

*File copy only
Folder 42
Drawn 15 (16)*

**THE NCEER-91 EARTHQUAKE CATALOG:
IMPROVED INTENSITY-BASED MAGNITUDES
AND
RECURRENCE RELATIONS
FOR U.S. EARTHQUAKES EAST OF NEW MADRID
(Technical Report LDGO-91-3)**

by

Leonardo Seeber and John G. Armbruster

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York
10964

ABSTRACT

A new catalog of earthquakes in the eastern U.S. named The NCEER-91 Earthquake Catalog stems from existing compilations and from new data. NCEER-91 is derived from the EPRI (1987) catalog (EPRI), and is based on new data from archival searches and on published compilations not fully incorporated in EPRI. This paper presents results from work in progress and is primarily concerned with earthquake magnitudes. The data discussed is from the area east of the New Madrid seismic zone (east of 85.5W and north of 30.0N) and magnitudes $M \geq 3$. In this area, NCEER-91 drastically reduces the number of magnitudes based on maximum intensity (Mmi) by providing data for more reliable felt-area magnitudes (Mfa), most of them during the periods of completeness. Many Mmi were overestimated, particularly in EPRI/V, where intensity magnitudes are computed according to Veneziano and Van Dyck (1985). The bias is manifested by higher rates of $M=3-4$ events in the period when magnitudes are primarily determined from intensity data than in the instrumental period. This anomaly is largely eliminated in NCEER-91 and is partially eliminated in EPRI/S, where magnitudes are calculated from EPRI data using relationships proposed by Sibol et al. (1987). The b-values and repeat times vary more smoothly with different low-magnitude limits and are internally more consistent in NCEER-91 than in EPRI/V. Differences in b-value and repeat times between NCEER-91 and EPRI/S, however, are insignificant. The assignment of seismicity rates and completeness periods for a set of magnitude windows is a subjective step in the procedure which introduces large uncertainties. Results also depend on the choice of algorithm to line-fit the magnitude distribution. If the assumption of linearity in this distribution can be taken for granted, the preferred approach is a maximum likelihood fit where statistically more significant data at low magnitudes are preferentially weighted. On the other hand, if deviations from a linear distribution are plausible, as suggested by geologic and earthquake data from the Manhattan Prong, then extrapolation should be minimized and recurrence rates at large magnitudes should rely more on data at intermediate than at small magnitudes. In this case, a least-squares fit with appropriate weighting may be desirable. While error estimates from maximum-likelihood fits are commensurable with variations derived from alternative procedures, error estimates from least-squares fits tend to be unrealistically small. Overall rates of seismicity in sub-regions (1) east and (2) west of the Appalachian crystalline front are similar and no significant differences in their statistical parameters are detected. Considering statistical confidence limits as well as variations derived from alternative

procedures, we conclude that b-values are 1.05 ± 0.05 and repeat times for $M \geq 6$ are 200 ± 100 years in each of these regions.

ACKNOWLEDGEMENTS

Klaus Jacob reviewed the manuscript and improved it with stimulating discussions and many helpful suggestions. The analytical work and part of the searches for historic data in this paper were supported by the National Center for Earthquake Engineering Research (NCEER grant 90-1303). Searches for historical earthquake data were also funded by National Science Foundation grants EAR 82-05860 and 83-16589 and Nuclear Regulatory Commission grant NRC G-04-84-010

TABLE OF CONTENTS

<i>SECTION</i>	<i>TITLE</i>	<i>PAGE</i>
1	INTRODUCTION.....	1-1
1.1	Earthquake Data for Definition of Seismic Sources -- Intraplate versus Interplate Environments	1-1
1.2	Improved Constraints on Magnitudes and their Effects on Recurrence Relations	1-2
1.3	Alternative Analytical Procedures and Source Zones: Effects on Recurrence Parameters.....	1-3
2	SOURCES OF BIAS IN THE CATALOG RELEVANT TO EARTHQUAKE STATISTICS AND HAZARD ESTIMATES	2-1
2.1	Maximum-Intensity versus Felt-Area Magnitudes	2-1
2.2	Systematic Bias in Magnitudes Derived from Intensity Data	2-2
2.3	Magnitude-Recurrence Relationships and Frequency of Large Earthquakes	2-2
3	OUTLINE OF DATA SOURCES AND PROCEDURES	3-1
3.1	Data Sources	3-1
3.2	Statistical Analysis	3-1
	<i>Rates</i>	3-2
	<i>Magnitude Distribution Plots</i>	3-2
	<i>Low-Magnitude cut-off</i>	3-3
	<i>High-Magnitude cut-off</i>	3-3
4	REVISED MAGNITUDES IN THE NCEER-91 CATALOG -- DATA SOURCES	4-1

TABLE OF CONTENTS (continued)

<i>SECTION</i>	<i>TITLE</i>	<i>PAGE</i>
5	RATES OF SEISMICITY AND COMPLETENESS FOR DIFFERENT MAGNITUDES	5-1
6	MAGNITUDE DISTRIBUTION AND RECURRENCE TIMES	6-1
7	DISCUSSION	7-1
7.1	NCEER vs. EPRI: Data Comparison	7-1
7.2	Is the Magnitude Distribution Linear?	7-3
7.3	Magnitude Distributions in Cratonic Eastern U.S. (WEST) vs. Appalachians (EAST)	7-6
7.4	Effects on the Statistical Parameters (a- and b- values) from Differences in Curve-Fitting and other Procedures ..	7-8
	<i>Subjective Picks of Seismicity Rates</i>	7-8
	<i>Least Squares vs. Maximum Likelihood</i>	7-9
8	CONCLUSIONS	8-1
9	REFERENCES	9-1

LIST OF ILLUSTRATIONS

<i>FIGURE</i>	<i>TITLE</i>	<i>PAGE</i>
3-1	Map of epicenters from the NCEER-91 catalog.....	3-7
3-2	Map of epicenters from the EPRI-S catalog.....	3-8
3-3	Map of epicenters from the EPRI-V catalog.....	3-9
3-4	Map of epicenters showing magnitude differences.....	3-10
	NCEER-91 vs. EPRI-S.	
3-5	Sample histogram of seismicity for a specific magnitude window ΔM	3-11
3-6	Six sets (a-f) of histograms of seismicity within cumulative 0.25M windows for EAST and WEST from NCEER-91, EPRI-S and EPRI-V.....	3-12
3-7	Six sets (a-f) of histograms of seismicity within cumulative 0.20M windows for EAST and WEST from NCEER-91, EPRI-S and EPRI-V.....	3-19
3-8	Six sets (a-f) of histograms of seismicity within incremental 0.25M windows for EAST and WEST from NCEER-91, EPRI-S and EPRI-V.....	3-26
3-9	b-values and repeat-times of $M \geq 6$ for sets of low-magnitude cut- offs determined from least-squares fits of the log-normal distribution of magnitudes obtained from cumulative steps of 0.25 magnitude units, in each of two regions: EAST (A) and WEST (B). Weights according to cumulative number of events.....	3-33
3-10	b-values and repeat-times of $M \geq 6$ for a set of low-magnitude cut- offs determined from least-squares fits of the log-normal distribution of magnitudes obtained from cumulative steps of 0.25 magnitude units, in each of two regions: EAST (A) and WEST (B). Weights according to incremental number of events.....	3-36
3-11	b-values and repeat-times of $M \geq 6$ for a set of low-magnitude cut- offs determined from least-squares fits of the log-normal distribution of magnitudes obtained from cumulative steps of 0.25 magnitude units, in each of two regions: EAST (A) and WEST (B). Uniform weights.....	3-39

LIST OF ILLUSTRATIONS (continued)

<i>FIGURE</i>	<i>TITLE</i>	<i>PAGE</i>
3-12	b-values and repeat-times of $M \geq 6$ for a set of low-magnitude cut-offs determined from least-squares fits of the log-normal distribution of magnitudes obtained from cumulative steps of 0.20 magnitude units, in each of two regions: EAST (A) and WEST (B). Weights according to cumulative number of events.....	3-42
3-13	b-values and repeat-times of $M \geq 6$ for a set of low-magnitude cut-offs determined from maximum-likelihood fits of the log-normal distribution of magnitudes obtained from incremental steps of 0.25 magnitude units, in each of two regions: EAST (A) and WEST (B).....	3-45
3-14	b-values and repeat-times of $M \geq 6$ for a set of low-magnitude cut-offs determined from maximum-likelihood fits of the log-normal distribution of magnitudes obtained from incremental steps of 0.25 magnitude units. NCEER WEST; three different upper magnitude limits.....	3-48
3-15	b-values and repeat-times of $M \geq 6$ for a set of low-magnitude cut-offs determined from maximum-likelihood fits of the log-normal distribution of magnitudes obtained from incremental steps of 0.25 magnitude units. NCEER EAST; three different recurrence time for the 1886 $m=6.8$ earthquake.....	3-50

LIST OF TABLES

<i>TABLE</i>	<i>TITLE</i>	<i>PAGE</i>
3-I	Earthquake compilations in the Eastern U.S. used as sources for the NCEER-91 catalog.....	3-4
3-II	Comparison between EPRI and NCEER-91 in terms of magnitudes.....	3-5
3-III	Number of events in each decade with magnitudes determined from maximum intensity (Mmi), felt area (Mfa), and instrumental recordings (Mis) through the historic period (1800-1980), in NCEER-91(N) and EPRI(E).....	3-6

SECTION 1
INTRODUCTION

1.1 Earthquake Data for Definition of Seismic Sources -- Intraplate versus Interplate Environments

The available record of prior earthquakes is invariably the most important data base for characterizing earthquake sources and assessing hazard. Prior earthquakes are known from a variety of data, such as their effects on man-made and geologic structures, from felt reports, or from instrumental records. Data requirements for hazard assessment may vary according to the nature of the seismogenic environment. Along the well-studied plate boundary in California, the succession of known historic and prehistoric ruptures on major seismogenic faults provides the main basis for time-dependent probabilistic hazard assessments. Background seismicity in this area originates from several distinct classes of faults, such as major faults that can generate large earthquakes, secondary faults too small to generate significant earthquakes, and creeping faults that slip aseismically for the most part. Background seismicity has a complex spatial and temporal relationship to large earthquake ruptures and is not directly considered for long-term earthquake forecasting. It may, however play a crucial role in short- and intermediate-term earthquake prediction.

In contrast, relatively little is known about large earthquakes in the eastern U.S. Large historic events are few, occurred before instrumentation, and are believed not to have ruptured the surface. Paleoseismology in the Eastern U.S. is just beginning to produce some data on prehistoric earthquakes. Moreover, faults have not been associated to known earthquakes, except in very few cases, and, generally, a tectonic model accounting for intraplate seismicity is still lacking.

Thus, earthquake hazard in the eastern U.S. is not based on the probability of specific ruptures, rather it is derived from the level of seismicity within poorly defined source regions where the seismicity is assumed to be homogeneous. This seismicity 'level' contains information about the recurrence time and the magnitude distribution of earthquakes. Thus, the source component of earthquake hazard determinations in the eastern U.S. is based almost exclusively on catalogs in the most basic form, a compilation of earthquakes characterized by origin times, locations and magnitudes. Where, how often,

and how big damaging earthquakes are expected in the eastern U.S. is determined almost exclusively on the basis of available earthquake catalogs.

Geologic data play a relatively minor role in determining earthquake hazard in the eastern U.S. They help to delineate earthquake sources and are beginning to yield information about prehistoric earthquakes. Geology will increasingly contribute to the assessment of earthquake hazard as the understanding of tectonic processes responsible for intraplate seismogenesis improves. Improved understanding, however, requires reliable earthquake data that can be compared, for example, to structural features and other geologic parameters. The limiting factor for most applications of the catalog in both applied and fundamental seismology are the reliability of earthquake parameters and the uniformity and completeness of the seismicity coverage.

The quality of the coverage of the seismicity varies with time, location, and magnitude. In particular, the type of data available to determine earthquake parameters changes over the historic period. Intensity data provide most of the information on background seismicity before the mid 20th century. Reliable instrumental data become available for most felt earthquakes only in the 1970's when regional seismic networks were installed. As instrumentation brings the threshold of detection below felt reports, many small events are detected. In California most of them are earthquakes; in the eastern U.S., some of these small events are earthquakes, but most of them are blasts or other artificial seismic sources. The differentiation of small earthquakes from other sources on the basis of seismic signatures is not always trivial and the increased sensitivity provided by regional networks introduces a new source of noise in the catalog. Generally, earthquake catalogs provide the main basis for site-specific earthquake hazard estimates in the eastern U.S. These catalogs are subject to a variety of limitations, some of which can be at least partially removed. Given their fundamental role in the parametrization of hazard and in other applications, considerable investments in the improvement of these catalogs are warranted.

1.2 Improved Constraints on Magnitudes and Their Effects on Recurrence Relations

Recent efforts to improve constraints on historical seismicity in the eastern U.S. (e.g., Bollinger and Hopper, 1972; Seeber and Armbruster, 1987; Reimbold and Johnston, 1987) have shown: 1) that the current catalog is uneven in coverage, reflecting its origin as a patchwork of different compilations; and 2) that substantial improvements can be

achieved by a systematic re-examination of available data. Following this lead, we are re-examining earthquake data by searching archival material and cross checking the original earthquake compilations. Such an effort is very labor intensive, but also rewarding because significant changes are brought about in both locations and magnitudes. In particular, systematic changes in the magnitude distribution caused by the revisions outweigh some of the effects brought about by refinements in statistical analysis of earthquake data which have been recurrently debated (e.g., Algermissen, et al , 1990 (US Geological Survey); Veneziano and Van Dyck, 1985 (Electric Power Research Institute); Bernreuter, et al.; 1984 (Lawrence Livermore National Laboratory)).

The main purpose of this report is to discuss results from a revision of the earthquake catalog for the eastern U.S. Only the effects of the revision on magnitudes in the area east of the New Madrid seismic zone are considered. The effects of the revision on earthquake locations will be discussed in a subsequent report. More changes are expected in the future from this ongoing re-examination. But the need for updated assessments of earthquake hazard as new data become gradually available through the time-consuming search of historic sources warrants presenting interim results from the ongoing re-examination. One of the purposes of presenting these results is to identify crucial factors affecting the parameters feeding into hazard assessments and to guide future efforts to improve the catalog.

1.3 Alternative Analytical Procedures and Source Zones: Effects on Recurrence Parameters

Parameters quantifying the seismicity for earthquake hazard estimates can be produced by several alternative procedures. Each procedure gives confidence limits, but how do these mathematically derived confidence limits compare to differences in the results obtained from alternative procedures? One of the main themes in this paper is to compare results from several alternative procedures. These comparisons show that the statistical confidence limits are often smaller than the range of results obtained from different procedures and can, therefore, underestimate the uncertainties.

Ideally, alternative data sets and procedures for the same seismogenic region should yield similar results. In contrast, data from distinct source zones may conceivably differ in their statistical characteristics. In order to test this hypothesis, seismicity from the eastern and the western sides of our study area were considered separately. As the boundary between

them, we chose the Appalachian crystalline thrust front, which separates the continental platform and shield area of North America from the Appalachian province characterized by an allochthonous crystalline slab emplaced by Paleozoic collisional orogenies. The crystalline front is also thought to separate two spatially distinct neotectonic source zones (Seeber and Armbruster, 1978). The qualifiers EAST and WEST are combined with the designations for the three catalogs considered in this paper to identify data from each of these zones. Unfortunately, differences in the magnitude distributions between EAST and WEST are not resolved. This negative result is not surprising, however, because the resolution of the method is poor, as indicated by the large procedure-dependent variations in the statistical parameters.

SECTION 2
**SOURCES OF BIAS IN DATA RELEVANT TO EARTHQUAKE
STATISTICS AND HAZARD ESTIMATES**

2.1 Maximum-Intensity versus Felt-Area Magnitudes

Earthquake magnitude was defined on the basis of instrumental measurements (Richter, 1935). A number of different instrumental magnitudes are used depending on earthquake size, distance, and recording instrument. For the purpose of a statistical evaluation of magnitude distribution, the largest of the instrumental magnitudes assigned to an event is chosen in this paper and simply referred to as instrumental magnitude (*M_{is}*). In cases where only intensity data are available or where instrumental determinations are not sufficiently reliable, magnitudes can be derived from either maximum intensity (*M_{mi}*), from felt-area (*M_{fa}*), or from more sophisticated parametrizations of the intensity distribution (also referred to as *M_{fa}* in this paper), using well established relationships (e.g., Sibol et al., 1987). Maximum intensity can usually be assigned from available reports and it is often listed in many of the classical earthquake compilations (e.g., Taber, 1914, 1915; Rockwood, 1872-1886). More detailed information and more laborious analysis are required for an assessment of felt-area. This effort, however, is usually worthwhile because felt-area is a reliable and robust measure of magnitude; even coarsely determined felt areas can yield useful constraints on the magnitudes (a factor of two in felt-area corresponds to about 0.2 of a magnitude). In contrast, maximum intensity tends to vary substantially for a set of earthquakes with the same magnitude (e.g., Sibol et al., 1987) and is a poor measure of magnitude.

The wide scatter in the relation between maximum intensity and magnitude depends on several factors, some observational in nature and related to inconsistencies in assigning maximum intensity, others phenomenological in nature and related to the non-uniqueness of the relation between magnitude and mesoseismal effects. Hypocentral depth, source mechanism, source spectrum and site response are some of the phenomenological factors. An important procedural factor is the criterion for assigning maximum intensity, which can be the maximum reported intensity or the mode of the intensity in the mesoseismal area. Population coverage relative to the size of the mesoseismal area is an important cultural factor that can introduce systematic bias. The maximum reported intensity is more likely to be less than the maximum 'potential' intensity for a sharply peaked intensity distribution than for a broad one. Furthermore, the intensity scale is coarse and discrete; magnitudes

derived from maximum intensity will cluster at the set of magnitudes corresponding to these discrete intensity levels and distort magnitude distribution plots. Sibol et al. (1987) obviate this problem by retaining Mmi as the magnitude distribution observed in the relation between magnitude and maximum intensity.

