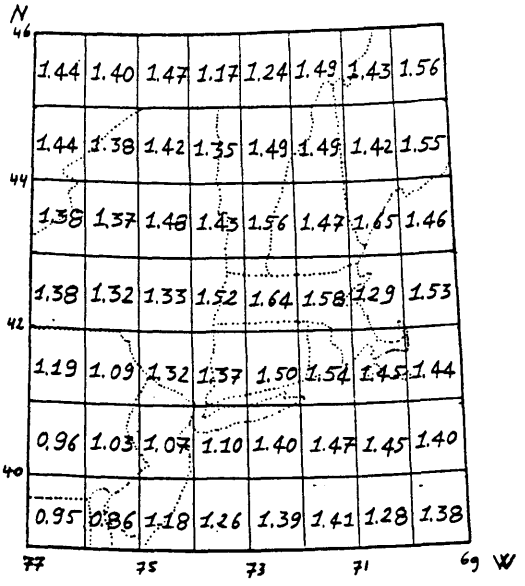
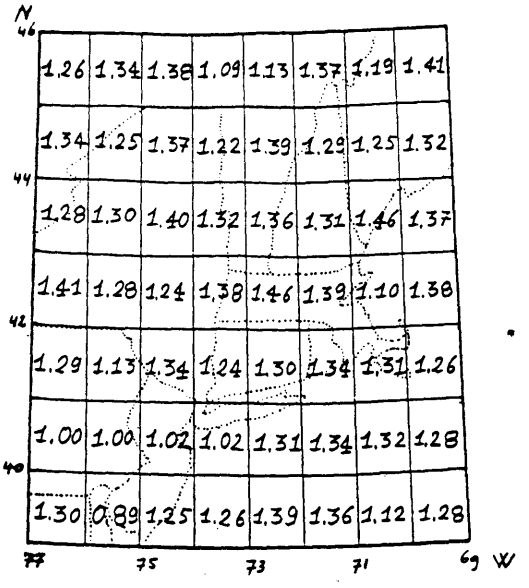
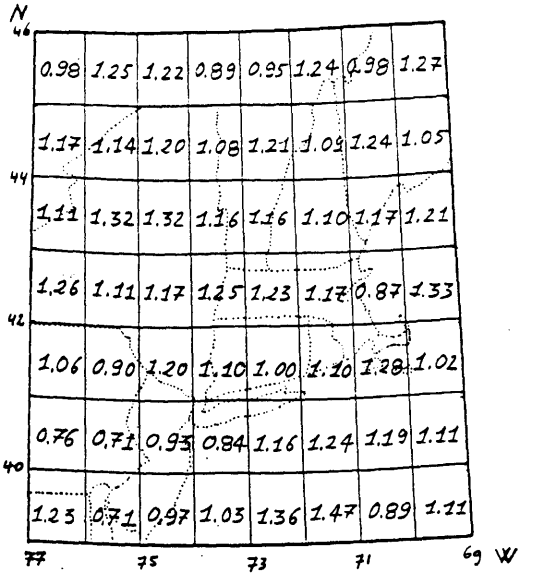


Figure 4.31 - Estimated $b_{\underline{x}}$ parameters for sensitivity cases

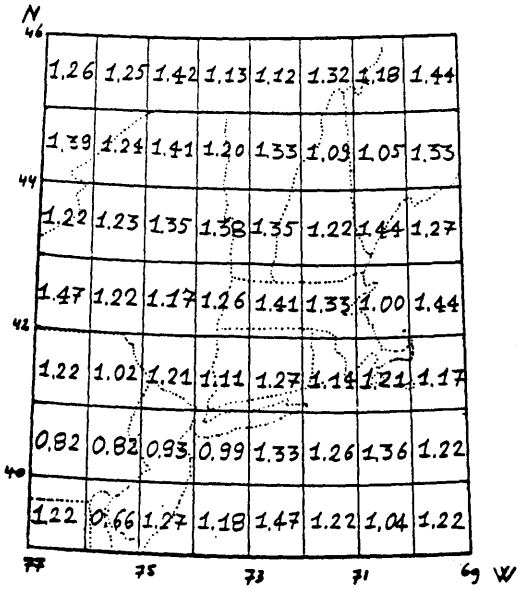
Case 5



Case 6



Case 7



Case 8

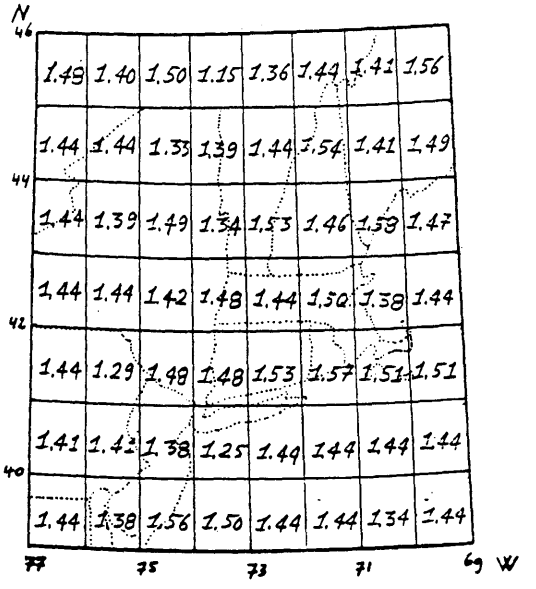


Figure 4.31 - (End)

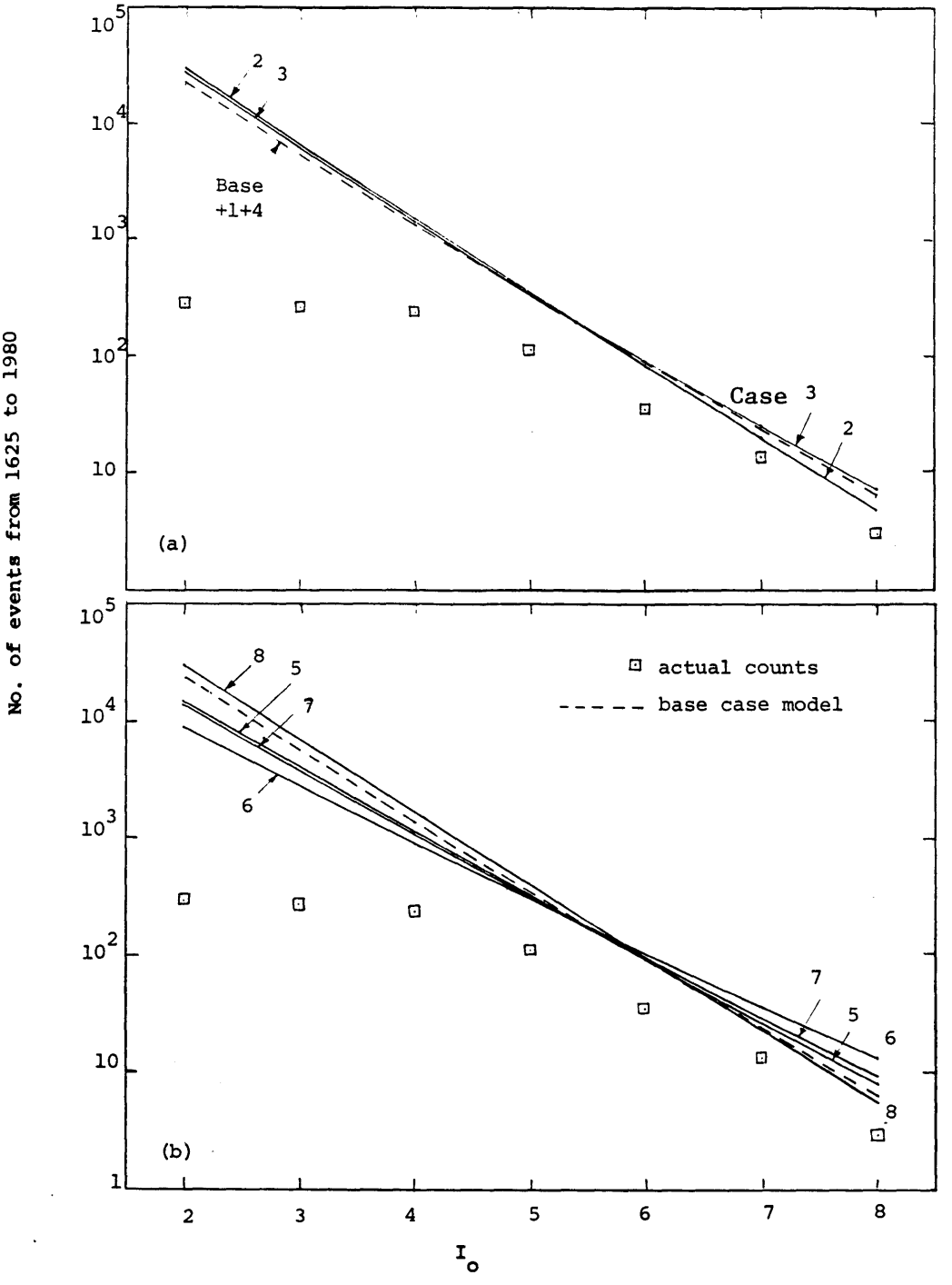


Figure 4.32 - Fitted counts, summed over the entire region, for different analysis cases