2.2 Systematic Bias in Magnitudes Derived from Intensity Data

Mmi's not only introduce large random errors in the catalog, but they probably also cause systematic errors leading to overestimated magnitudes. Mfa's are more credible than Mmi's and are generally adopted whenever available. Most catalogs contain a mixture of both Mmi and Mfa, but the distribution of these magnitudes is not random over the magnitude range. The larger the earthquakes, the more likely they are to have sufficient macroseismic information for a felt-area determination. Thus, of all the events associated with a given maximum intensity, the larger ones are more likely to be assigned an Mfa and are depleted from that category in the catalog; the events characterized only by an Mmi left in that maximum intensity category will be biased toward small size. This bias will cause Mmi for these events to be overestimated.

Conversely, the earthquakes chosen for a regression between Mmi and instrumental or felt-area magnitudes will be the ones for which reliable magnitudes are determined; they will probably tend to be the largest ones in each class of Mmi, particularly in the low range of maximum intensities (e.g., Sibol et al. 1987). Thus, the magnitude distribution obtained from the regression is expected to reflect a bias toward high magnitudes. Such a bias will further contribute to the tendency for Mmi to be over-estimated.

2.3 Magnitude-Recurrence Relationships and Frequency of Large Earthquakes

Earthquake hazard estimates depend on the inferred frequency of large destructive earthquakes. In the eastern U.S., data on large prehistoric earthquakes are still scarce and strain-based tectonic models for explaining seismogenesis are not available. Consequently, recurrence times for large events are inferred primarily from the space-time-magnitude distribution of historic seismicity. Inferences on the time-space behavior of large earthquakes based on the seismicity at smaller magnitudes require a number of fairly

arbitrary assumptions and tend to be characterized by large uncertainties stemming from poor constraints on historic seismicity.

A common assumption particularly vulnerable to criticism is that seismicity is stationary, i.e., that characteristics of the seismicity are time invariant. Stationarity prescribes that the distribution of seismicity (excluding foreshocks and aftershocks) in space and magnitude is statistically stable over a time period of interest for earthquake hazard. Thus, patterns of seismicity would evolve with time constants which are much longer than the historic period; during this period, changes would be insignificant and projections would be reliable, at least over a time sufficiently long to satisfy engineering concerns. Arguments against the validity of this assumption have been raised on the basis of historical data showing substantial changes in patterns of seismicity (e.g., Seeber and Armbruster, 1987) as well as data on paleoseismicity showing a discrepancy between repeat times of large events and geologic data (e.g., Coppersmith, 1989; Leffler and Wesnowsky, 1991).

Although stationarity is implicitly assumed in the magnitude-recurrence analysis carried out in this paper, we do not wish to imply an acceptance of this assumption. In fact, the improved constraints on the catalog may offer new tests for stationarity, but this issue is beyond the scope of this paper. The main purpose of the magnitude-recurrence analysis is to compare results with previous statistical studies of the EPRI catalog and to evaluate the changes reflected in the NCEER catalog in terms of earthquake recurrence. Generally, the results show that improved data change some of the hazard factors significantly. Another central theme is the comparison of uncertainties derived from the data with variations in results derived from alternative modeling techniques and with statistical fluctuations. While the emphasis is often on statistical uncertainties, our results show that uncertainties from the basic data and from the procedures to model these data can outweigh statistical uncertainties.

Another basic assumption in the statistical analysis of the catalogs is that the logarithm of the number of earthquakes N of a given magnitude M is inversely proportional to that magnitude, i.e., that the Gutenberg and Richter (1954) relationship $\text{Log}N=a-bM$ holds with constant a and b over the entire magnitude range. This relationship has fundamental implications concerning self-similarity in the size-distribution of earthquake sources. On the practical side, the assumption of linearity may be justified even if the data appear to be linear only because they represent the superposition of many source zones with distinct and possibly non-linear size distributions. The assumption of linearity dictates the choice of

procedures to fit the data and to extrapolate these data to large poorly sampled magnitudes which are significant for hazard.

We do not see any evidence in the data presented here of significant deviations from a linear magnitude distribution, but the resolution is poor. Locally, deviations may be expected on the basis of structural evidence discussed below. According to this evidence, the distribution may be clustered (i.e., non-linear) at all magnitudes, but the earthquake population within each cluster may be independent of magnitude and insufficient for resolution. For a given data set, the earthquake distribution may appear linear in the range of magnitude where several clusters are superimposed, but deviate from linearity at large magnitudes where only one cluster is sampled. In this case, recurrences of large and rare earthquakes estimated on the basis of a linear magnitude distribution may be wrong.

SECTION 3
OUTLINE OF DATA SOURCES AND PROCEDURES

This work sets out to improve the earthquake catalog in terms of magnitude determinations for the eastern U.S. (east of the New Madrid seismic zone). It also analyzes the data in a number of standard statistical procedures: a) to test the effect of the changes on parameters directly relevant to hazard estimates; and b) to compare the results obtained by alternative procedures. The results are given in a new catalog named NCEER-91. Figure 3-1 shows epicenters from this catalog.

3.1 Data Sources

All entries in the EPRI (1987) earthquake catalog were compared to other data sources listed in Table 3-I. The main improvement on the EPRI catalog are many additional felt areas as well as new earthquakes. As a result, the size of most of the earthquakes with $M \geq 3$ within the periods of completeness can be computed either with felt-area or with instrumental data. Felt-area magnitudes are computed from the formulas by Sibol et al. (1987). Figures 3-2 and 3-3 show epicenters for two versions of EPRI's catalog with different schemes to calculate magnitudes; Tables 3-II and 3-III compare EPRI's catalog with NCEER-91; Figure 3-4 compares these catalogs and shows differences at the appropriate epicenters.

3.2 Statistical Analysis

The main purpose of this analysis is to obtain reliable parameters of the level of seismicity and to assess the dependence of the results on the choice of basic data and on the analytical procedure. For this purpose results obtained with different data and different procedures are compared. Fundamental assumptions in the statistical analysis are: 1.) the seismicity is stationary, i.e., the spatial distribution of earthquakes does not change in time except for statistical fluctuations; and 2.) the size-distribution of earthquakes is linear, i.e., it has the form $\text{Log}N = a - bM$, where a and b are constants (Gutenberg and Richter, 1954). The following analyses are carried out in order.

Rates

- A. Seismicity is classified in a set of equal size steps in magnitude, either 0.20 or 0.25 magnitude units. This classification can be either 1.) incremental, where the magnitudes of the events in each magnitude window are within the limits of that window; or 2.) cumulative, where all magnitudes above the lower limit of the window are included.
- B. Magnitude-time plots are constructed for each magnitude window. These are log-log plots of rate (events /year) vs. backward time. These plots are designed to display as much of the information leading to the next step as possible (e.g., Figure 3-5).
- C. A rate and a completeness period are assigned from the magnitude-time plots to each magnitude window.

Three sets of procedures are carried out for each of six data sets. The three procedures consist of incremental grouping at steps of 0.25M and 0.20M plus cumulative grouping at steps of 0.25M. Six data sets result from three earthquake listings, each separated geographically into the eastern and the western parts of the study area. The three listings are our new catalog NCEER-91 and two versions of the EPRI catalog, EPRI-S, where maximum-intensity magnitudes M_{mi} are assigned according to Sibol et al. (1987) and EPRI-V where M_{mi} are assigned according to Veneziano and Van Dyck (1985). 18 sets of magnitude-dependent rates are produced at this stage of the procedure (Figures 3-6, 3-7, and 3-8).

Magnitude Distribution Plots

The 18 sets of magnitude-dependent rates are fit to Gutenberg-Richter's law, $\text{Log}N=a-bM$, to recover best-fitting values of the parameters a and b . This fitting is accomplished by the maximum likelihood and by the least-squares methods. In the least-squares method the rates for each magnitude window are weighted by three different schemes: 1.) weight proportional to the cumulative number of events in that magnitude window; 2.) to the incremental number of events; and 3.) all rates have the same weight. The total at this stage adds up to 30 fitting procedures: three catalogs, each split into two regions, yield six data sets. Magnitude distributions are plotted according to five different schemes for each data set: magnitude windows can be cumulative with 0.25M[a] and 0.20M[b] window size plus incremental with 0.25M[c] window size; [a] is fit by the least-squares method and weighting schemes 1., 2., and 3.(Figures 3-9, 3-10, and 3-11); [b] is fit by the least-

squares method and weighting scheme 1.(Figure 3-12); and [c] is fit by the maximum likelihood method (Figure 3-13).

Low-Magnitude cut-off

Each of the 30 fitting procedures is carried out for a set of low-magnitude cut-offs, from $M=3$ to $M=5$, to test the stability of the statistical parameters over the magnitude range. The results are presented as plots of b-values and recurrence times for a $M=6$ earthquake (rather than the a-value) as functions of low-magnitude cut offs (Figures 3-9 to 3-13).

High-Magnitude cut off

Finally, NCEER-91 incremental rates for the western part of the study area are fit by maximum-likelihood with three alternative upper magnitude limits ($M=6, 7,$ and 8) and relative temporal constraints (Figure 3-14). Furthermore, NCEER-91 cumulative (magnitude steps $0.25M$ and $0.20M$) and incremental (magnitude steps $0.25M$) rates for the eastern part of the study area are fit by least-squares and maximum-likelihood methods, respectively, with three alternative recurrence times for the $M=6.8$ (Charleston 1886-like) earthquake (Figure 3-15).

TABLE 3-I Earthquake Compilations in the Eastern U.S. used as sources for the NCEER-91 catalog. The numbers are the felt areas from each source which are either different from, or new relative to EPRI.

Data Source	Felt Areas
Armbruster, unpublished data	51
Armbruster and Seeber, 1987	18
Barstow et al (Roundout/NRC), 1981	0
Hopper and Bollinger, 1971; Bollinger and Hopper, 1972	45
Boston-Edison (Weston Geophysical), 1976	0
Dewey and Gordon, 1984	(no felt area info)
EPRI (Electric Power Research Institute), 1987	--
Reinbolt and Johnston, 1987	46
NOAA Earthquake Effect File, through 1981	102
Rockwood, 1871-1886	12
Stover (USGS), 1982	(no felt area info)
US Earthquakes, 1928-1983	60
Visvanathan (South Carolina Geol. Surv.), 1980	6

TABLE 3-II Comparison between EPRI and NCEER-91 in terms of magnitudes

	NCEER-91	EPRI
Events	1554	1508
no a/s	1206	1214
N. of Mis	392	391
no a/s	342	355
N. of Mfa	405	193
no a/s	288	166
N. of Mmi	747	924
no a/s	576	693
Magnitudes changed	399	
Magnitudes unchanged	1073	
Added events (M>3 and new)	82	
Eliminated events (M<3 and false)	36	

a/s=aftershocks; N.=number; Mis=instrumental magnitudes; Mfa=felt-area magnitudes; Mmi=maximum-intensity magnitudes

		1800	10	20	30	40	50	60	70	80	90	1900	10	20	30	40	50	60	70	80	
Mmi	VII													0	0	0					
														3	1	1					
	VI			1			0		0	0			1	2	0	0		0			
				1			1		2	11			3	3	4	1		4			
	V	4		1	2	2	0	1	3	0	1		4	5	5	2	1	3	1	0	0
	4		1	2	3	3	2	8	12	3		7	8	9	8	6	8	6	2	1	
IV	N	9	5	6	3	4	21	5	16	20	59	14	24	38	20	17	21	11	14		
	E	9	5	7	3	4	21	5	21	60	59	16	29	47	38	28	38	18	17		
III	N	3	5	2	1	18	23	8	14	21	6	10	10	26	28	16	4	1	3	1	
	E	3	5	2	1	18	23	8	15	23	6	10	10	28	45	20	5	2	3	1	
Mfa	N	1	5	6	3	11	23	8	31	109	17	22	39	47	70	36	65	51	19	1	
	E		5	4	2	7	13	4	15	18	15	15	27	25	16	2	9	3	1	0	
Mis	N													1	3	28	8	34	123	90	
	E													2	14	42	37	70	135	90	
		1800	10	20	30	40	50	60	70	80	90	1900	10	20	30	40	50	60	70	80	

TIME ----->

TABLE 3-III Number of events in each decade with magnitudes determined from maximum intensity (Mmi), felt area (Mfa), and instrumental recordings (Mis) through the historic period (1800-1980), in NCEER-91(N) and in EPRI (E). The Mmi are grouped according to maximum intensity (MI). Note that Mis's are less in N than in E primarily because felt-area magnitudes are preferred over early instrumental magnitudes.

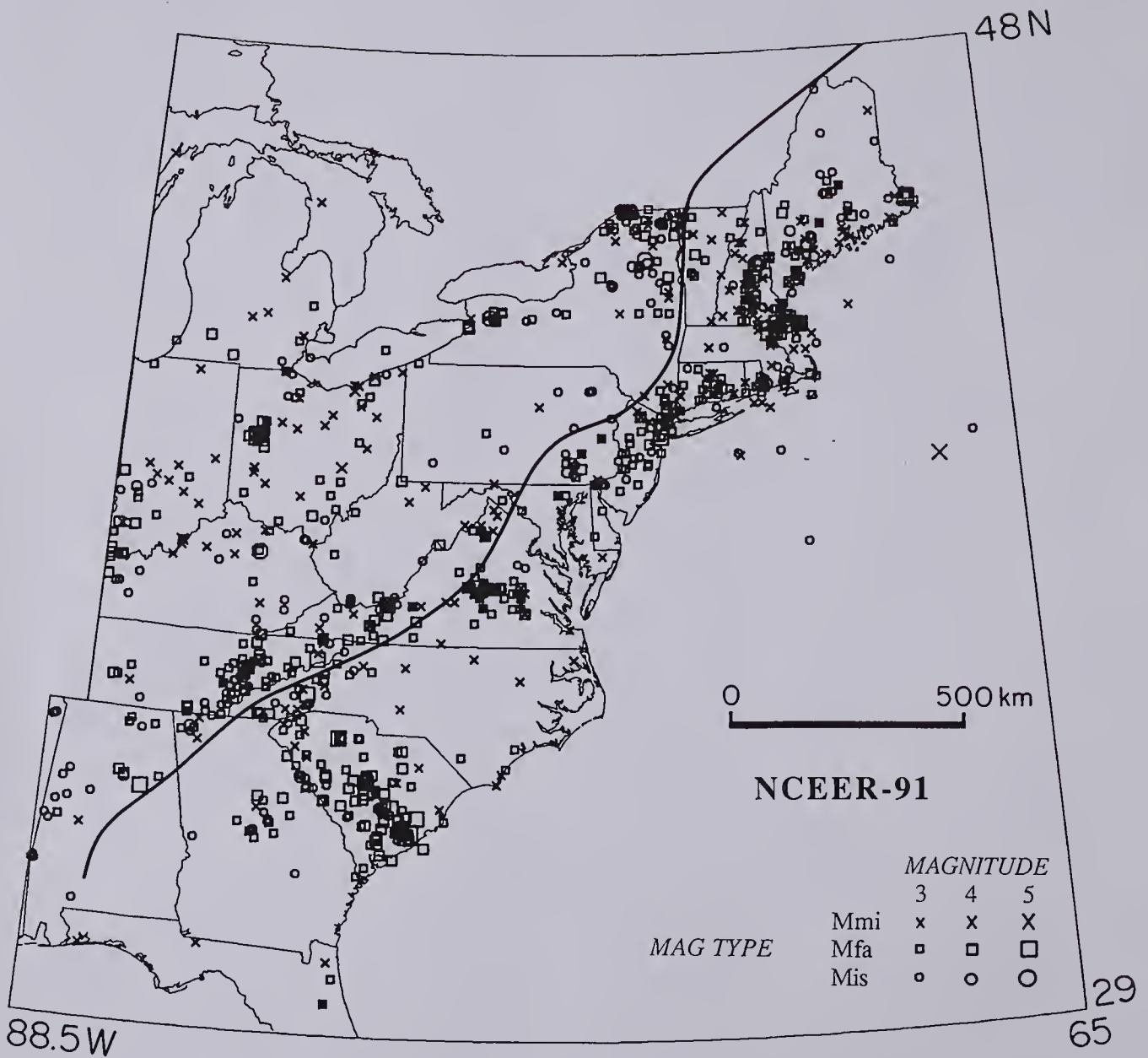


FIGURE 3-1 Seismicity map of the eastern half of the eastern U.S. (east of the New Madrid seismic zone). Epicenters and magnitudes ($M \geq 3$) from the NCEER catalog (NCEER-91; see text). Three kinds of magnitudes, Mis (instrumental), Mfa (felt-area), and Mmi (maximum intensity), are distinguished by symbols. The emphasis in NCEER-91 is on a revised set of magnitudes; most of the unreliable Mmi during the period of completeness have been upgraded to Mfa.