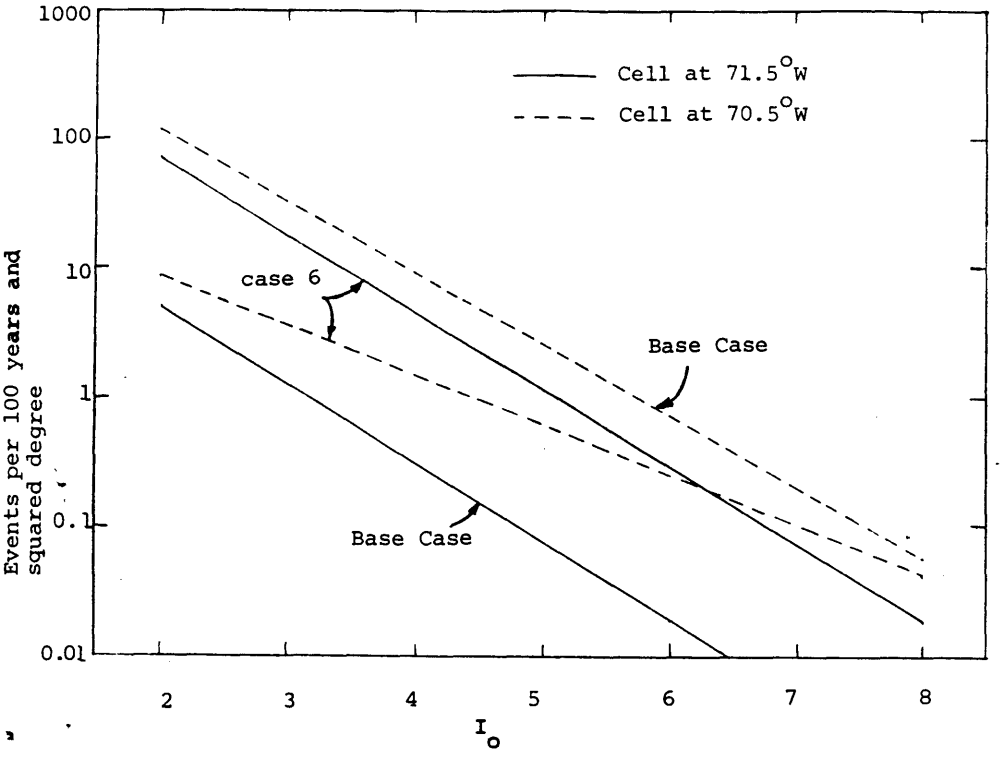


Figure 4.33 - Fitted exponential relation for two cells in Base Case and Case 6

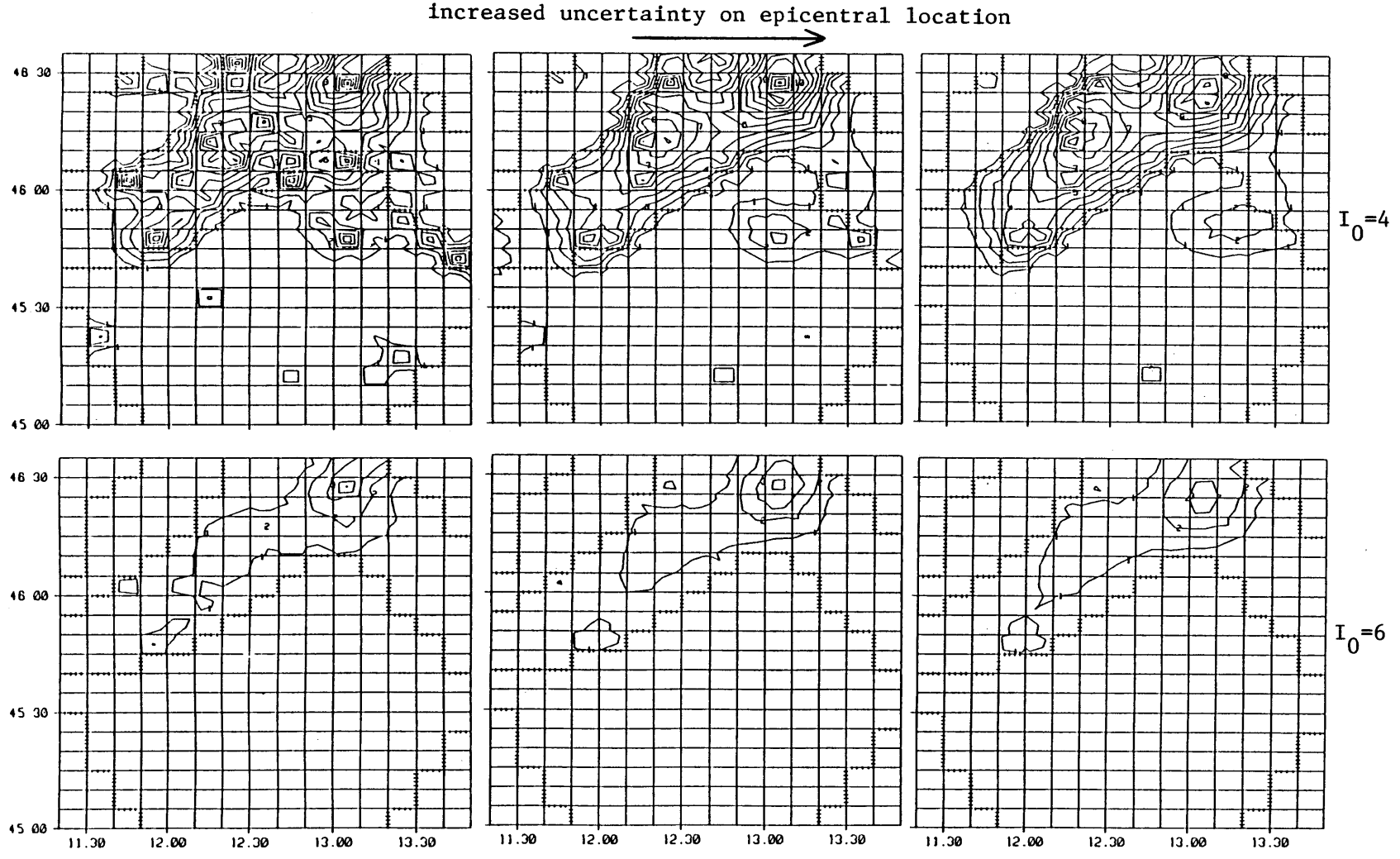
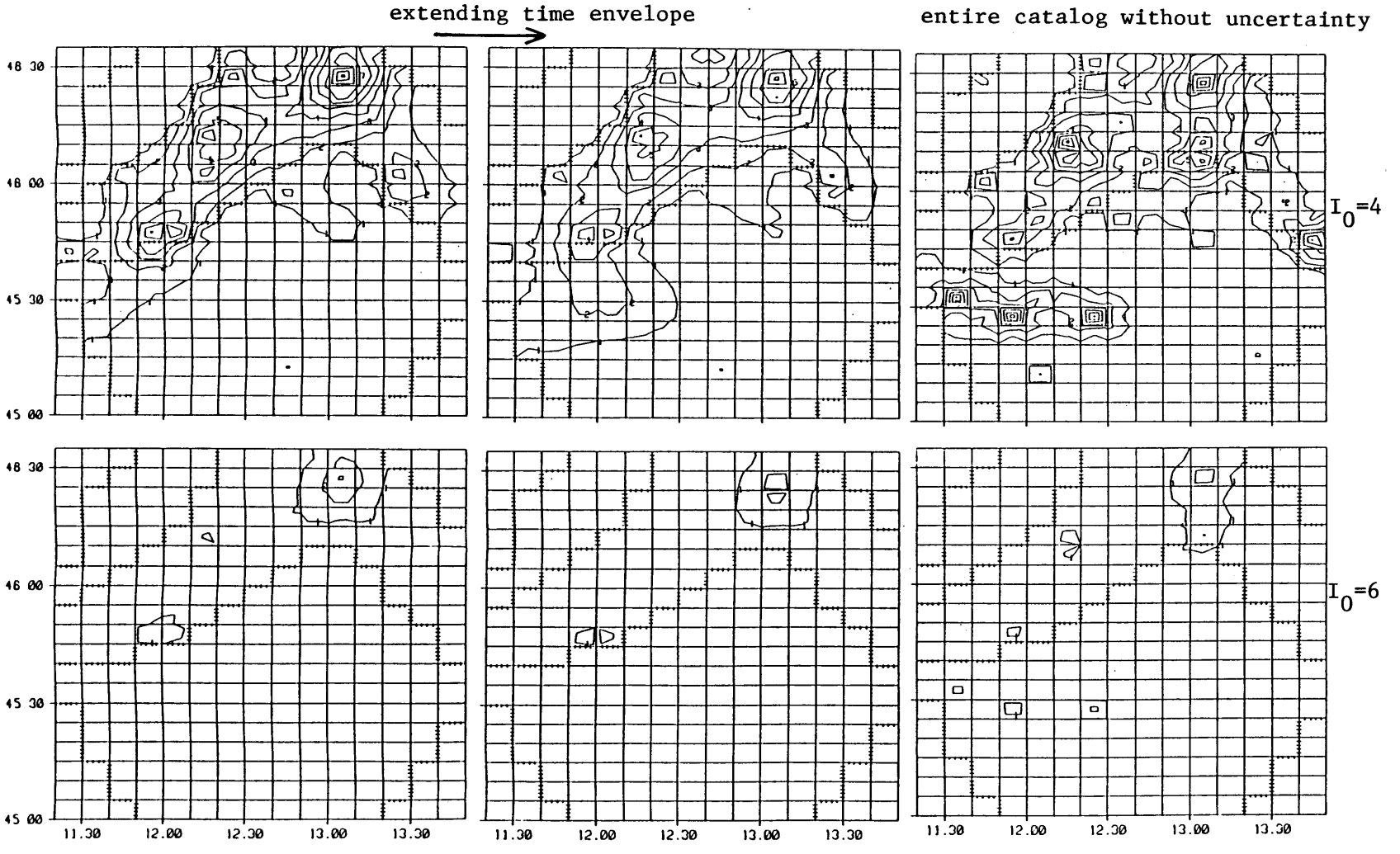
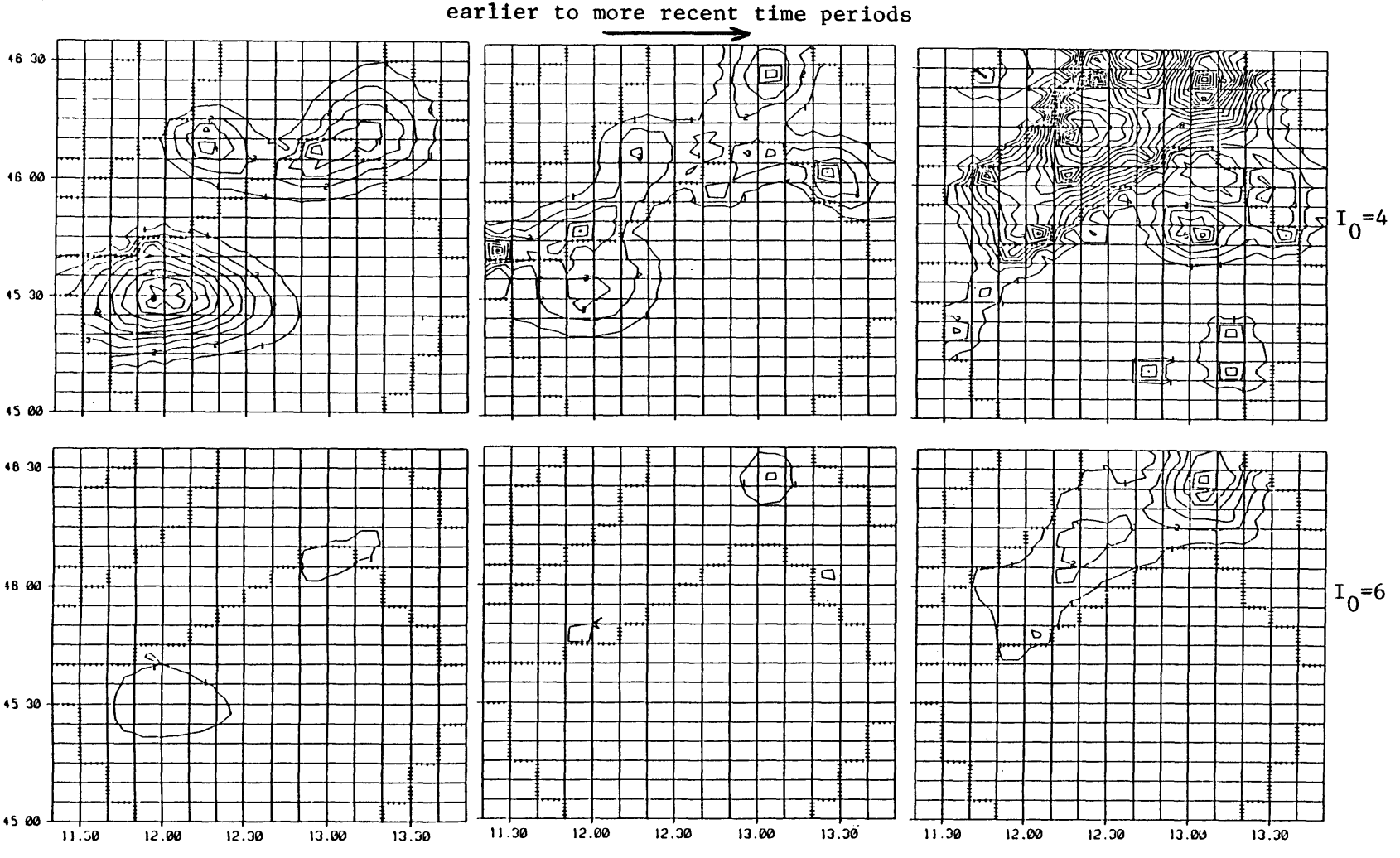