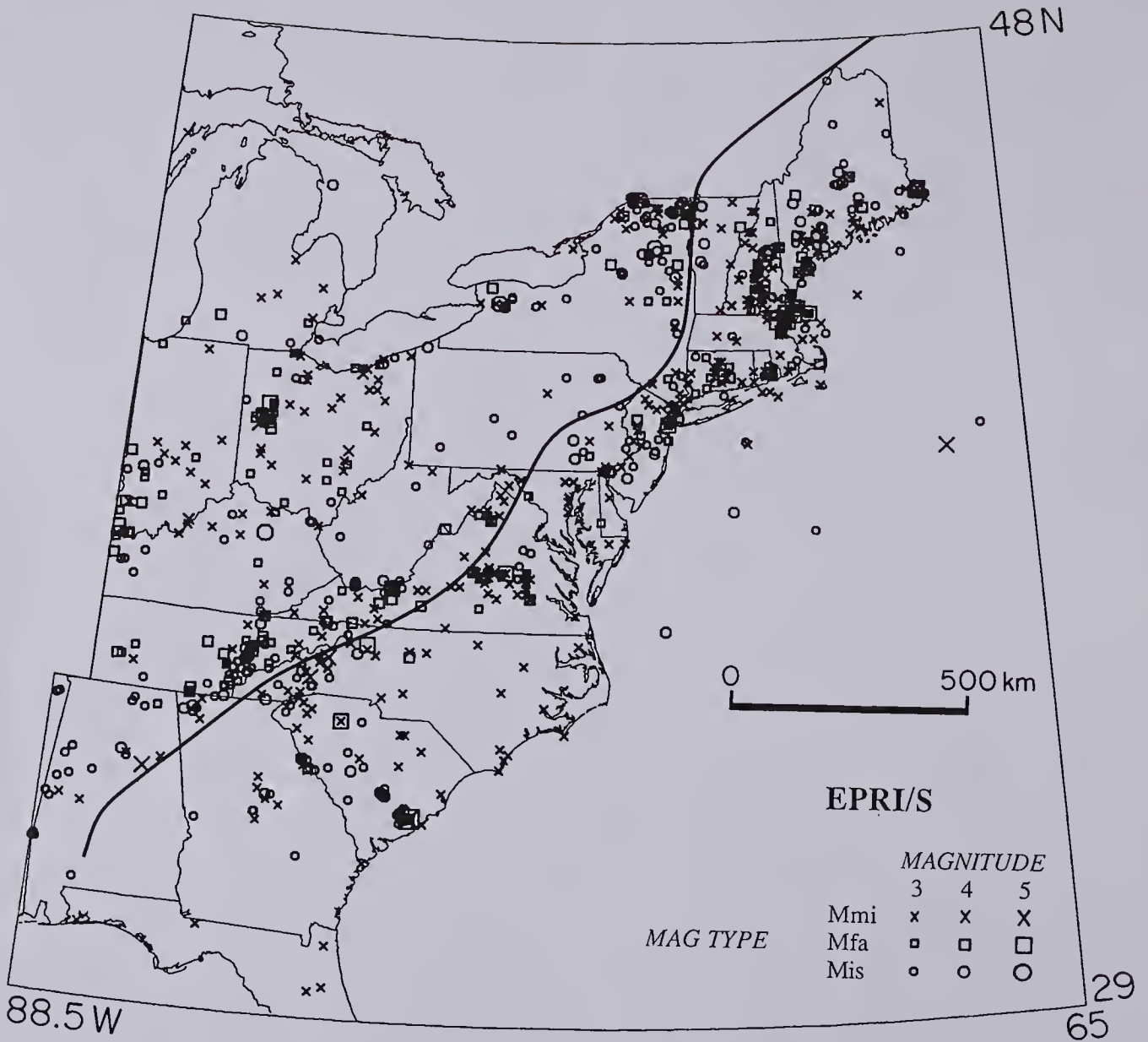


FIGURE 3-2 Same map as in Figure 3-1, with data from the EPRI/S catalog (see text; magnitude categories: Mis=instrumental; Mfa=felt area; Mmi=maximum intensity).

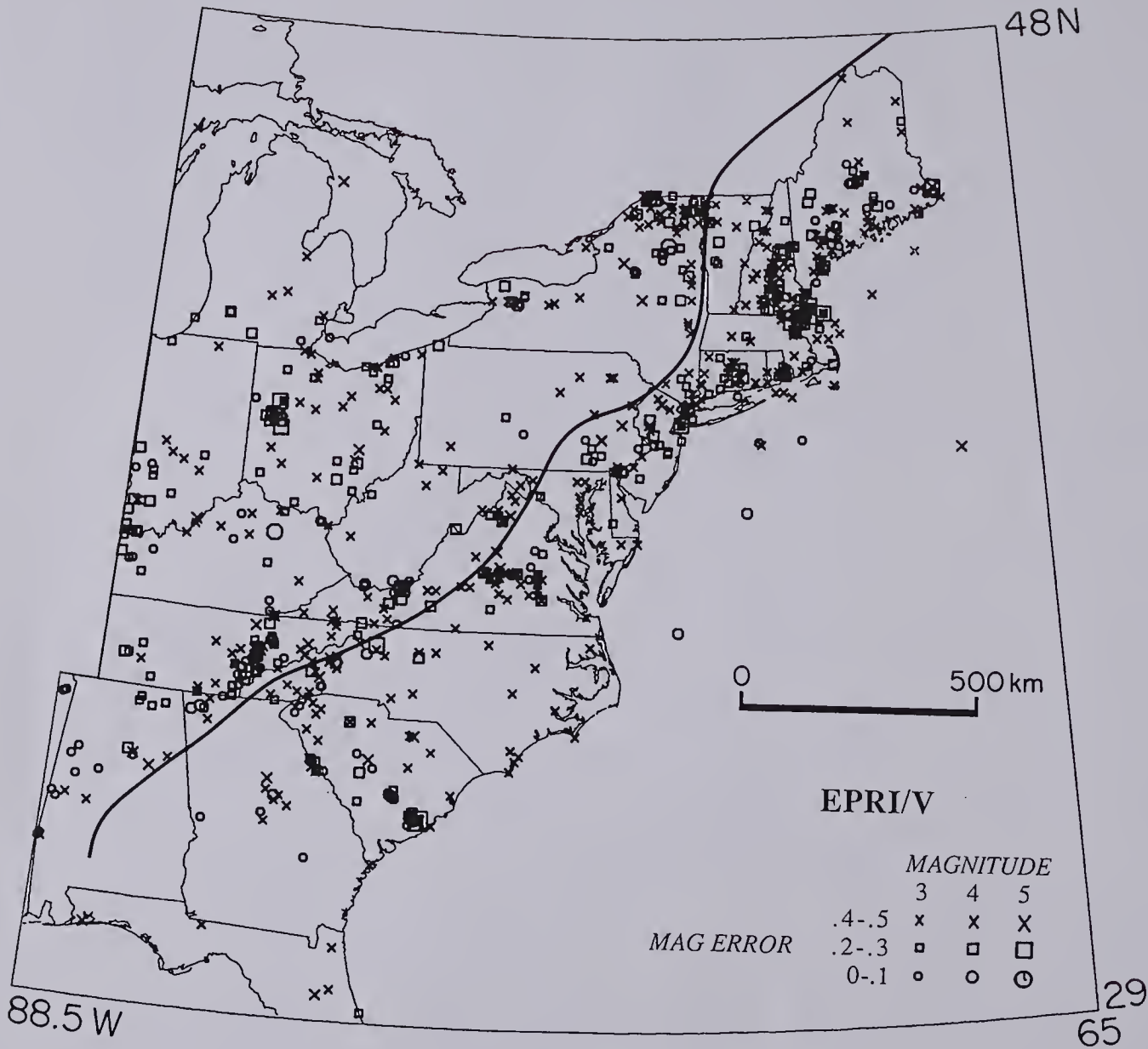


FIGURE 3-3 Same map as in Figure 3-1, with data from the EPRI/V catalog (see text). In this catalog magnitudes are not differentiated according to type, but they are assigned error estimates. The symbols indicating these error estimates have been calibrated to approximately correspond with the categories in Figures 3-1 and 3-2.

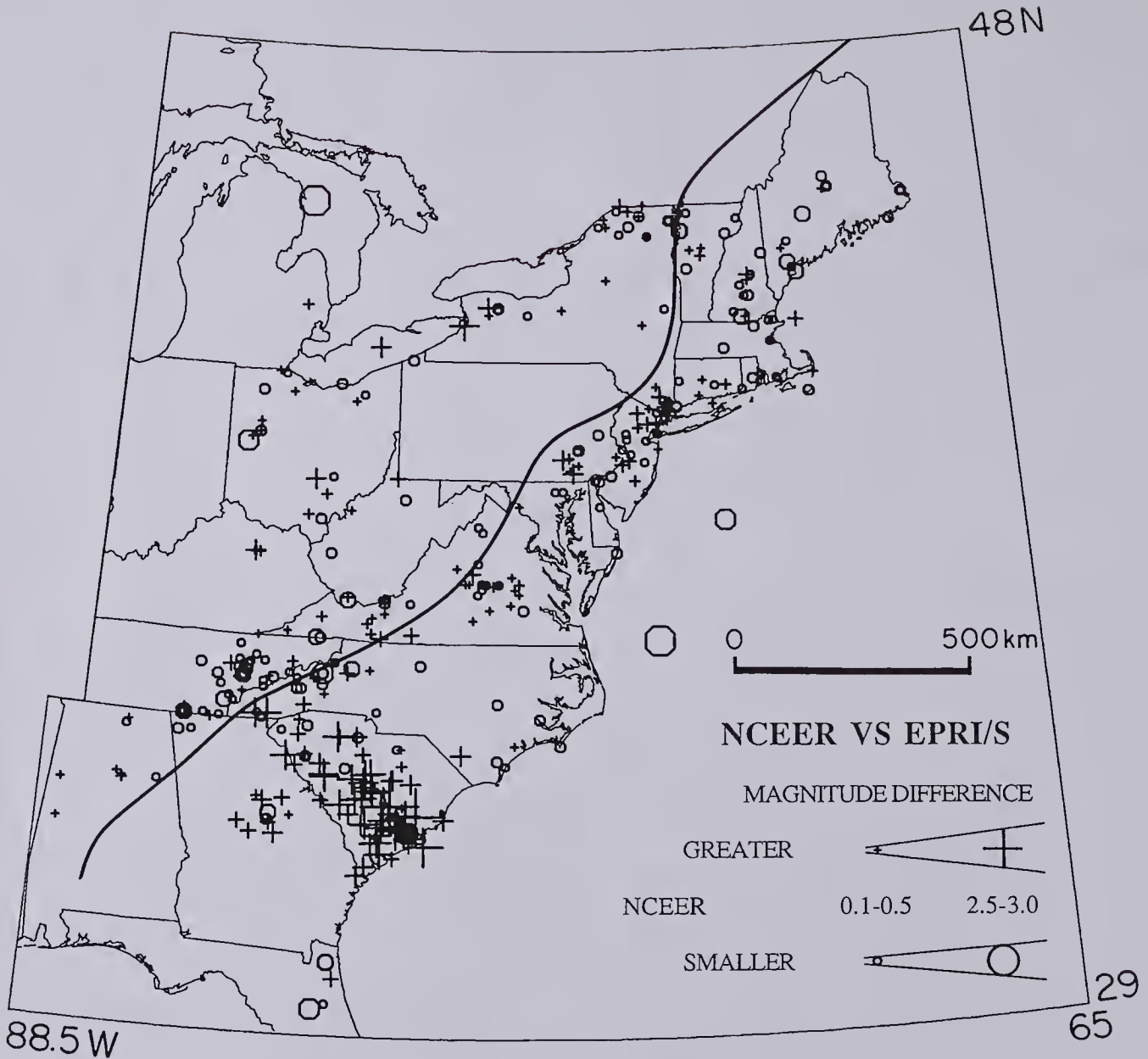
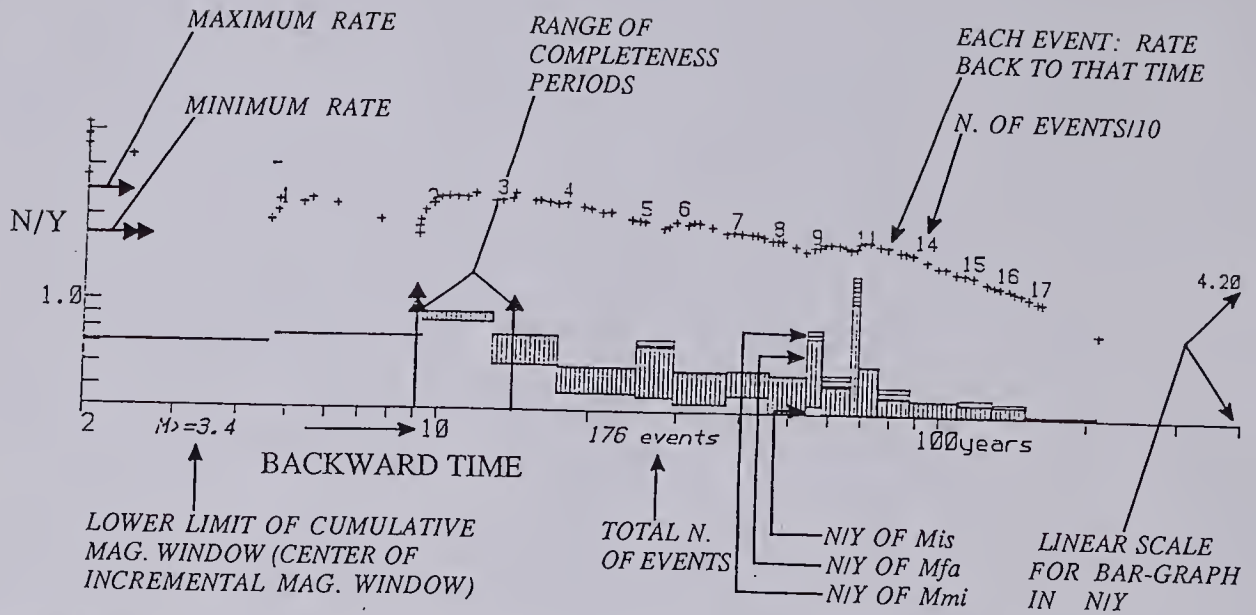


FIGURE 3-4 Differences in magnitudes between the NCEER-91 and the EPRI/S catalogs. Epicenter locations of some of the events common to both NCEER and EPRI may also differ in the two catalogs. Epicenters in this Figure are from NCEER (except for the events only in EPRI).



MAGNITUDES: Mis = INSTRUMENTAL; Mfa = FELT-AREA; Mmi = MAXIMUM INTENSITY
 N. = NUMBER ; N/Y = RATE (NUMBER PER YEAR)

FIGURE 3-5 Sample of log-log magnitude-time plot describing the temporal behavior of the seismicity in a given magnitude range. This labeled sample-plot for a single magnitude window serves as a legend for Figures 3-6, 3-7, and 3-8. The lower limit of the magnitude window is given below-left. Each plus represents an earthquake, but some events are omitted to avoid clutter in crowded portions of the plots. The ordinate of these points refer to the logarithmic scale on the left and gives the average rate of seismicity from the time of that event up (left) to the end of the catalog. In this case the rate is for a cumulative magnitude range, i.e., it includes all events with magnitudes above the lower limit for this plot. The numbers adjacent to some of the crosses give the number of events in tens back to that point. The bar-graphs also represent the average rate of seismicity, but only on the right of each plot. The bar-graphs refer to the linear scale represented by the number on the apparent rate of seismicity. These bar-graphs also indicate the type of magnitude: Mmi=horizontal shading; Mfa=vertical shading; Mis=no shading. The choice of each period in the bar-graphs is attempting to maximize information on apparent changes in the rate of seismicity. Each period terminates at an event, it always terminates at multiples of ten events and after a fixed linear distance in the abscissa. This distance is fixed to the logarithmic time scale. These plots provide the basis for a subjective estimate of the average rate of seismicity (horizontal arrow) and the period of completeness or the number of events back to the completeness limit (vertical arrow). In this sample, rate and completeness are given as a range of possible values. The completeness limit also reflects an independent judgement based on the history of seismology (see text).

FIGURE 3-6 Temporal distribution of seismicity within cumulative magnitude windows in steps of 0.25 magnitude units above $M=3$. See Figure 3-5 for a description of these plots. Note that the column of plots on the right has an abscissa scale expanded by a factor of two to display detail available at low magnitude. Six sets of plots result from three data sets and two areas. For each of NCEER, EPRI/S, and EPRI/V, two areas are considered: east of the Appalachian front (EAST); and west of the Appalachian front (WEST). For comparison sake, the same scale is adopted in corresponding plots for the three catalogs.