Figure 4.34 - Estimated recurrence rates for different sensitivity cases in Model C



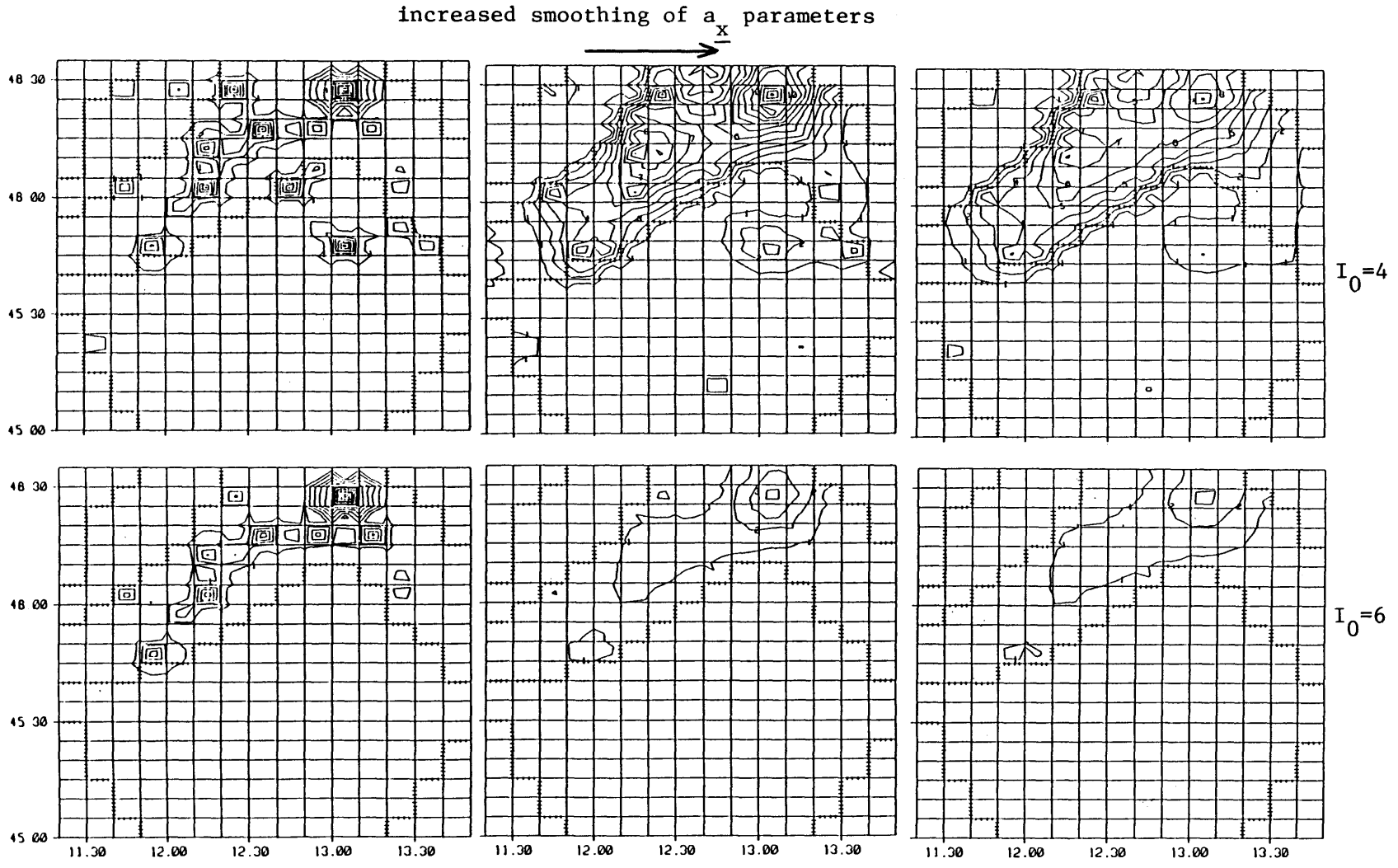
d. Case 4 e. Case 5 f. Case 6

Figure 4.34 - (Continued)



g. Case 7 h. Case 8 i. Case 9

Figure 4.34 - (continued)



j. Case 10

k. Case 2

1. Case 11

Figure 4.34 - (End)

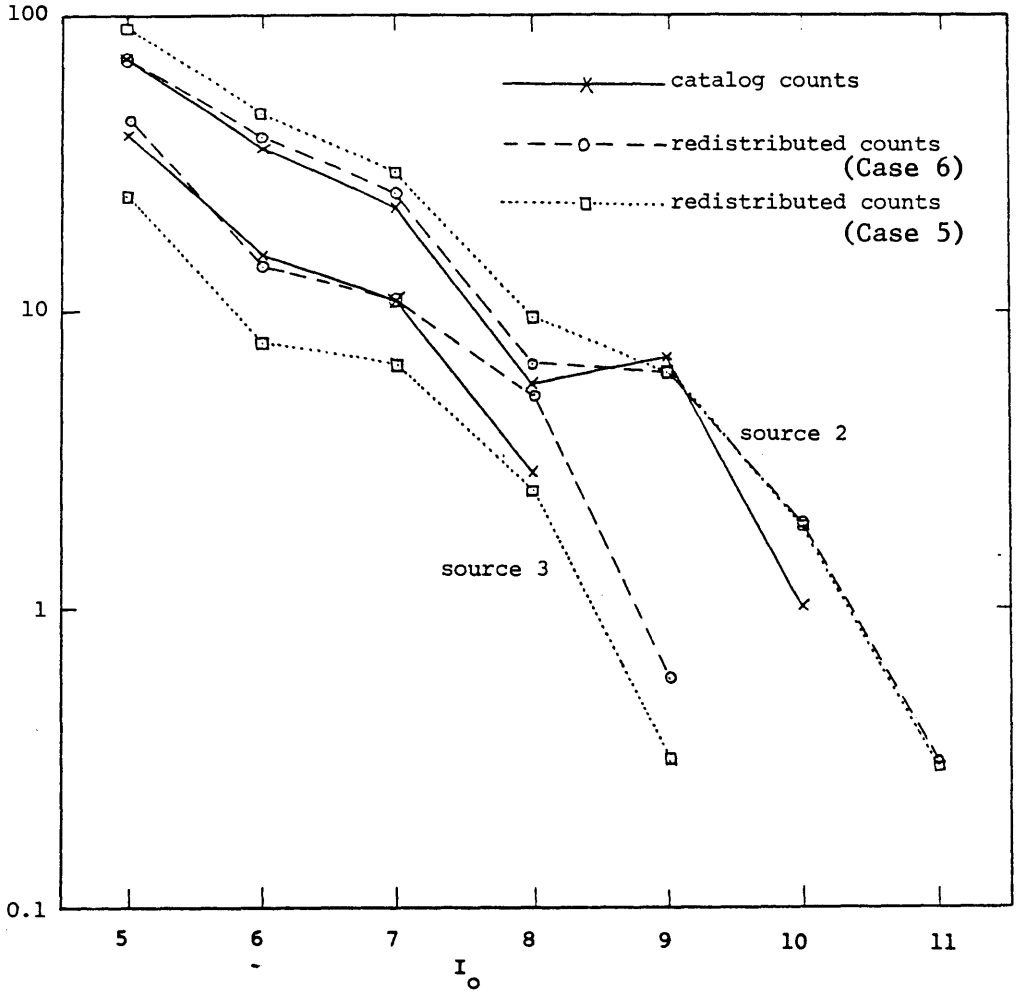


Figure 4.35 - Influence of uncertainty on location and size on the a-posteriori earthquake counts