SEISMICITY RATE - EVENTS PER YEAR

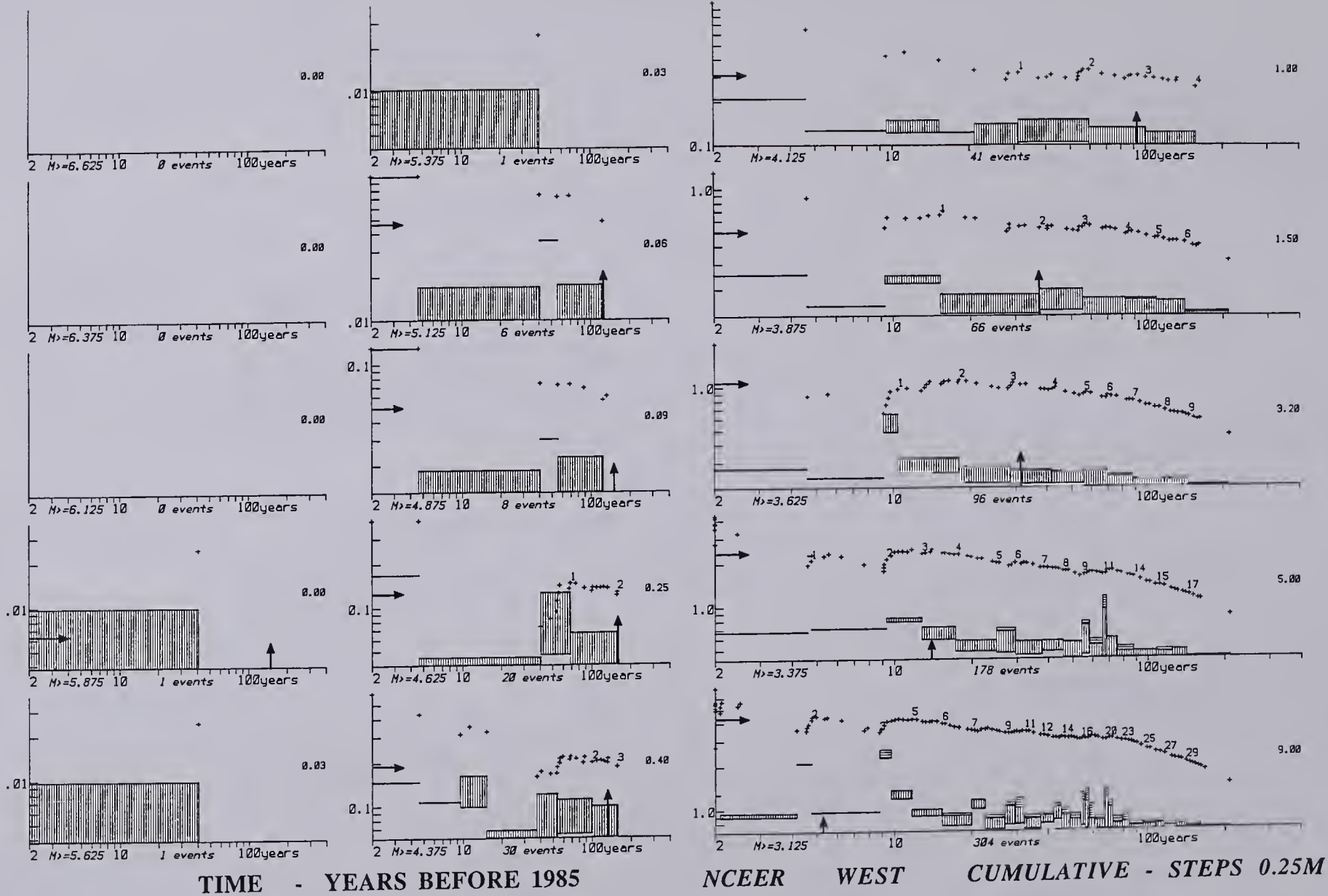
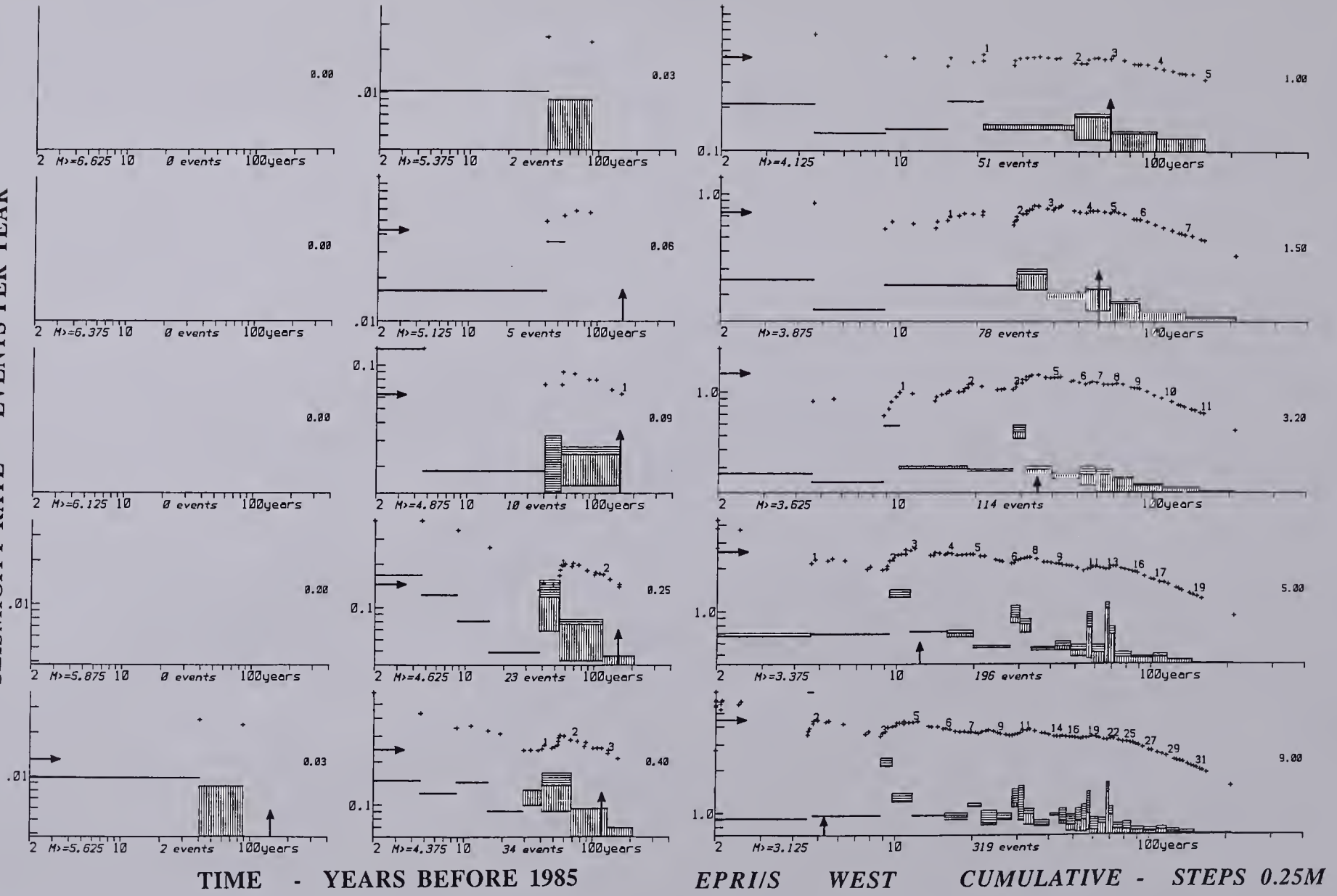


FIGURE 3-6a

SEISMICITY RATE - EVENTS PER YEAR

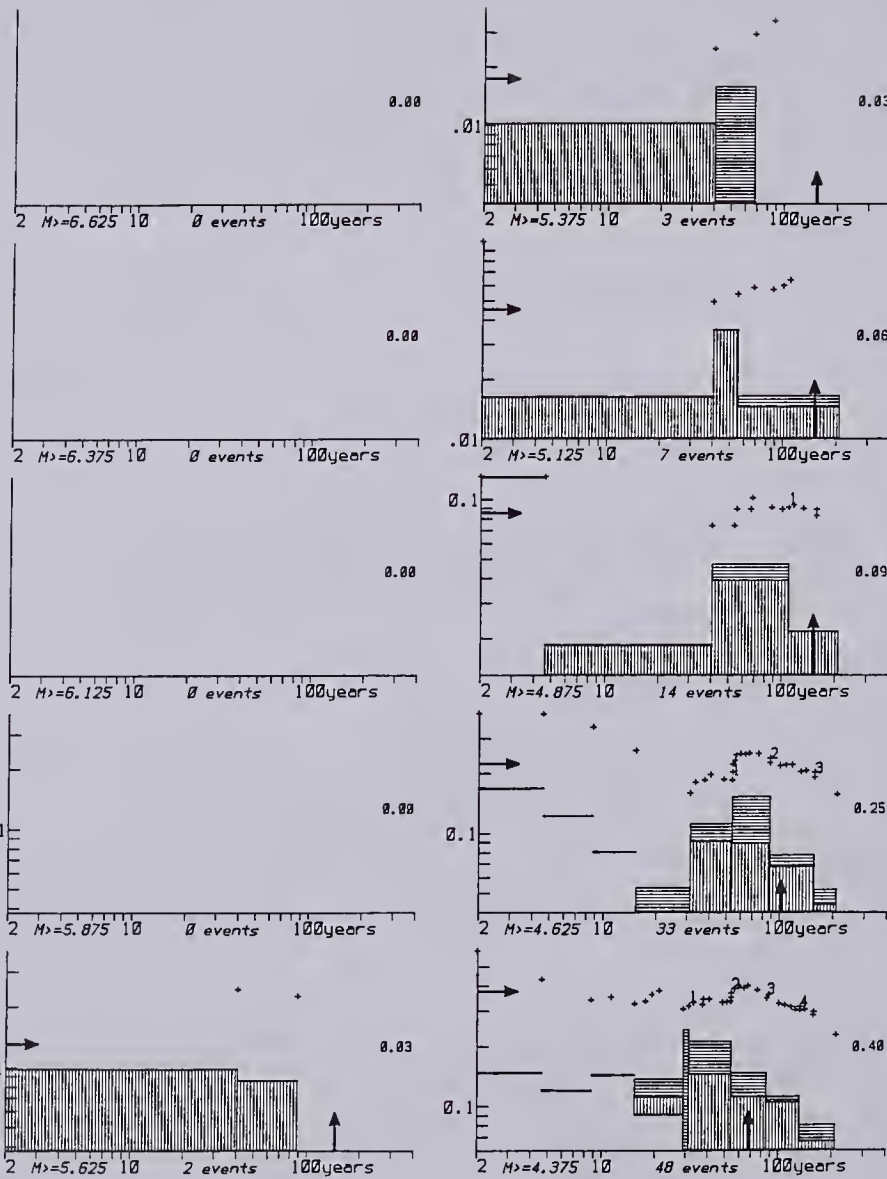


TIME - YEARS BEFORE 1985

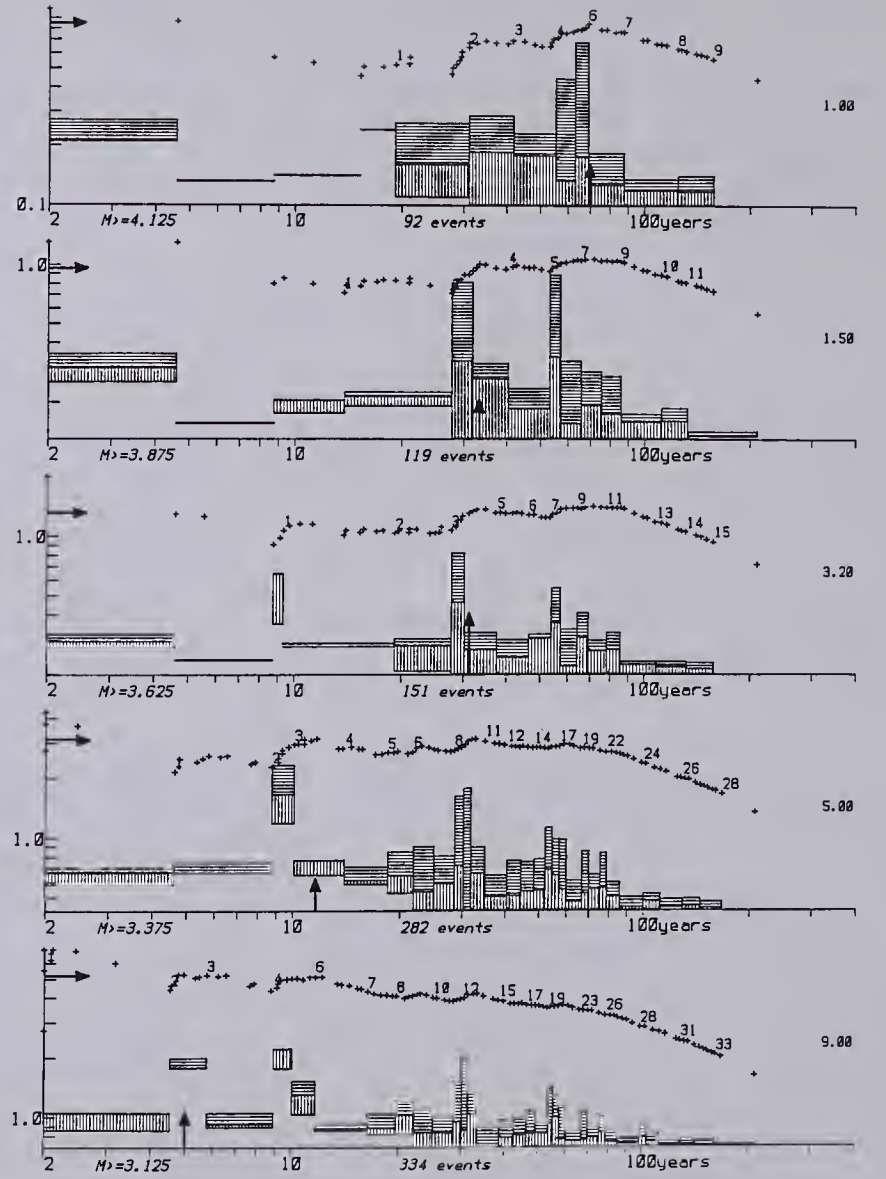
EPRI/S WEST CUMULATIVE - STEPS 0.25M

FIGURE 3-6b

SEISMICITY RATE - EVENTS PER YEAR



TIME - YEARS BEFORE 1985



EPRI/V WEST CUMULATIVE - STEPS 0.25M

FIGURE 3-6c

SEISMICITY RATE - EVENTS PER YEAR

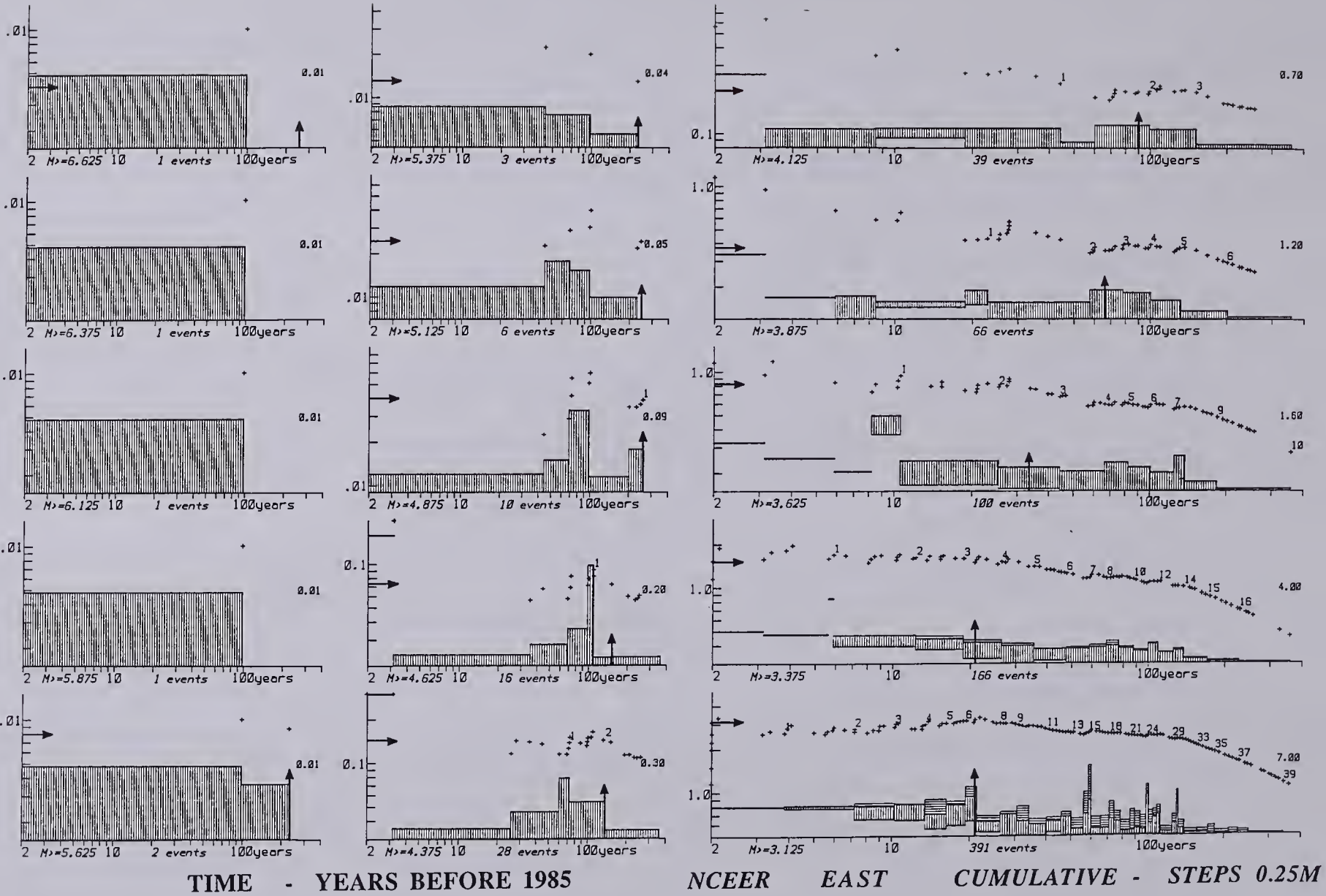
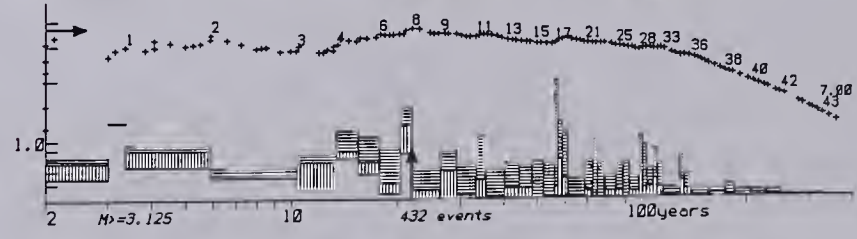
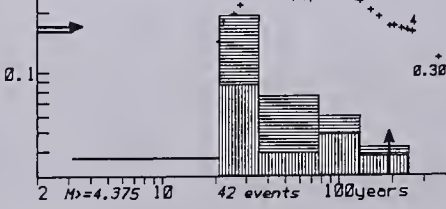
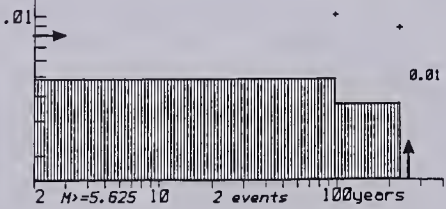
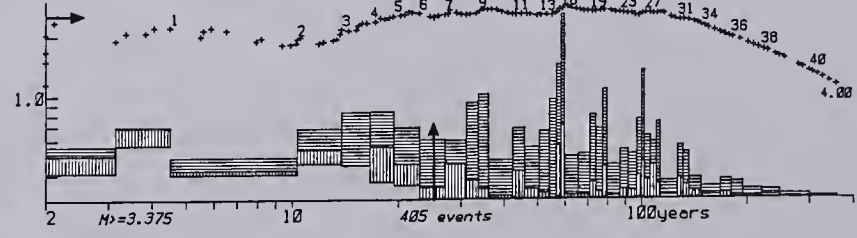
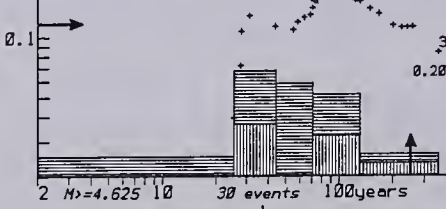
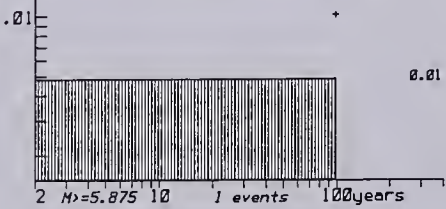
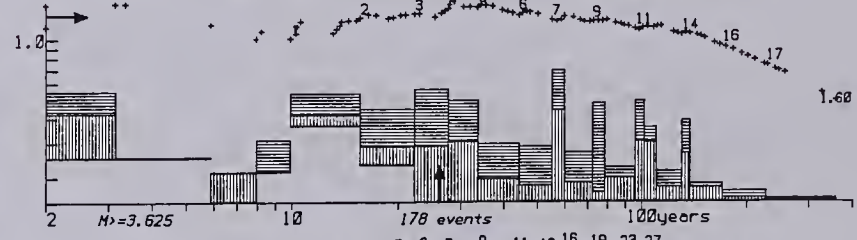
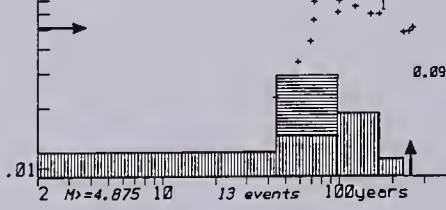
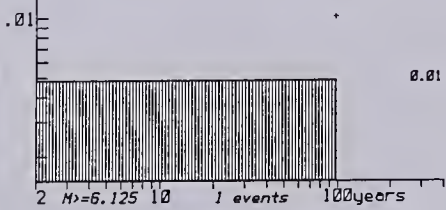
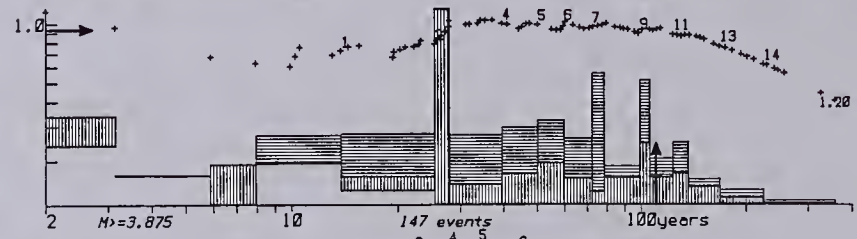
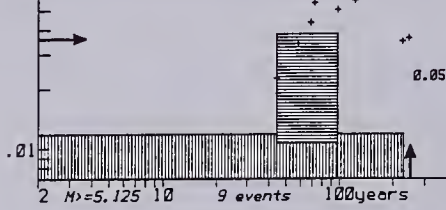
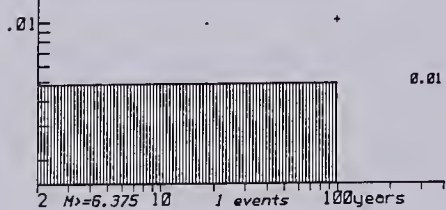
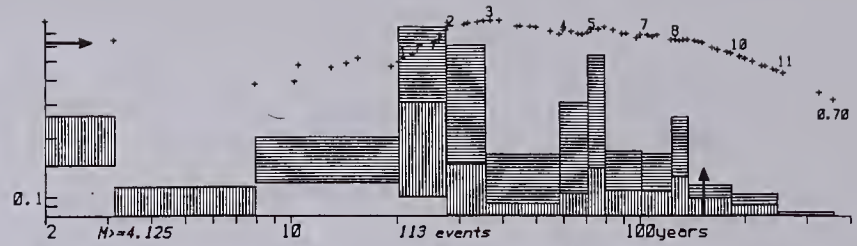
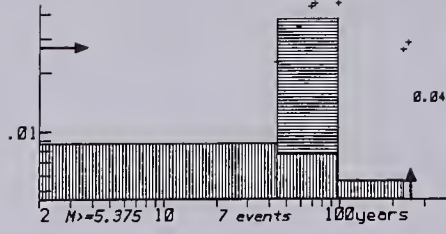
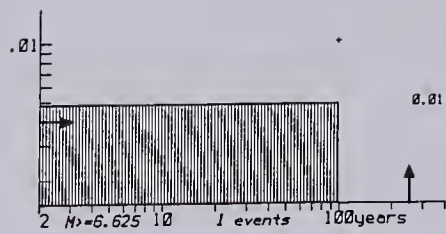