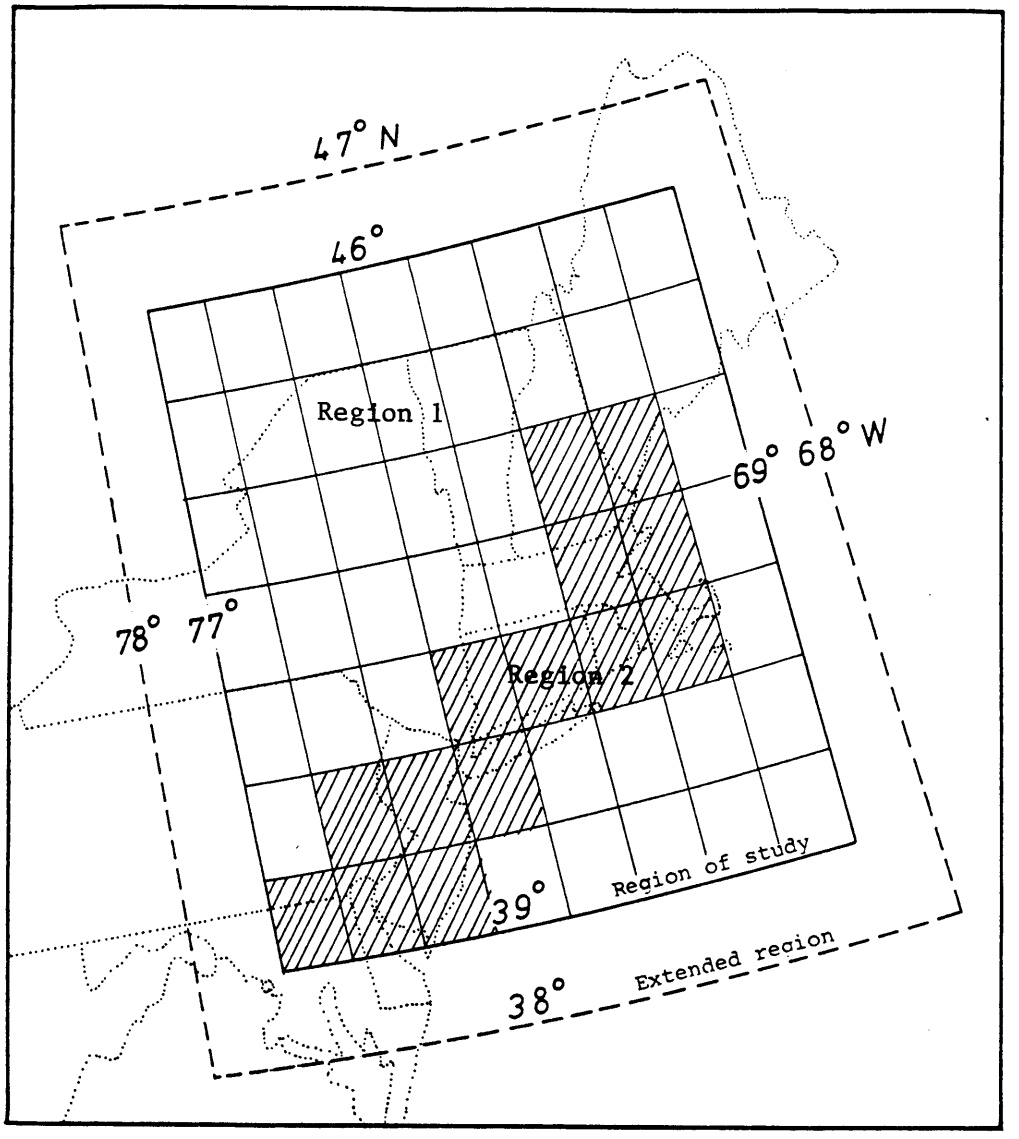
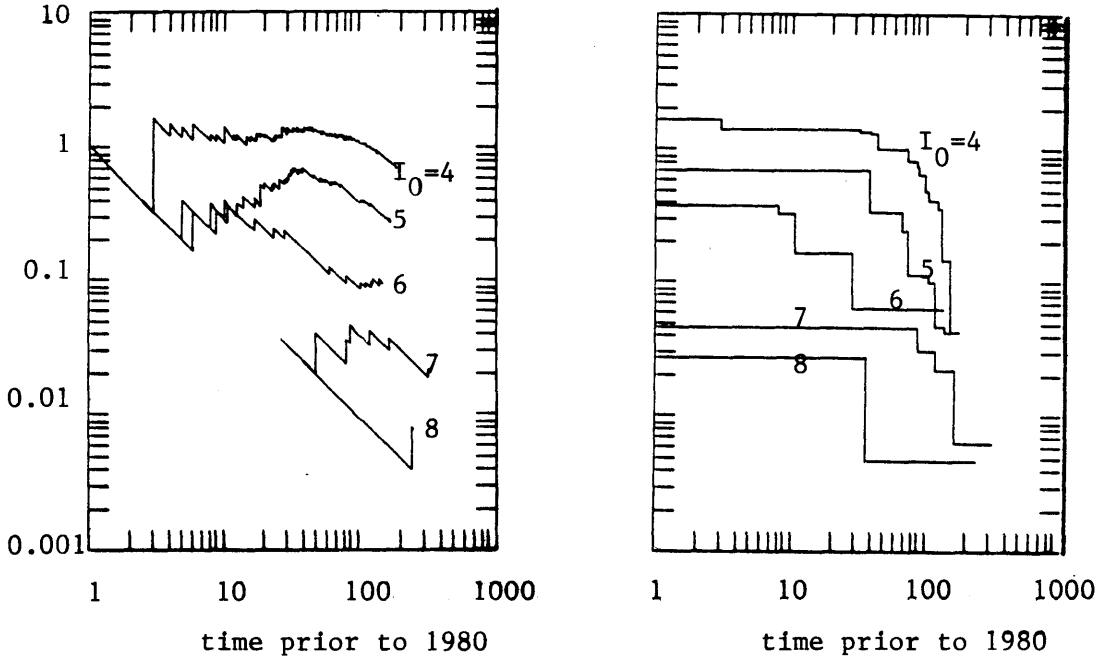
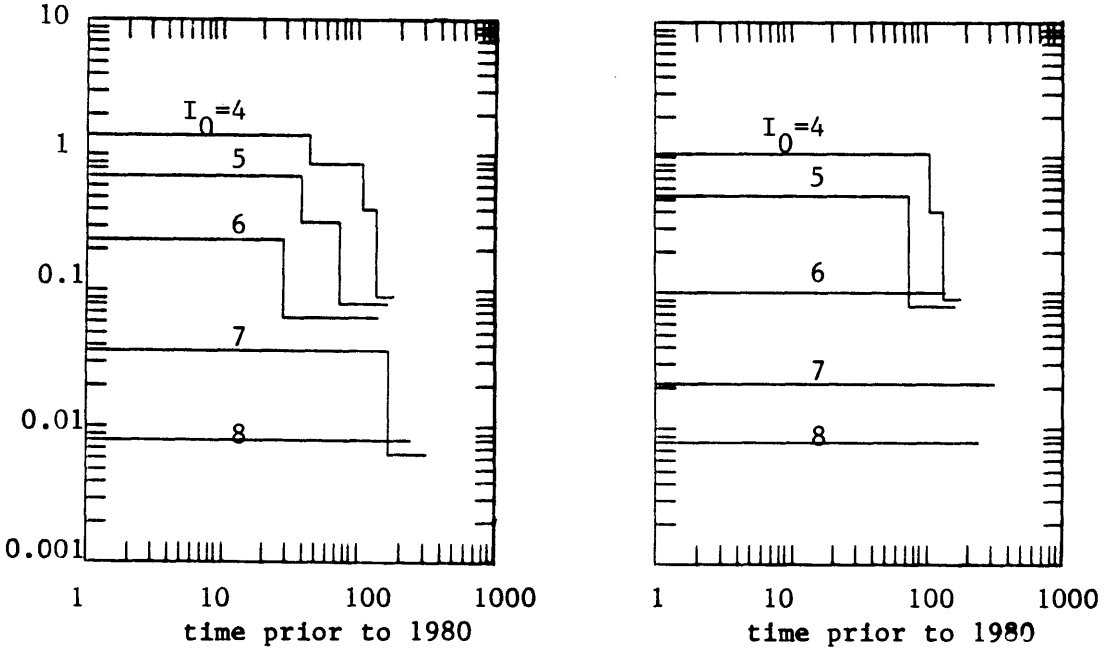


Figure 4.36 - Regions of uniform incompleteness



a. empirical recurrence rate for increasing observation periods

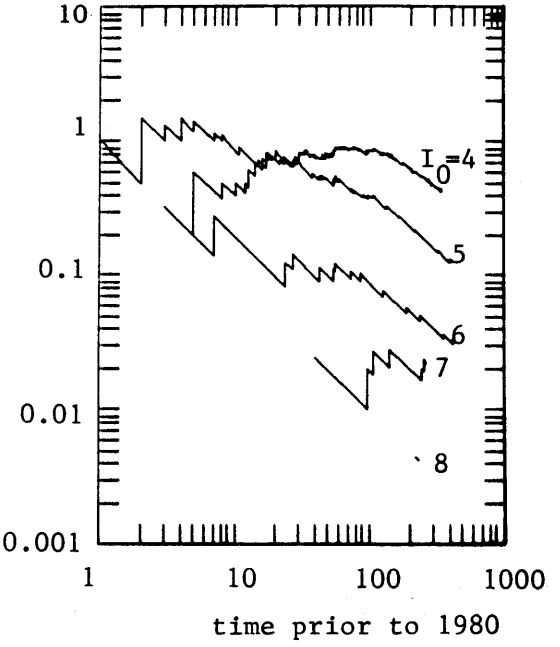
b. Estimate of monotonically decreasing recurrence rates as a function of time



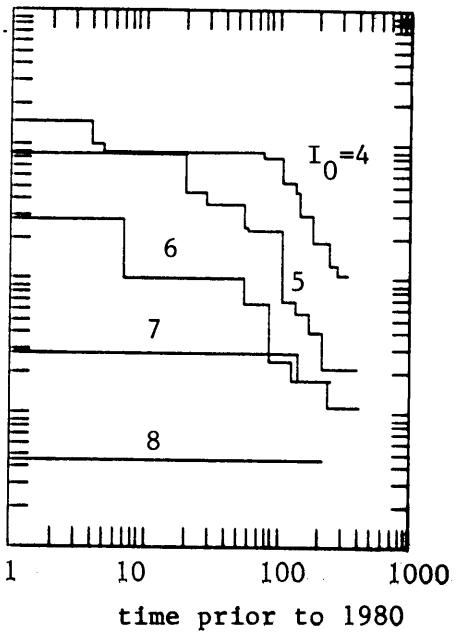
c. after removing non-significant deviations from b. (α=0.20)

d. after removing non-significant deviations from c. (α=0.05)

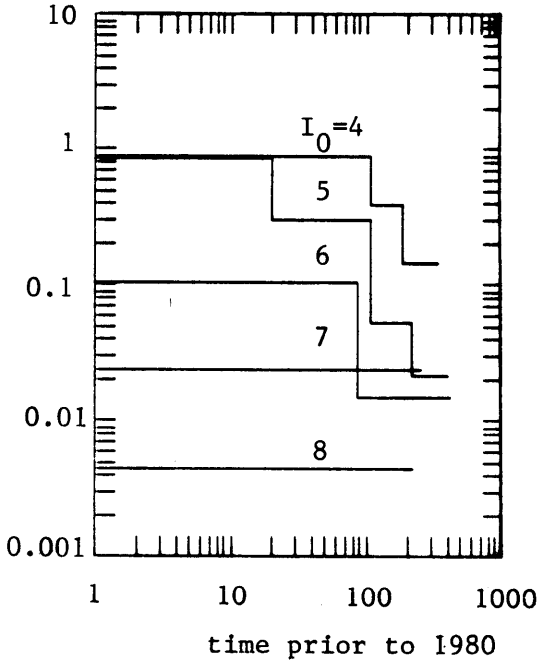
Figure 4.37 - Temporal variation of recurrence rates in completeness region 1.