FIGURE 3-6d



FIGURE 3-6e

SEISMICITY RATE - EVENTS PER YEAR



TIME - YEARS BEFORE 1985

EPRI/V EAST

CUMULATIVE - STEPS 0.25M

FIGURE 3-6f

FIGURE 3-7 Temporal distribution of seismicity within cumulative magnitude categories in steps of 0.20 magnitude units. This figure displays the same data in the same form as Figures 3-6 and 3-8: six sets of plots giving the distribution for NCEER, EPRI/V, and EPRI/S - east of the Appalachians (EAST) and west of the Appalachians (WEST); refer to the sample plot in Figure 3-5 for a detailed explanation. The main difference between Figures 3-6 and 3-7 is the size of the magnitude steps. Another peculiarity of Figure 3-7 is that estimates of seismicity rate and completeness are given as a range between a maximum and minimum value, rather than a single value as in Figure 3-6.

SEISMICITY RATE - EVENTS PER YEAR

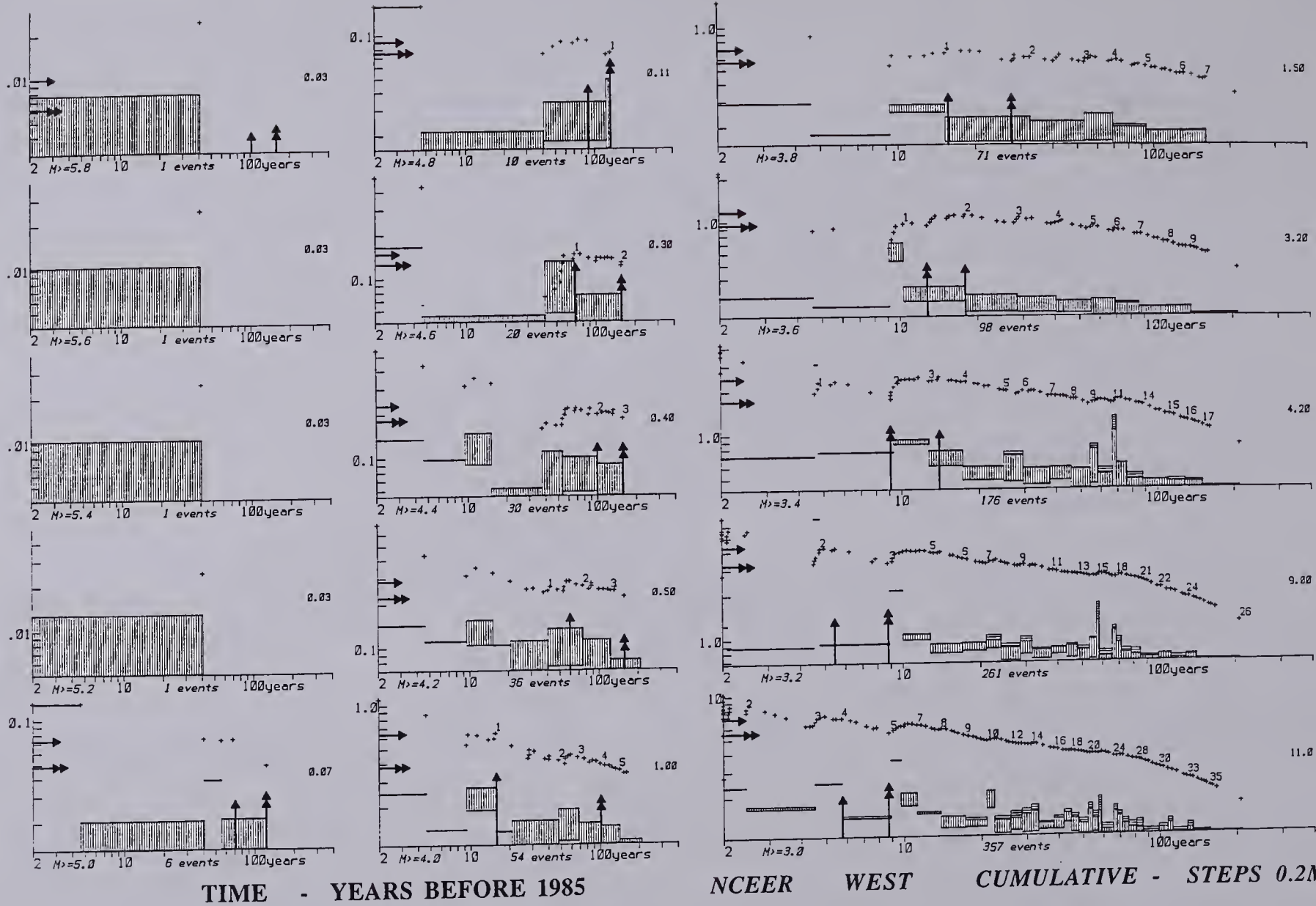
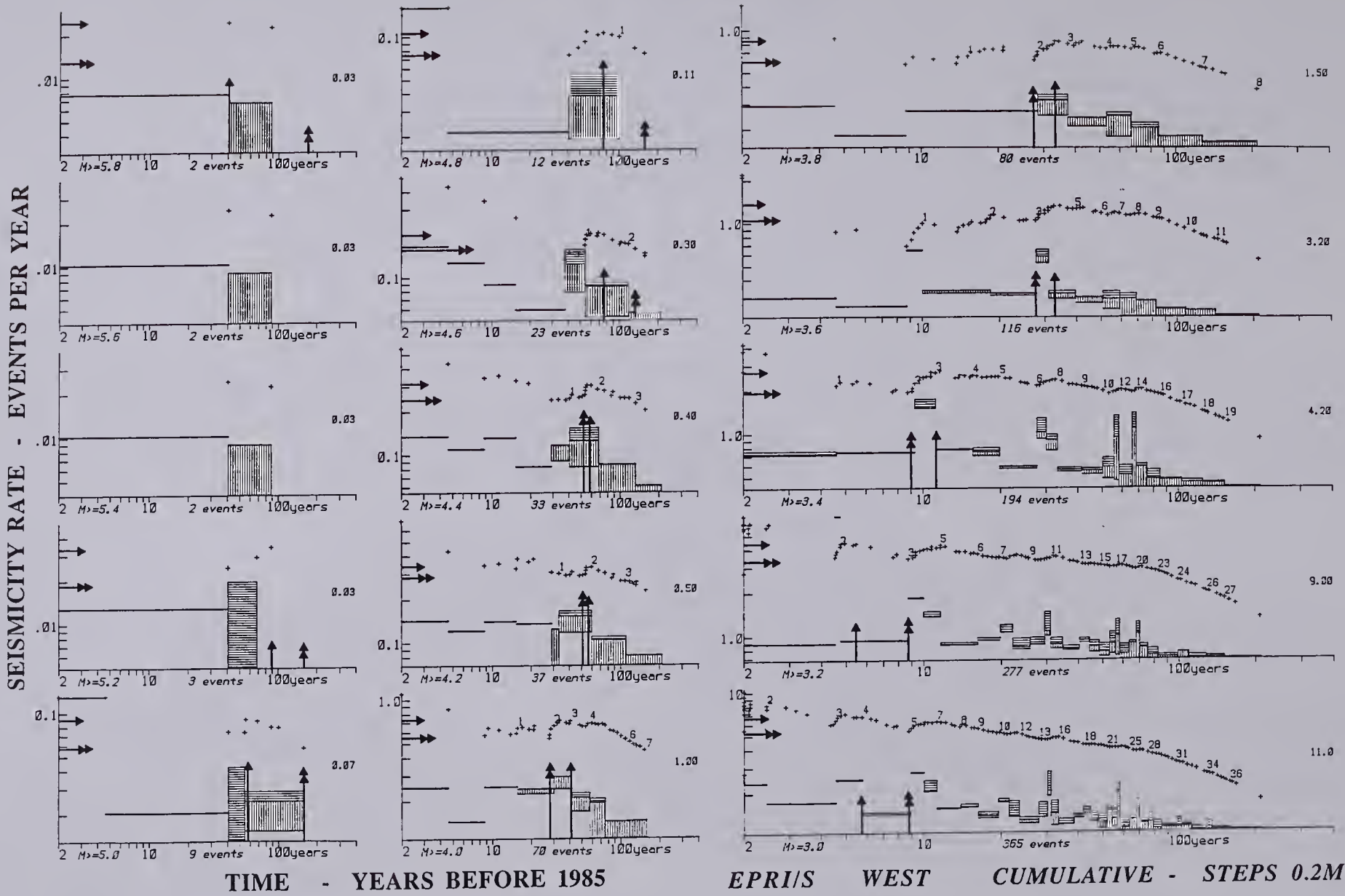