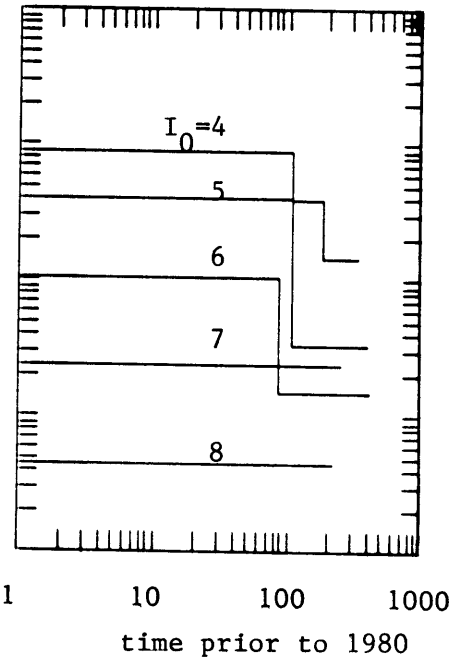
a. empirical recurrence rate for increasing observation periods



b. Estimate of monotonically decreasing recurrence rates as a function of time



c. after removing non-significant deviations from b. (α=0.20)



d. after removing non-significant deviations from c. (α=0.05)

Figure 4.38 - Temporal variation of recurrence rates in completeness region 2.

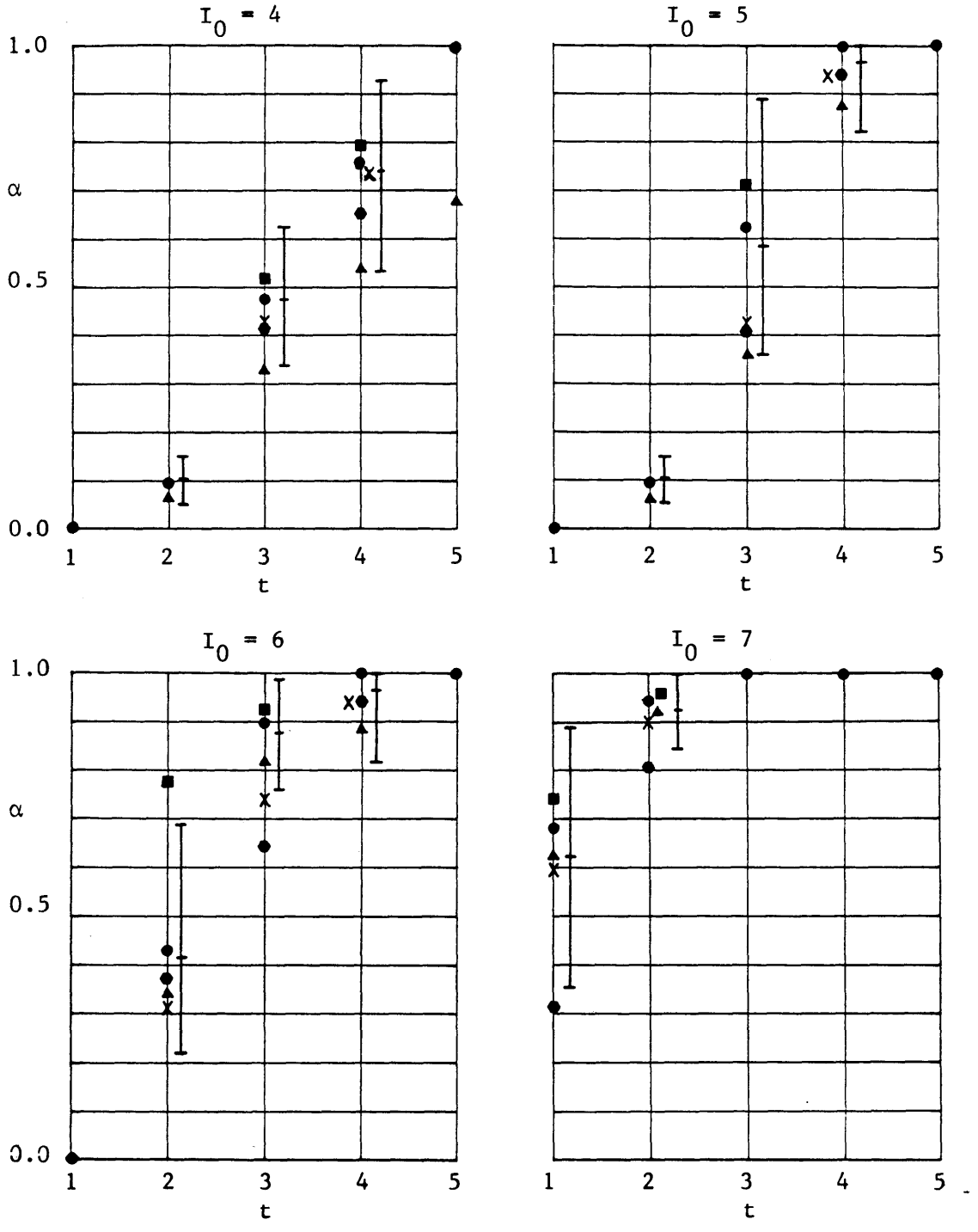


Figure 4.39a - Estimates of probability of detection for Completeness Region 1

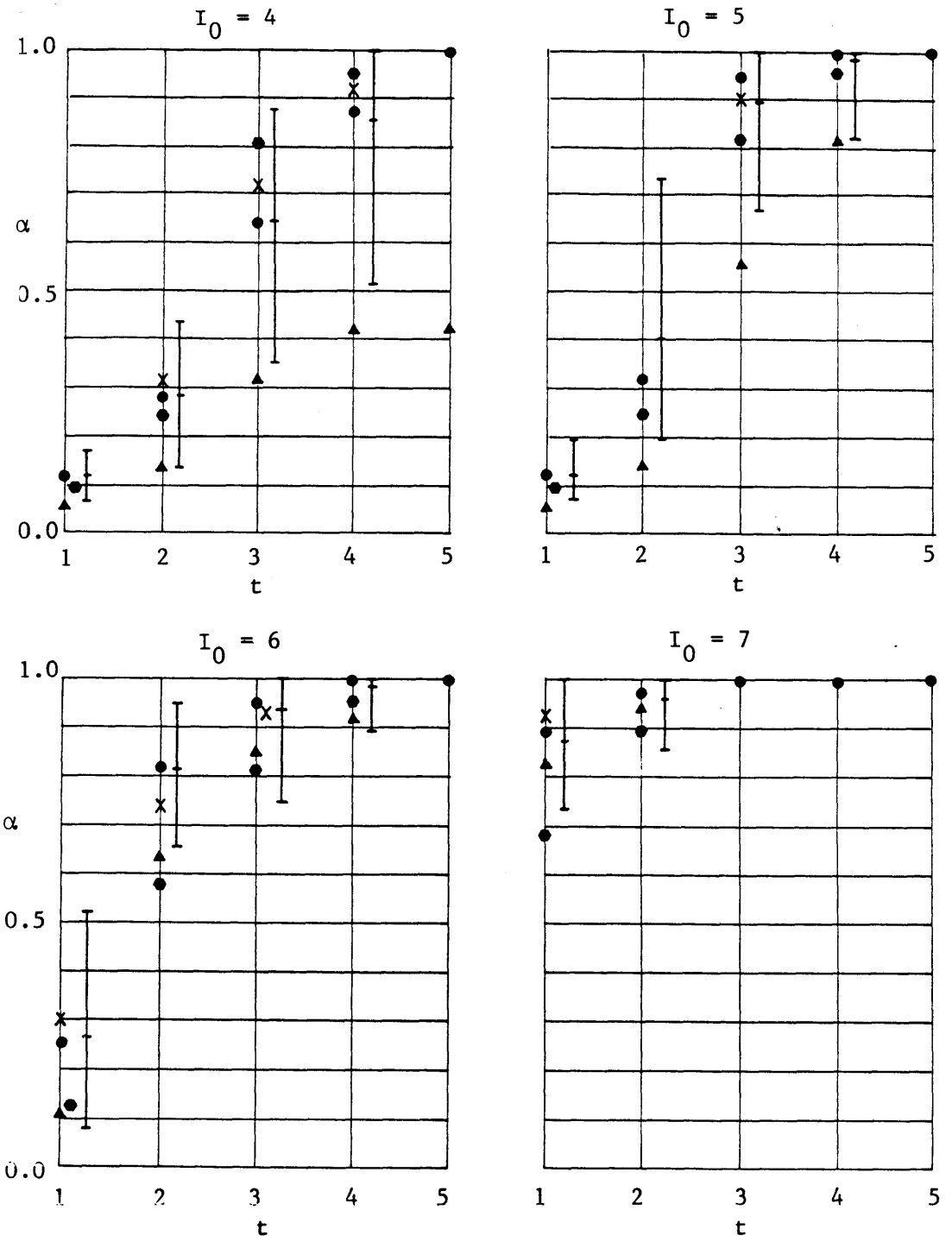
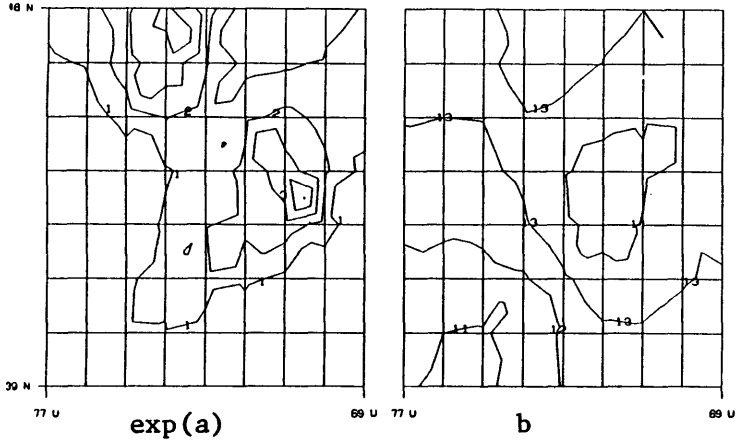
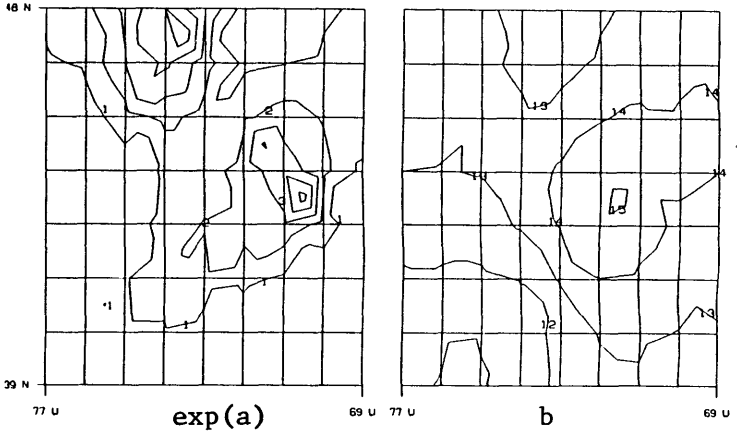