FIGURE 3-7a



TIME - YEARS BEFORE 1985

EPRI/S WEST CUMULATIVE - STEPS 0.2M

FIGURE 3-7b

SEISMICITY RATE - EVENTS PER YEAR

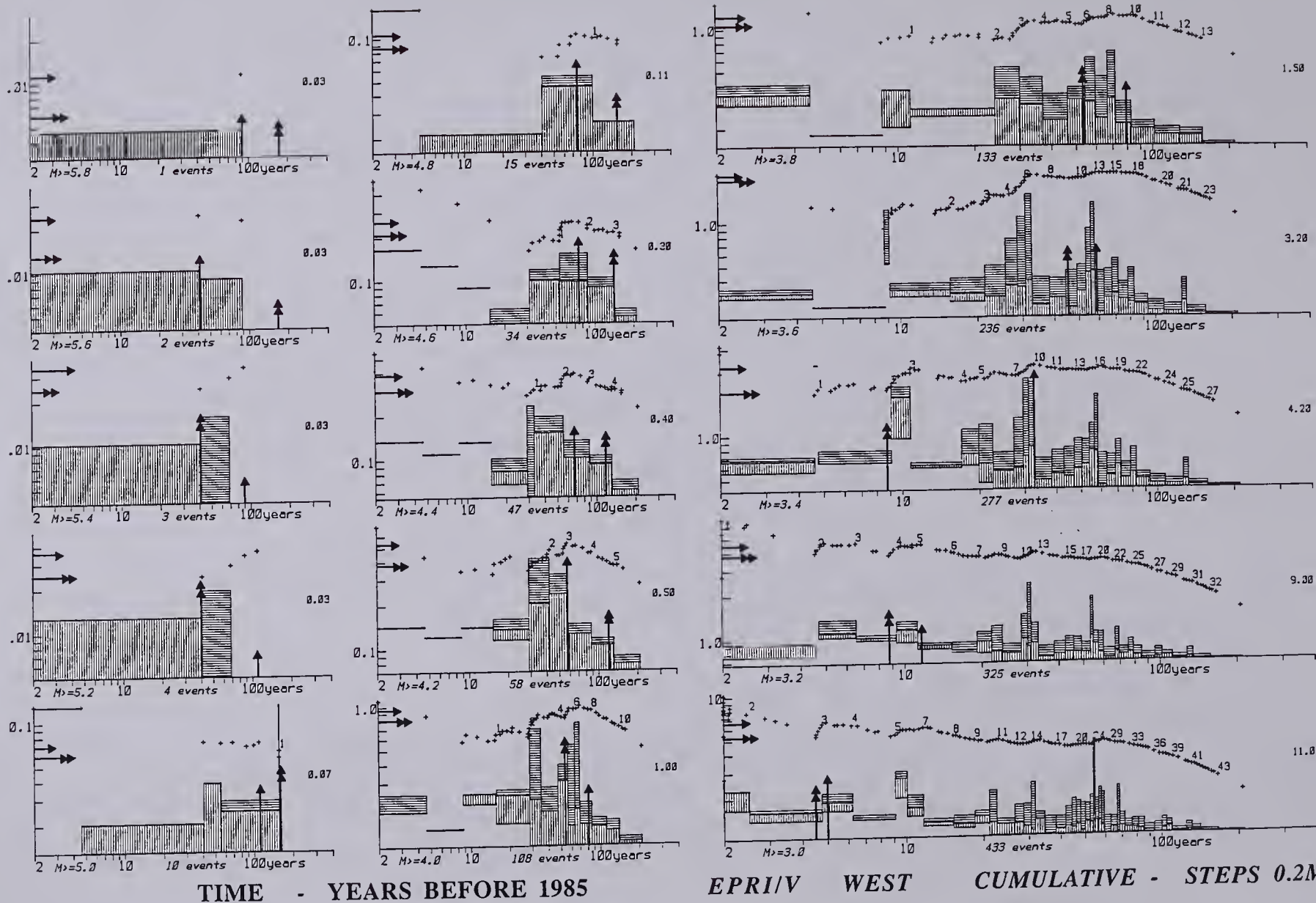


FIGURE 3-7c

SEISMICITY RATE - EVENTS PER YEAR

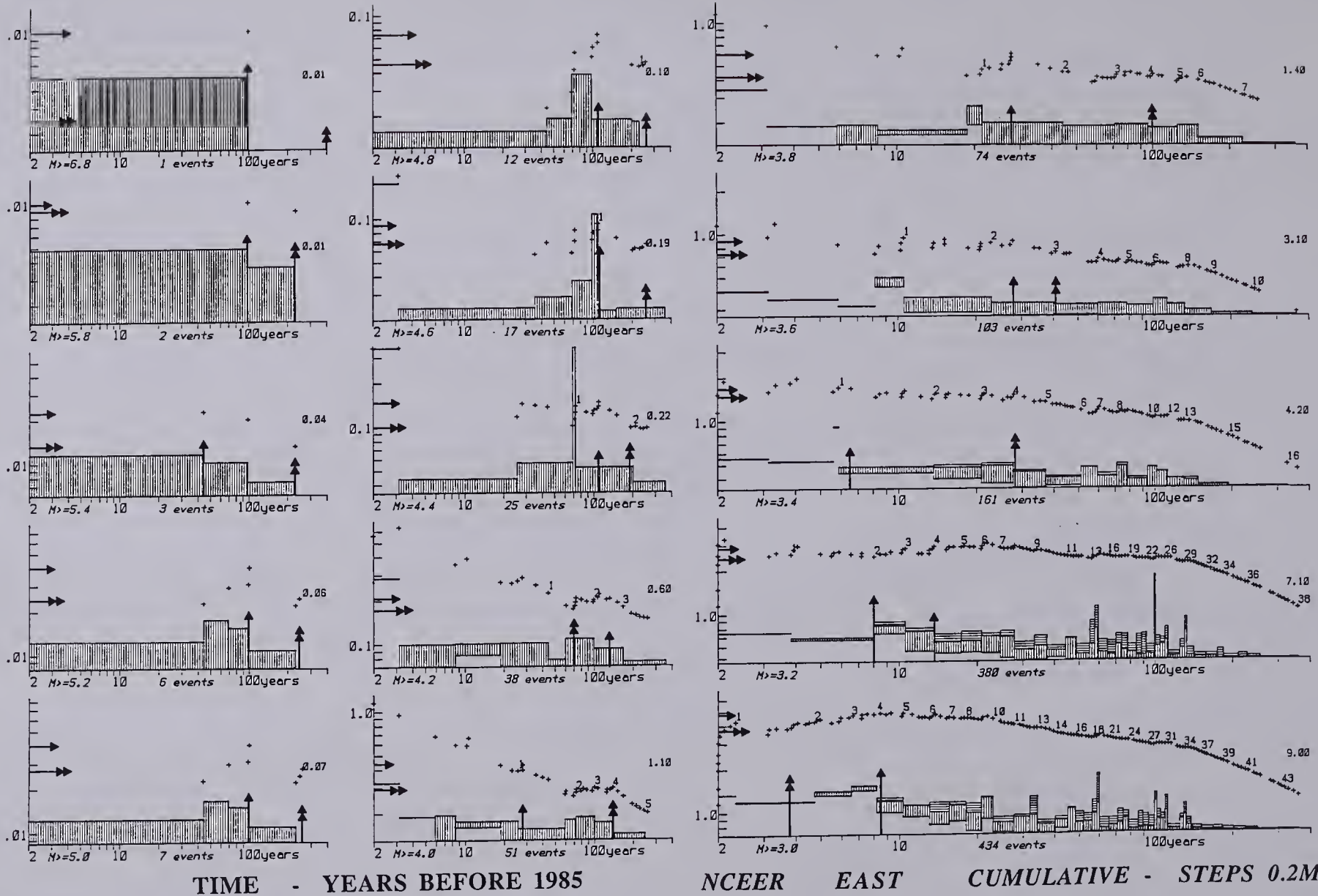


FIGURE 3-7d



FIGURE 3-7e

3-25

SEISMICITY RATE - EVENTS PER YEAR



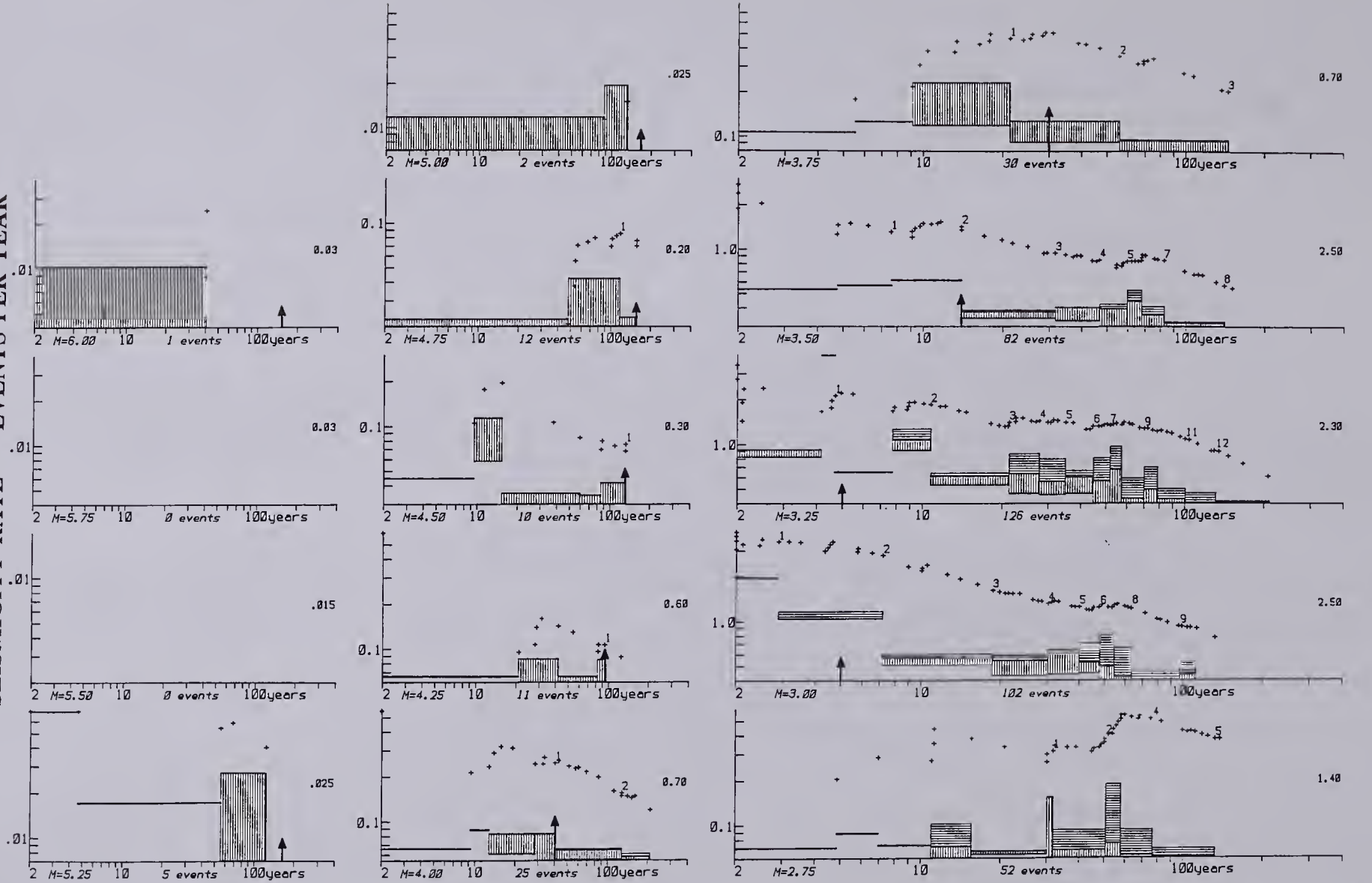
TIME - YEARS BEFORE 1985

EPRI/V EAST CUMULATIVE - STEPS 0.2M

FIGURE 3-7f

FIGURE 3-8 Temporal distribution of seismicity within incremental magnitude windows in steps of 0.25 magnitude units. Each magnitude window is centered at the value specified below the plot. Except for the incremental rather than cumulative nature of the magnitude windows, this figure displays the same data in the same form as Figures 3-6: six sets of plots giving the distribution for NCEER, EPRI/V, and EPRI/S - east of the Appalachians (EAST) and west of the Appalachians (WEST). Refer to the sample plot in Figure 3-5 for a more detailed explanation.