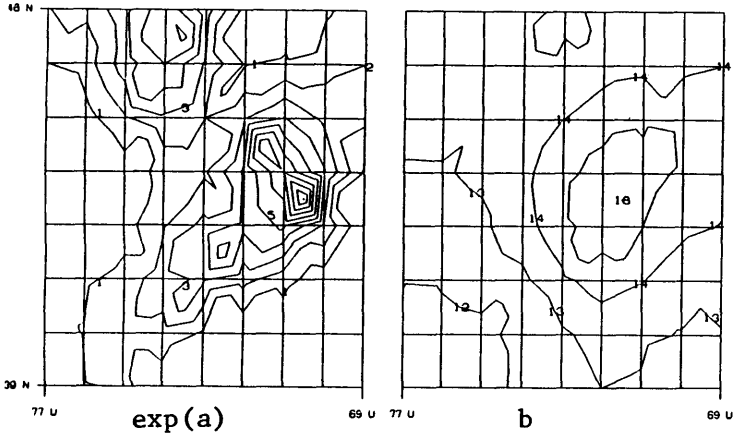
Figure 4.39b - Estimates of probability of detection for Completeness Region 2



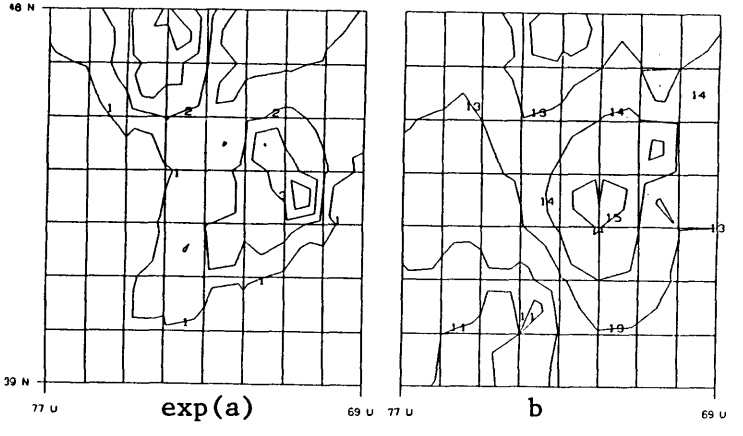
a. Reference Case



b. Case 1

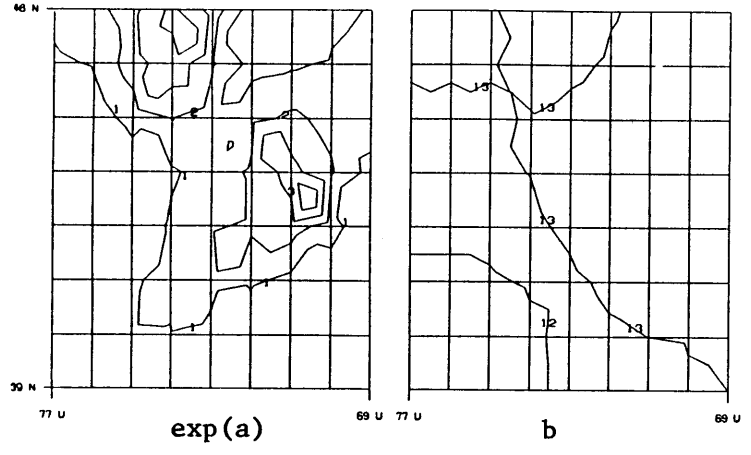


c. Case 4

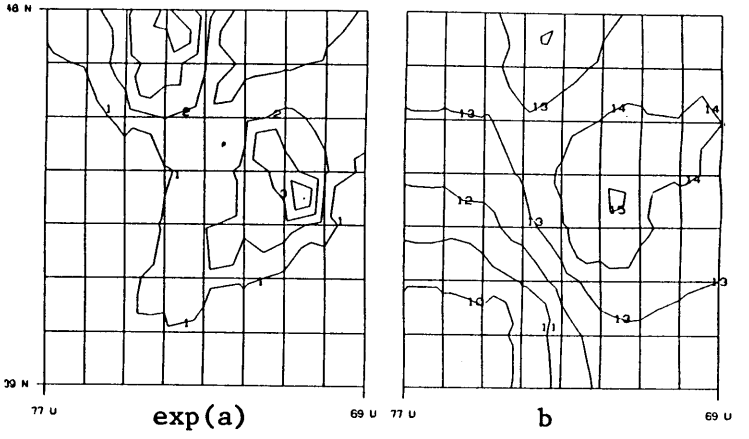


d. Case 5

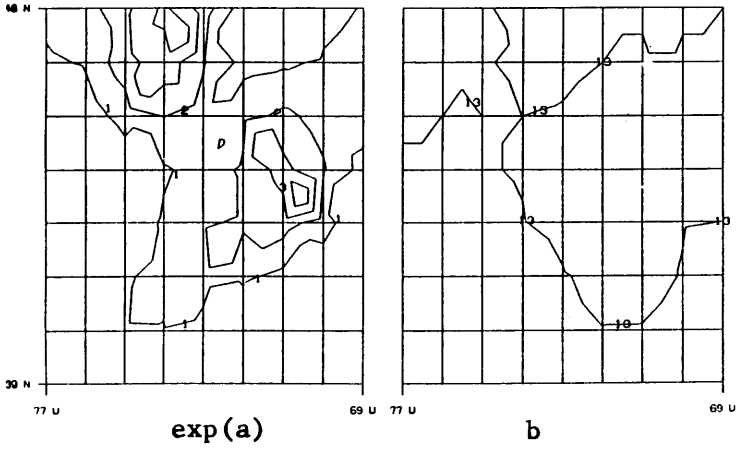
Figure 4.40 - Contourplots of the recurrence parameter estimates for different cases in Model D



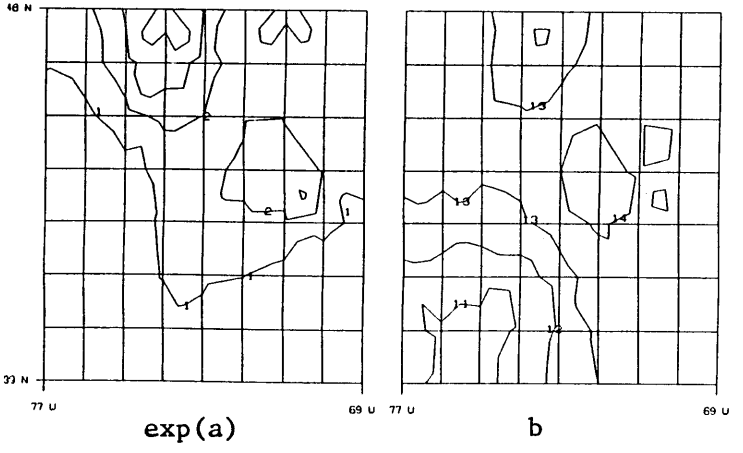
e. Case 6



f. Case 7

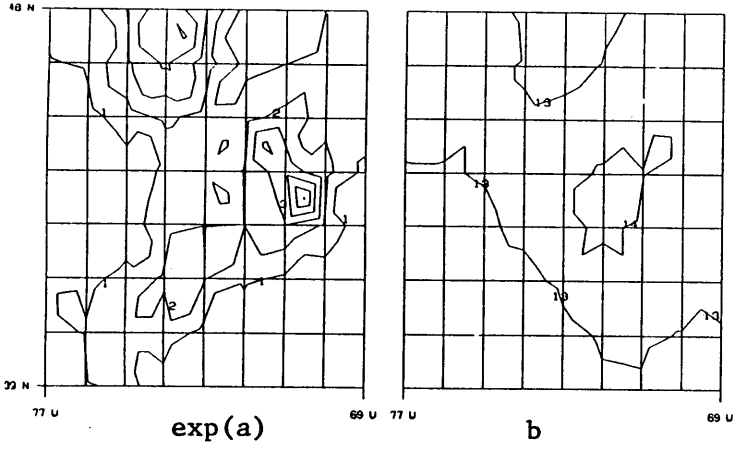


g. Case 8

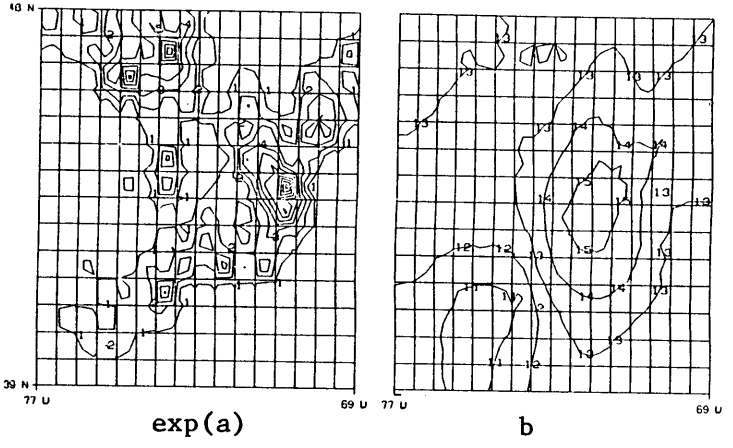


h. Case 9

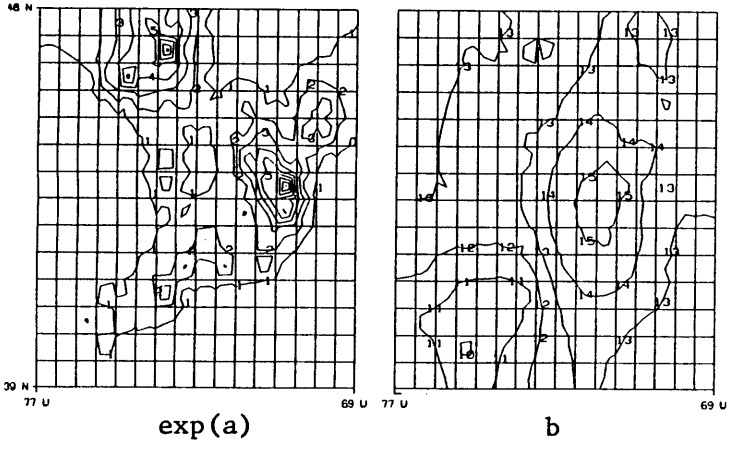
Figure 4.40 - (Continued)