SEISMICITY RATE - EVENTS PER YEAR

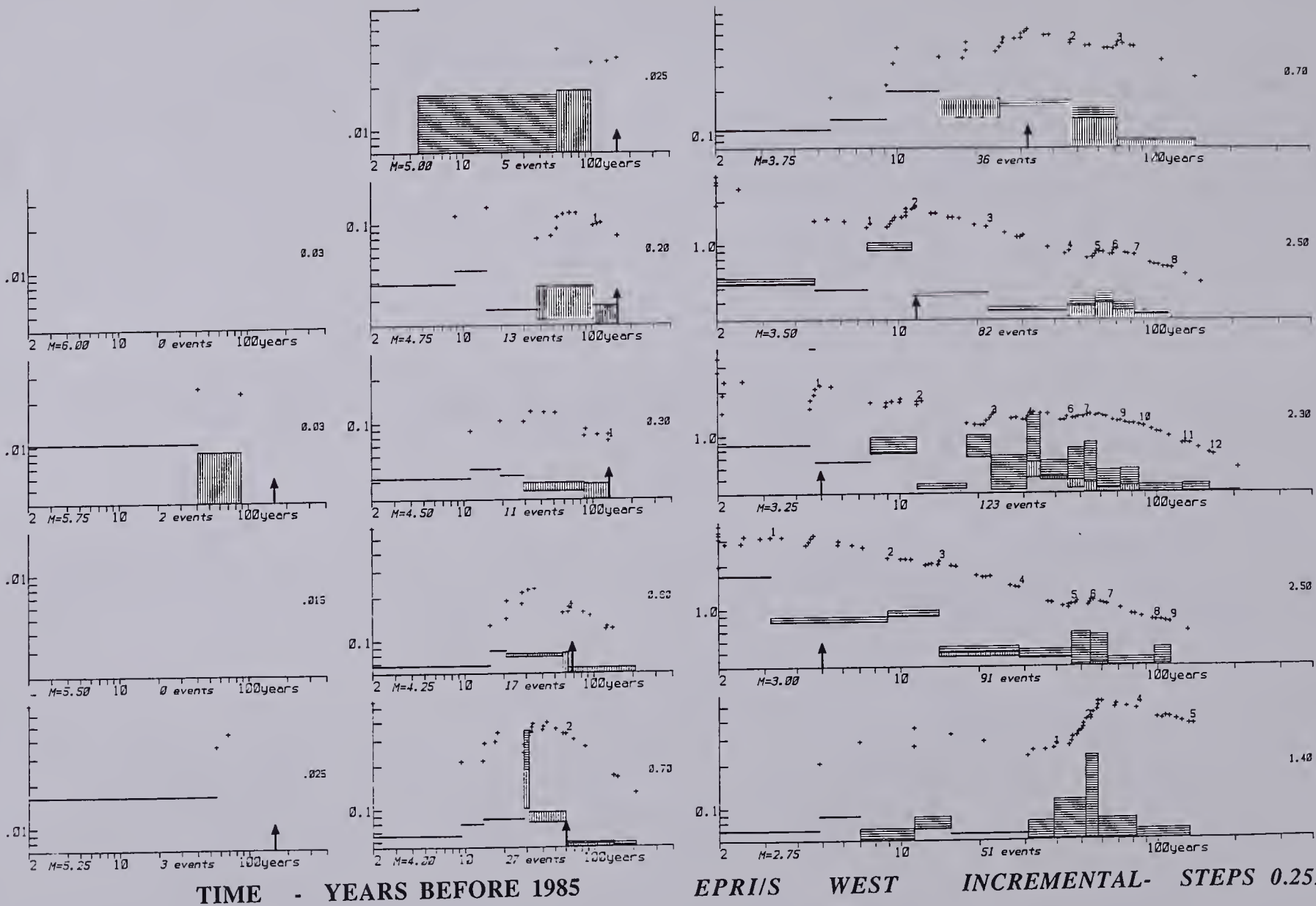


TIME - YEARS BEFORE 1985

NCEER WEST INCREMENTAL - STEPS 0.25M

FIGURE 3-8a

SEISMICITY RATE - EVENTS PER YEAR

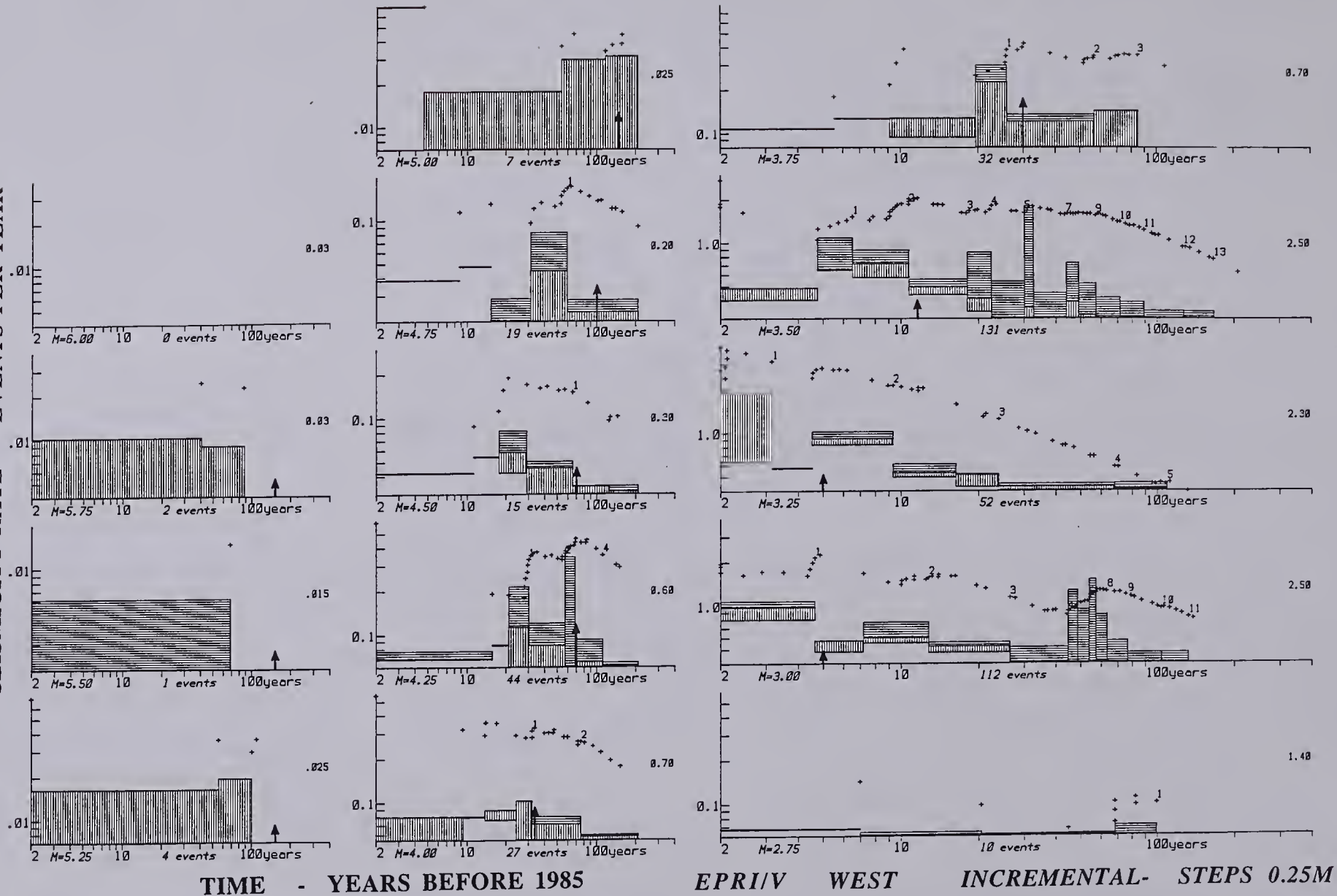


TIME - YEARS BEFORE 1985

EPRI/S WEST INCREMENTAL- STEPS 0.25M

FIGURE 3-8b

SEISMICITY RATE EVENTS PER YEAR



TIME - YEARS BEFORE 1985

EPRI/V WEST INCREMENTAL- STEPS 0.25M

FIGURE 3-8c

SEISMICITY RATE - EVENTS PER YEAR

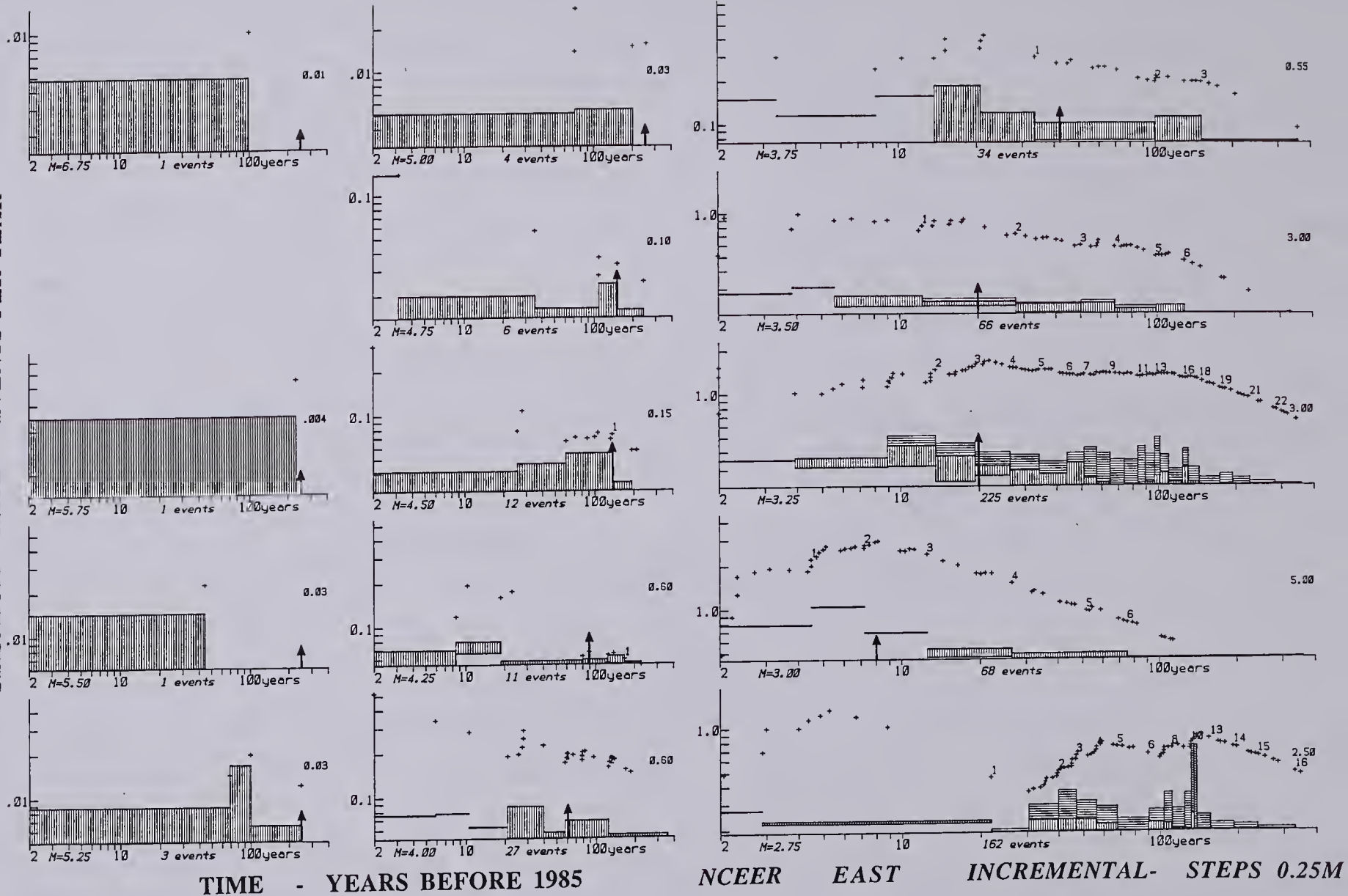
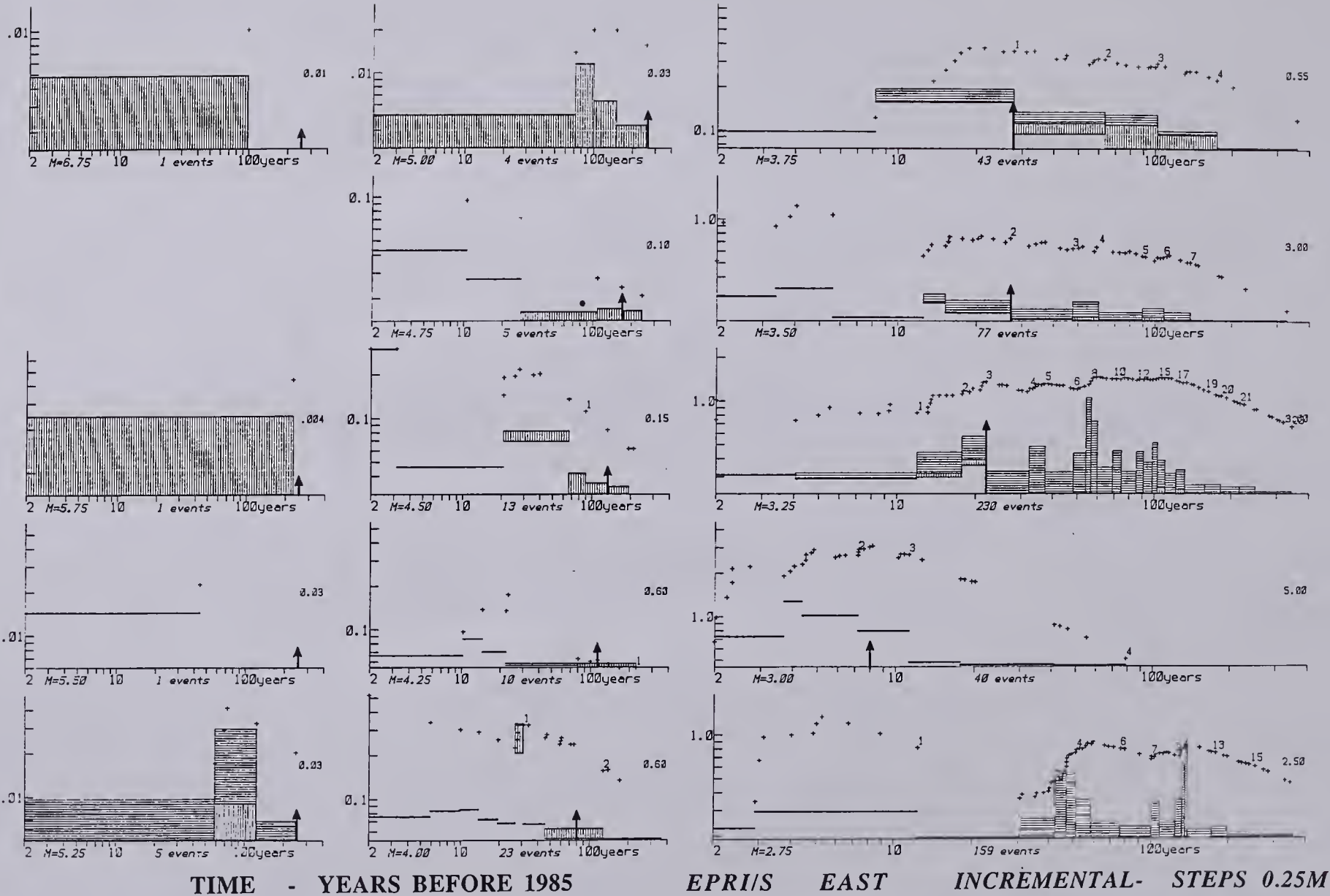


FIGURE 3-8d

SEISMICITY RATE - EVENTS PER YEAR

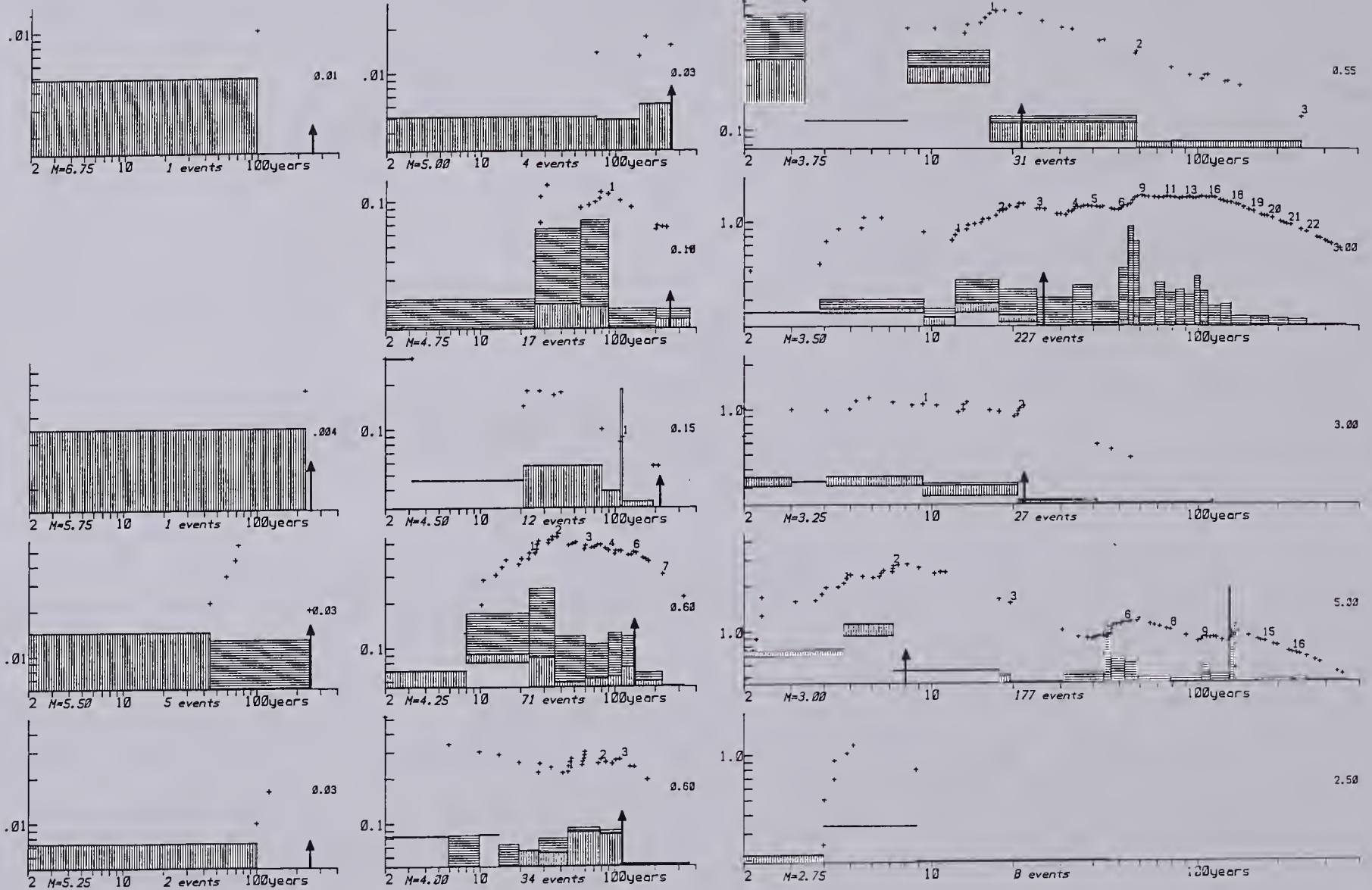


TIME - YEARS BEFORE 1985

EPRI/S EAST INCREMENTAL- STEPS 0.25M

FIGURE 3-8e

SEISMICITY RATE - EVENTS PER YEAR



TIME - YEARS BEFORE 1985

EPRI/V EAST INCREMENTAL- STEPS 0.25M

FIGURE 3-8f

FIGURE 3-9 b-value and repeat-time of $M \geq 6$ determined from a least-squares fit of the log-normal distribution of magnitudes obtained from cumulative steps of 0.25 magnitude units. Error bars are one standard deviation (uncorrelated, constant variance). Each point in the magnitude distribution plots (not shown) is given a weight according to the cumulative count of events down to the corresponding magnitude. The fits are calculated for a set of distributions with low-magnitude cut-offs at each of the steps. b-value and repeat-time are plotted as functions of these low-magnitude cut offs. The plots in each figure refer to the three catalogs, NCEER, EPRI/V, EPRI/S. One set of plots (A) are for data east of the Appalachians (EAST) and the other (B) are for data west of the Appalachians (WEST). Magnitude distributions are obtained from rates given in Figure 3-6.

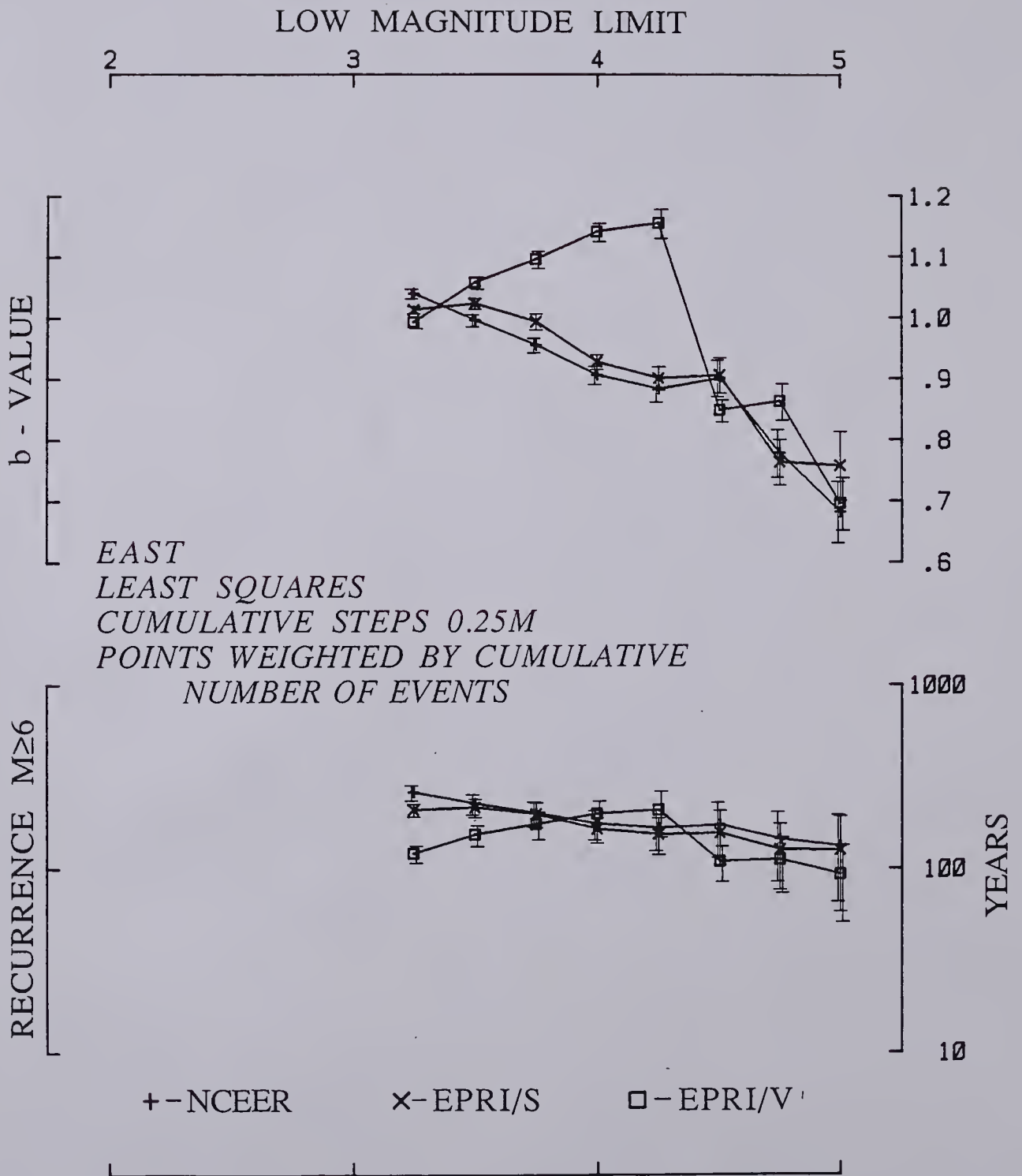


FIGURE 3-9a

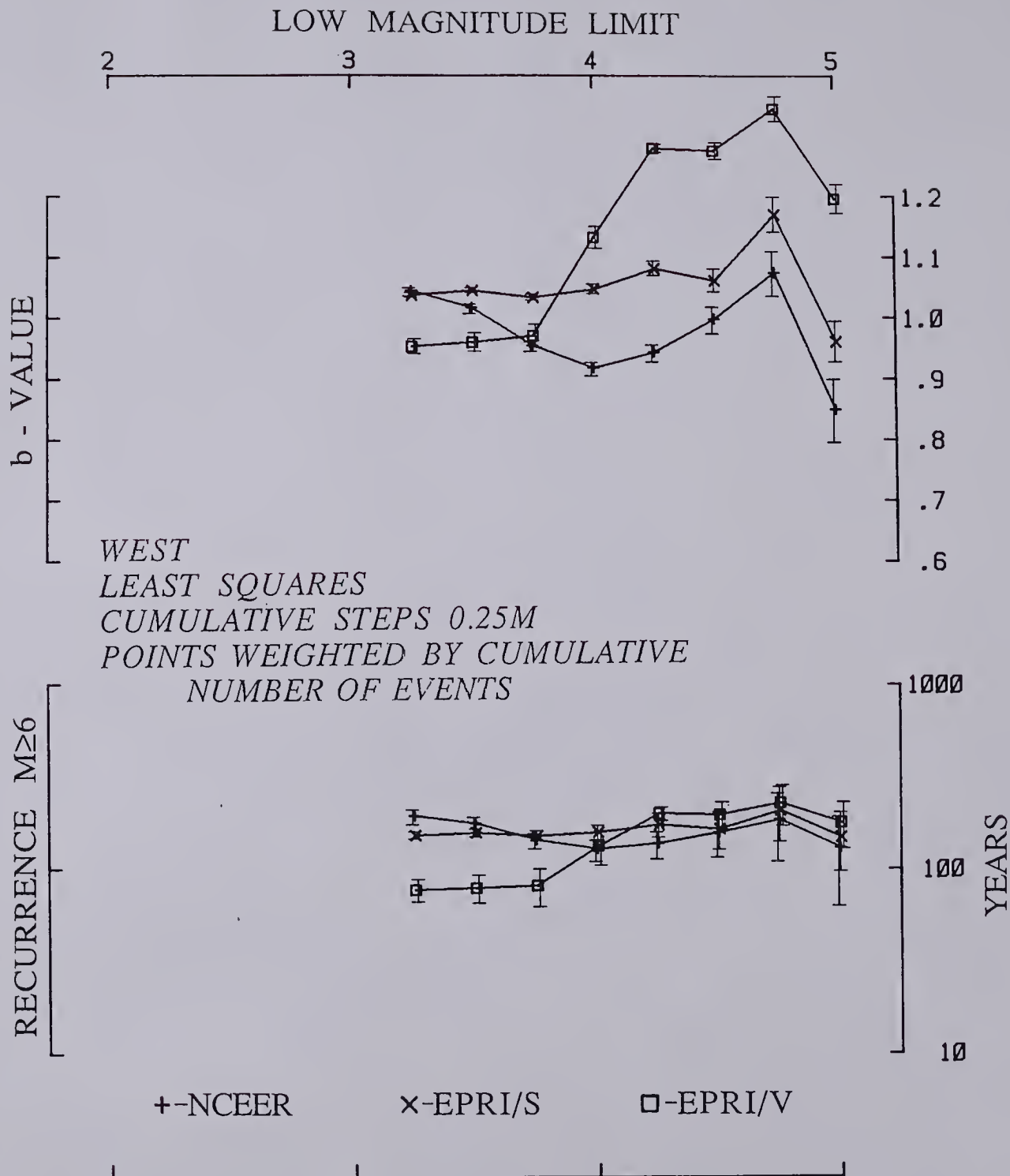


FIGURE 3-9b

FIGURE 3-10 Same as Figure 3-9 (see caption), except that each point in the magnitude distribution plots is weighted according to the incremental count of events between that point and the next point 0.25 magnitude units larger.