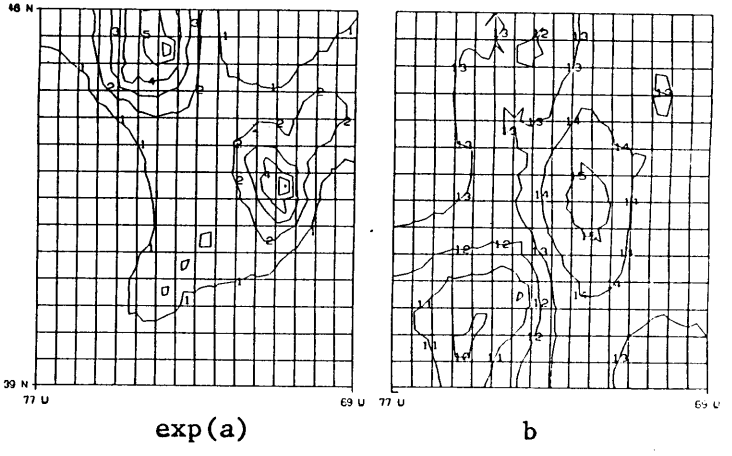
i. Case 10



j. Case 11



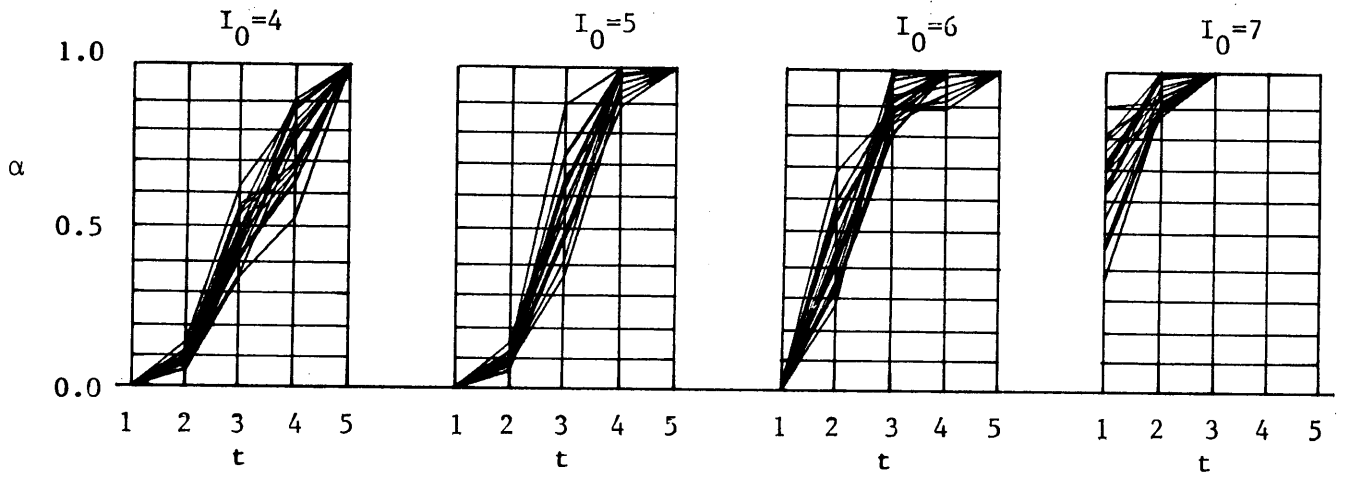
k. Case 12



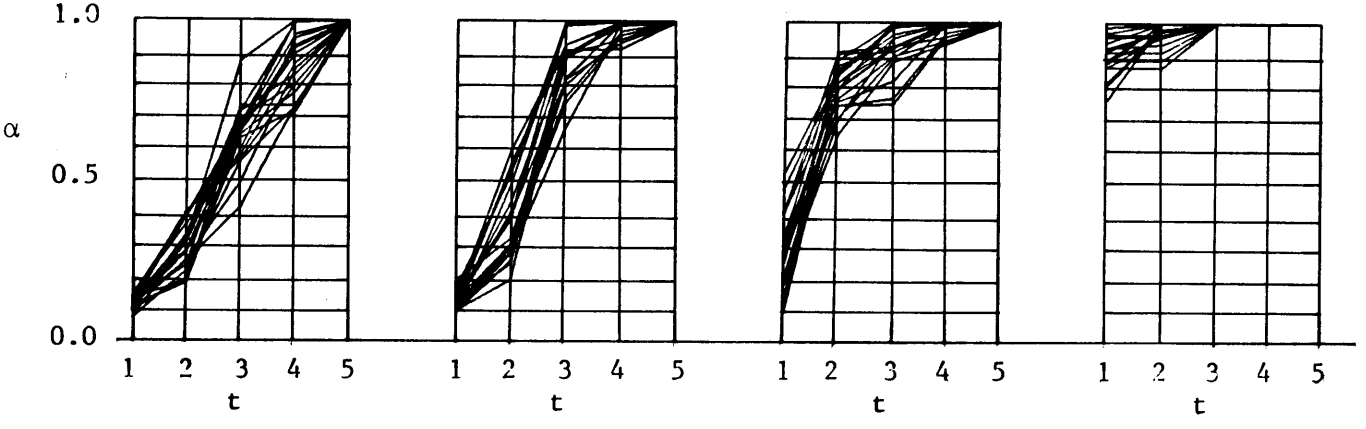
l. Case 13

Figure 4.40 - (End)

a. probability of detection



COMPLETENESS REGION 1



COMPLETENESS REGION 2

Figure 4.41 - Parameter estimates in the first 20 samples of empirical bootstrapping

b. recurrence rate estimates

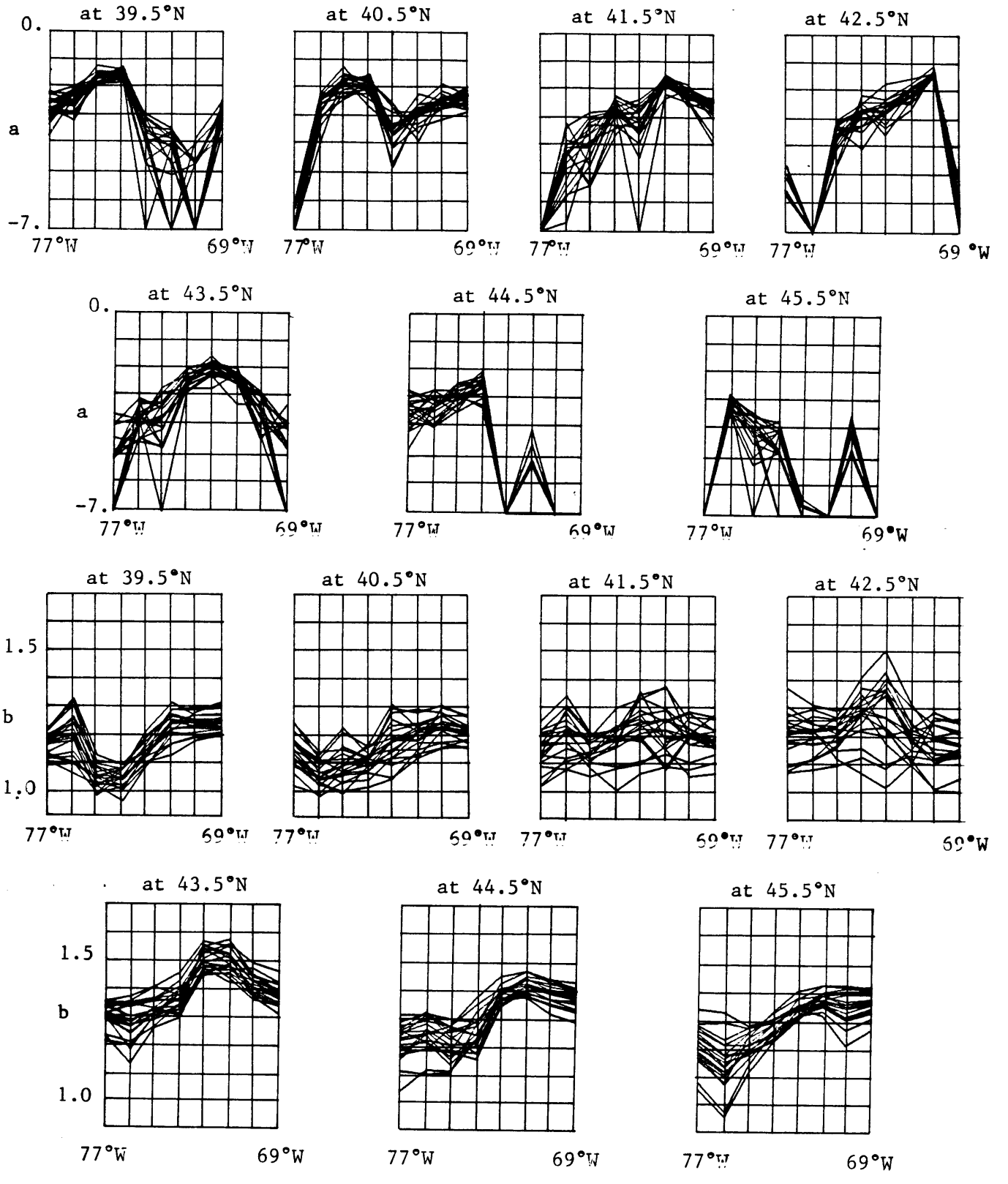
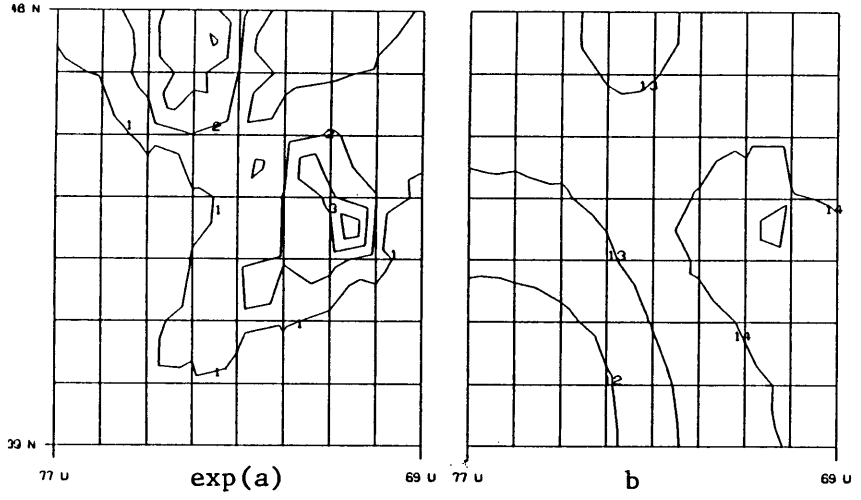
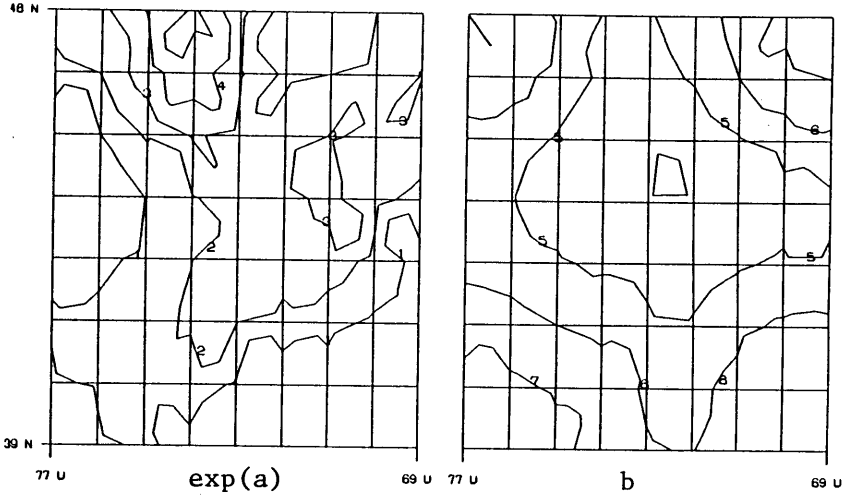


Figure 4.41 - (End)

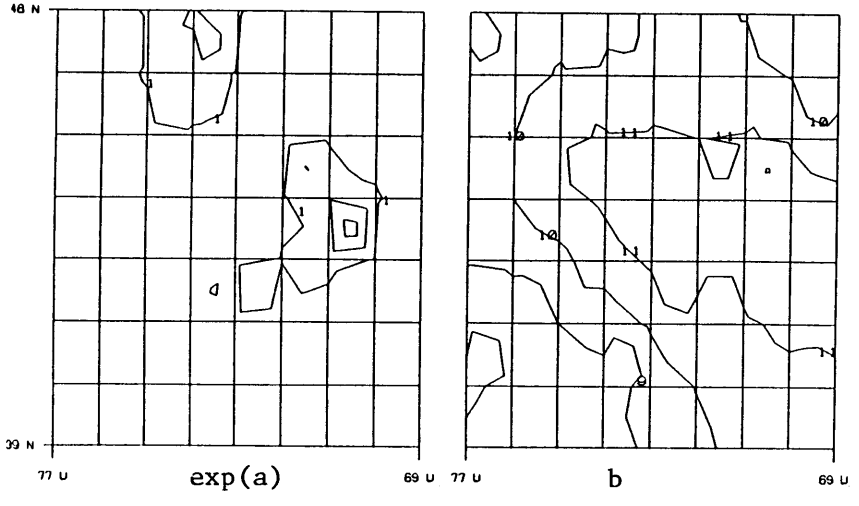
sample average



sample standard deviation



sample minimum



sample maximum

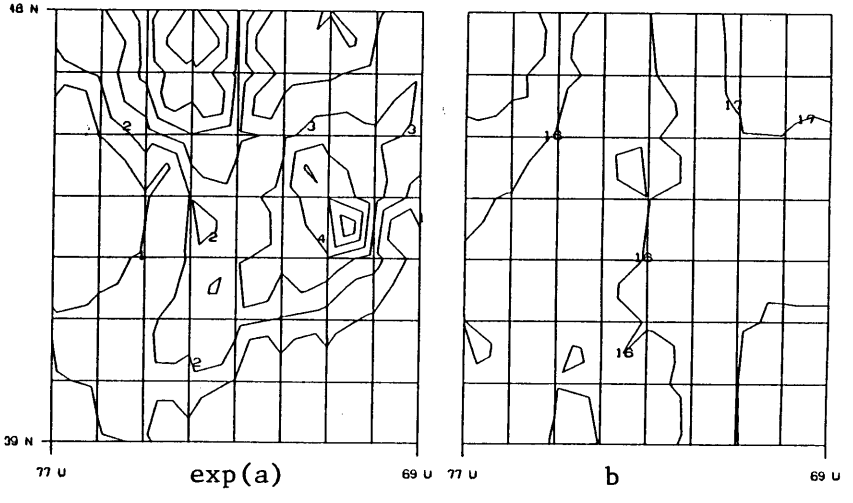


Figure 4.42a - Contourplots of sample statistics for recurrence parameter estimates (parametric bootstrapping)

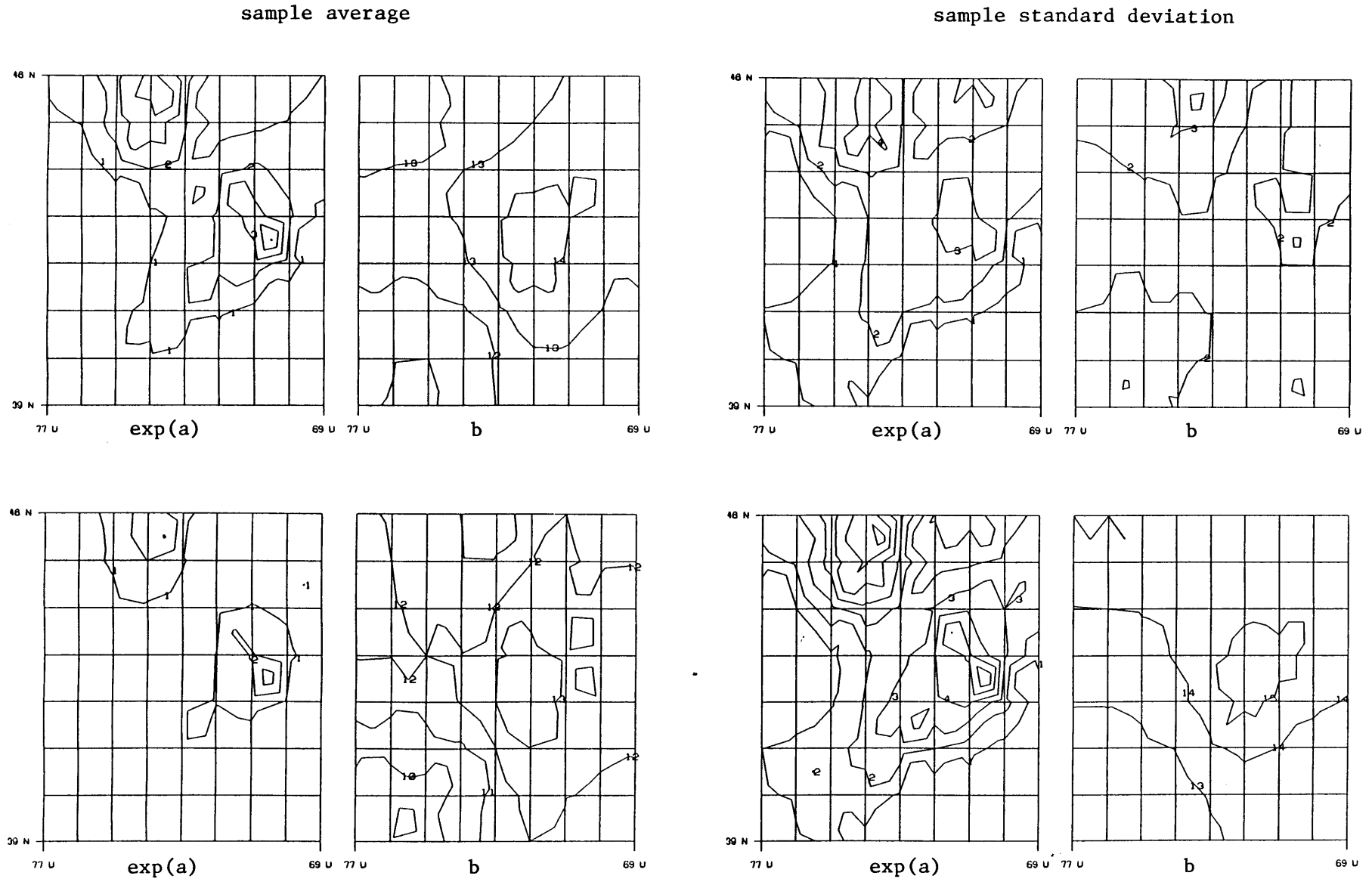
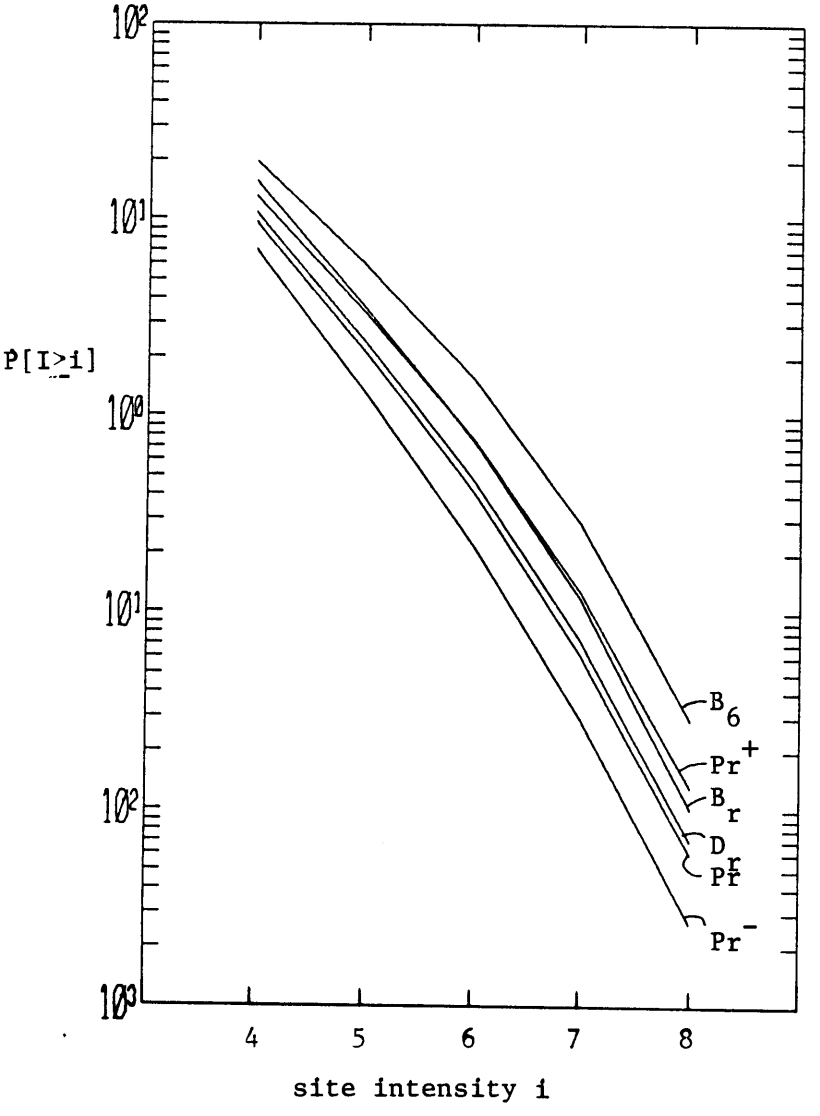


Figure 4.42b - Contourplots of sample statistics for recurrence parameter estimates (empirical bootstrapping)



D_r : reference case of Model D

B_r : reference case of Model B

B_6 : Case 6 of Model B

Pr, Pr^+, Pr^- : average and $\pm 2\sigma$ from parametric bootstrapping for Model D

Figure 4.43 - Illustration of sensitivity of seismic hazard to model assumptions