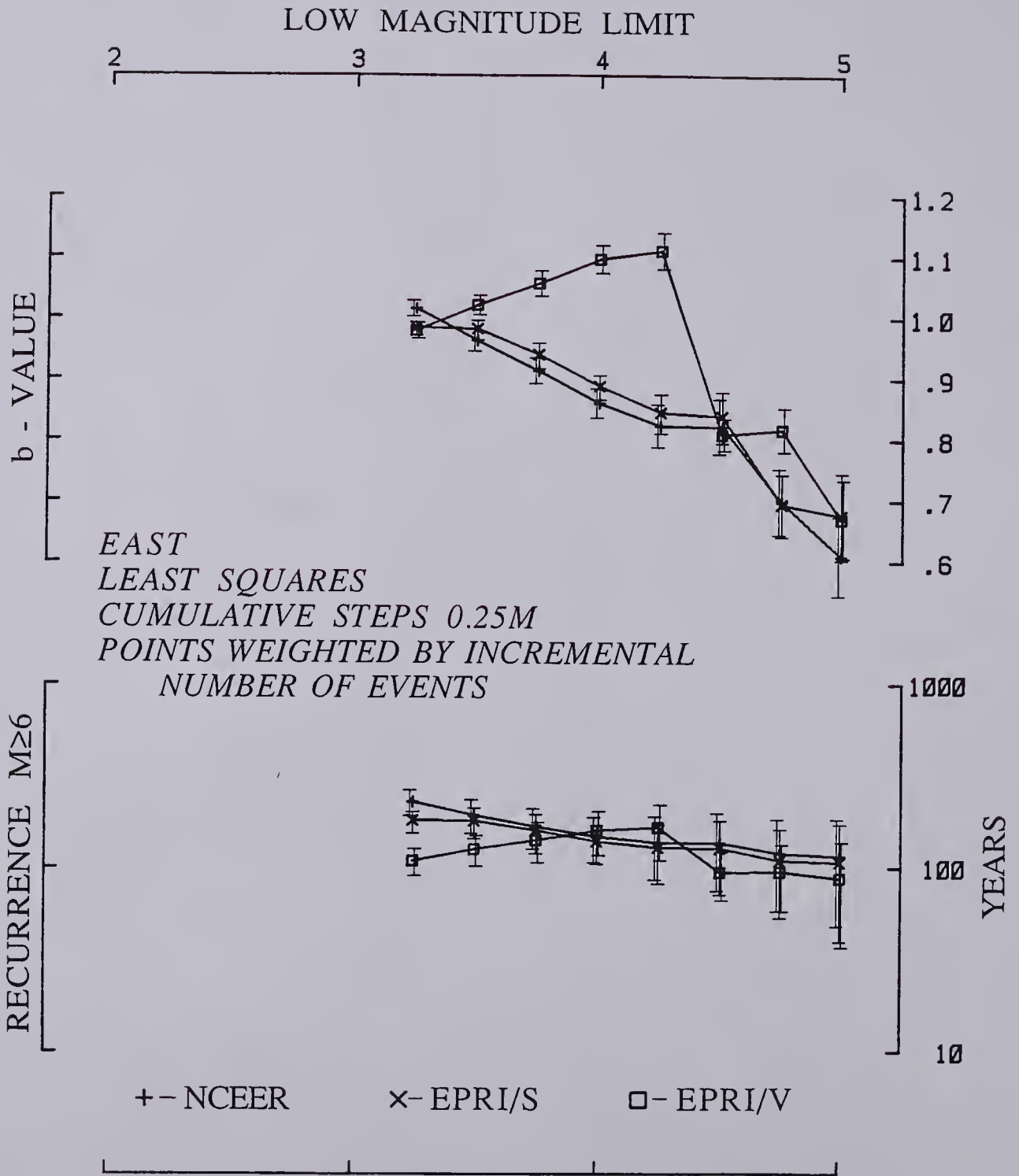


FIGURE 3-10a

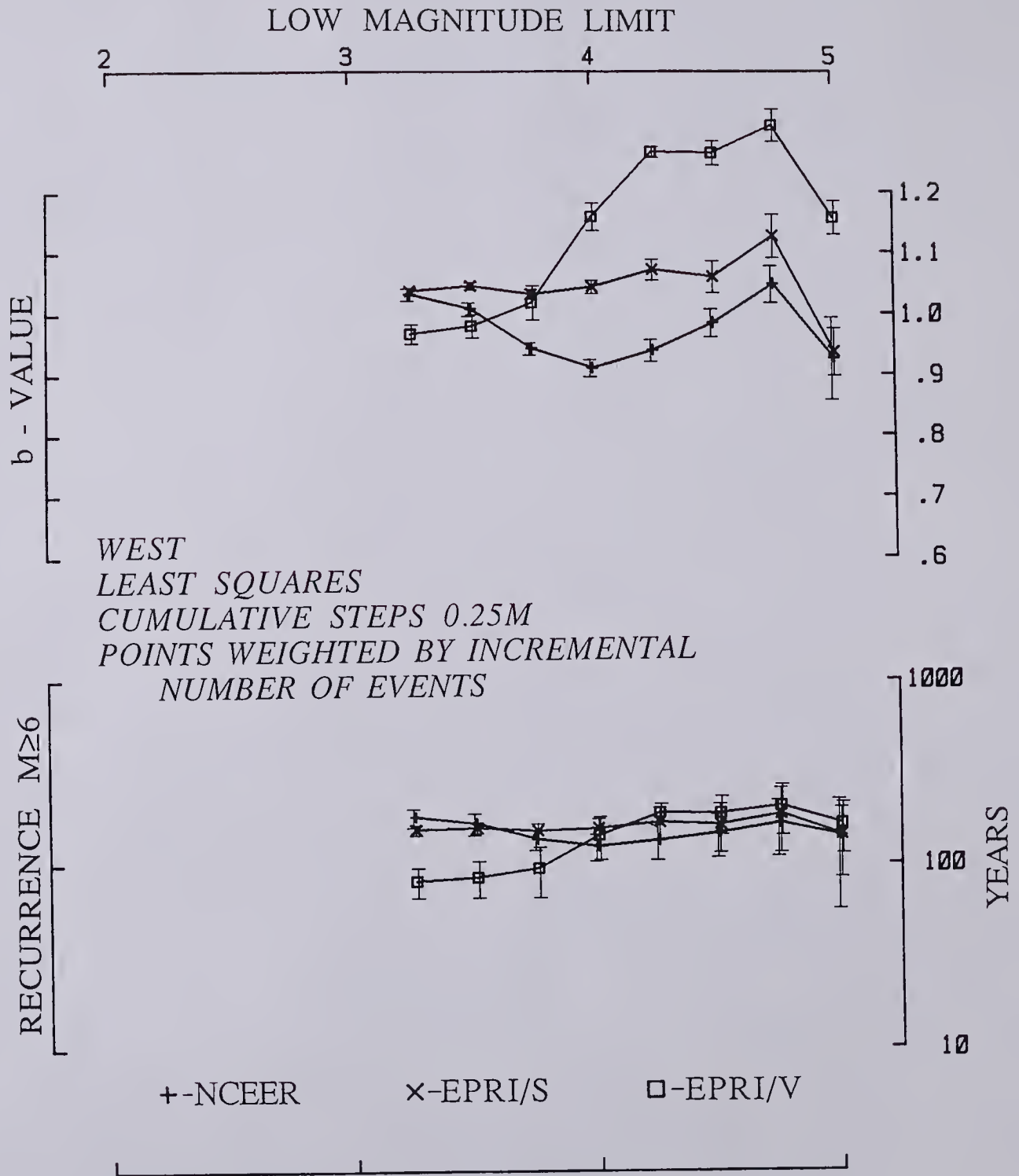


FIGURE 3-10b

FIGURE 3-11 Same as Figure 3-9 (see caption), except that all points in the magnitude distribution plot are weighted 1.

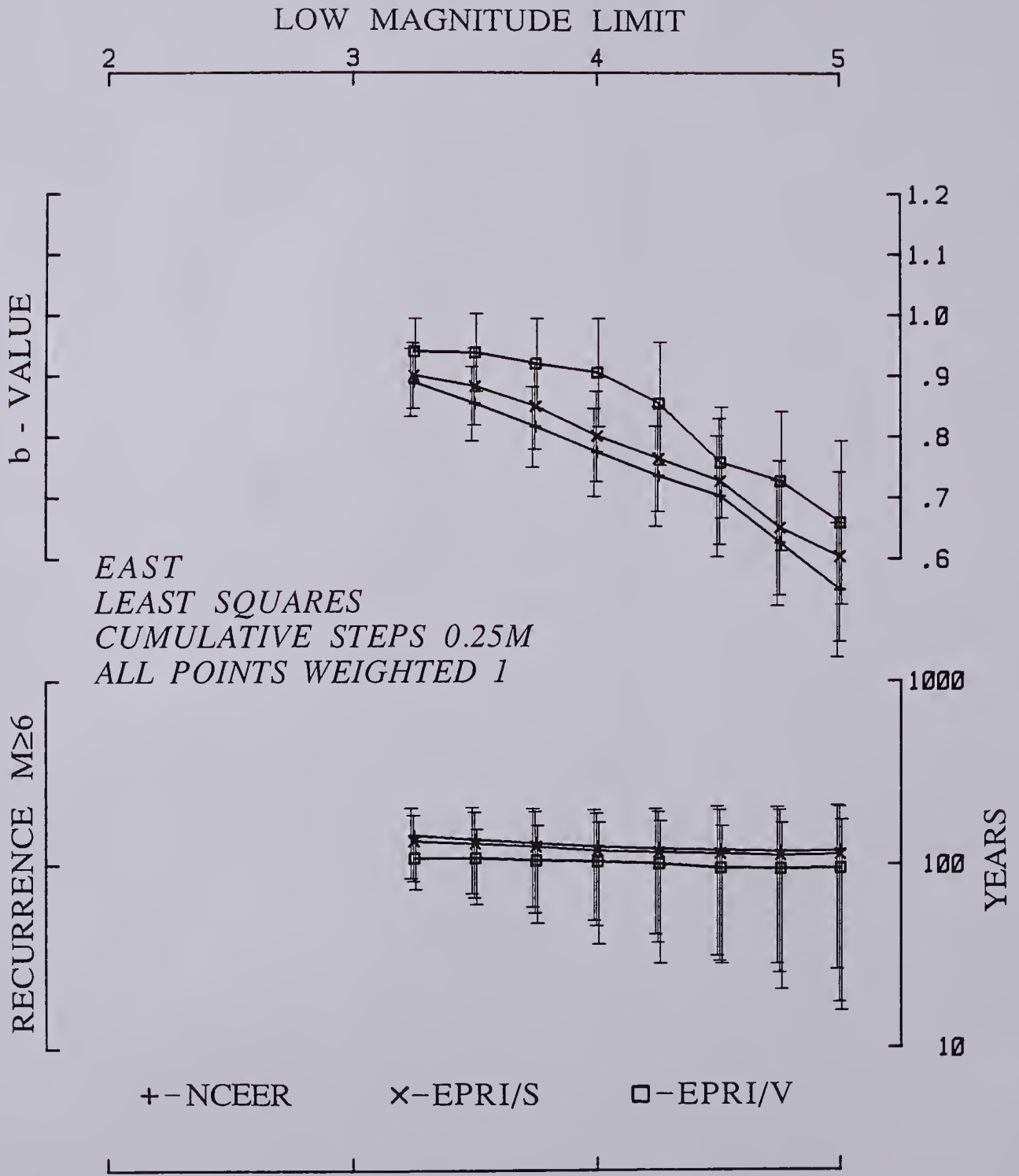


FIGURE 3-11a

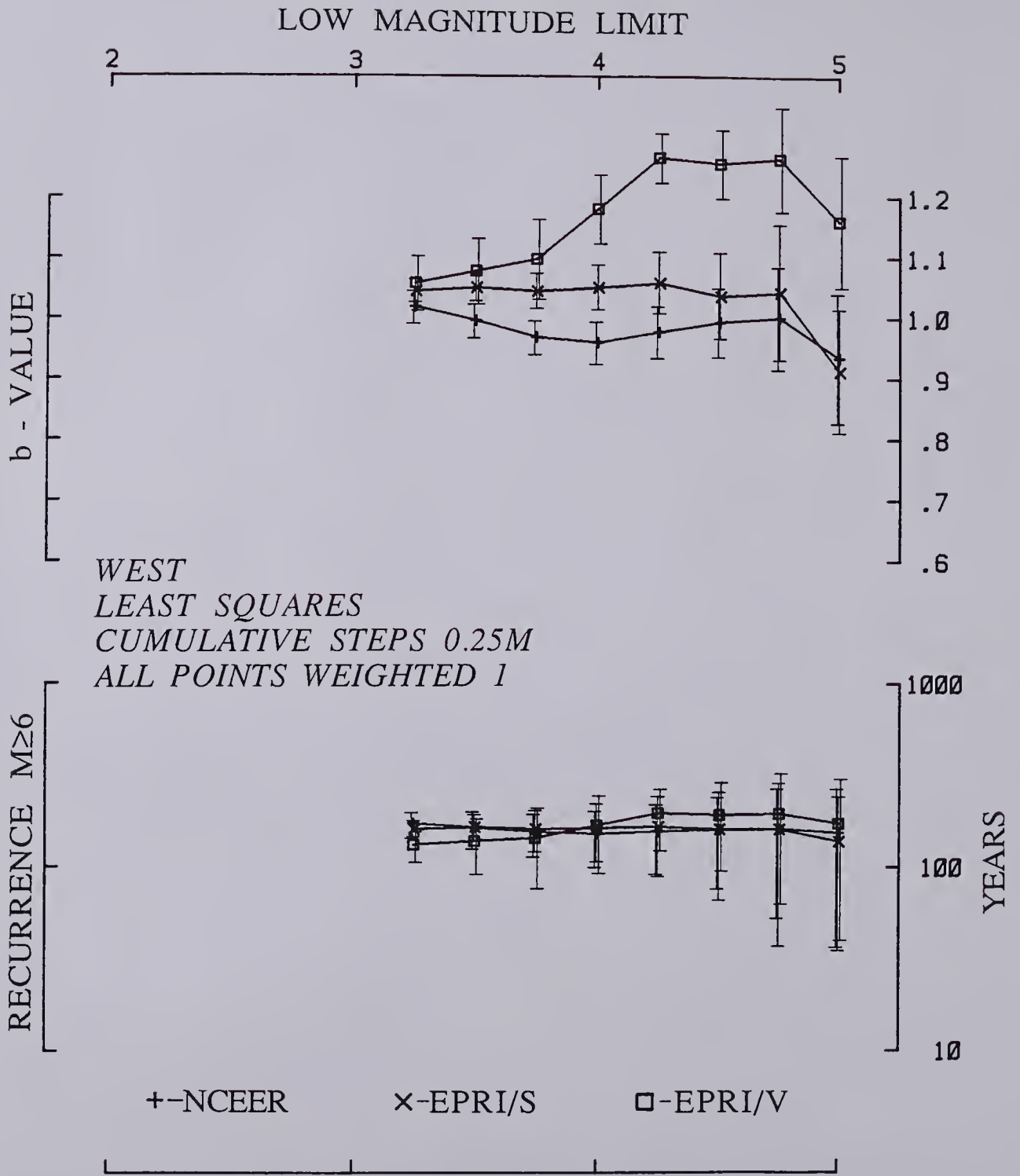


FIGURE 3-11b

FIGURE 3-12 b-value and repeat-time of $M \geq 6$ determined from a least squares fit of the log-normal distribution of magnitudes obtained from cumulative magnitude steps of 0.20 magnitude units. Each point in the magnitude distribution plots is given a weight according to the cumulative count of events down to that magnitude. The fits are calculated for a set of distributions with low-magnitude cut-offs at each of the steps. The plots in each figure refer to the three catalogs, NCEER, EPRI/V, EPRI/S. One set of plots (A) shows results for events east of the Appalachians (EAST) and the other (B) shows results for events west of the Appalachians (WEST). Magnitude distributions are obtained from Figure 3-7.

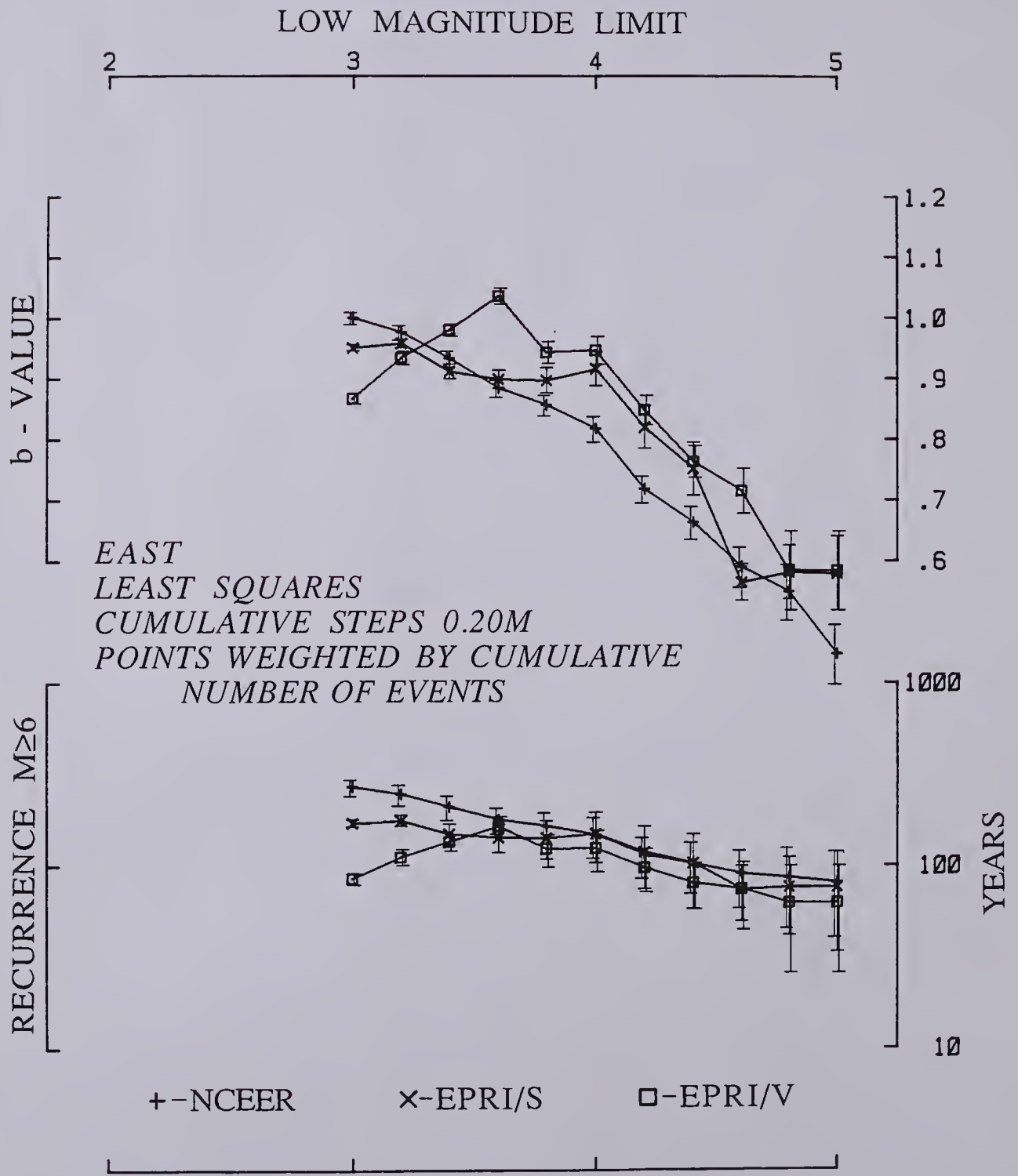


FIGURE 3-12a

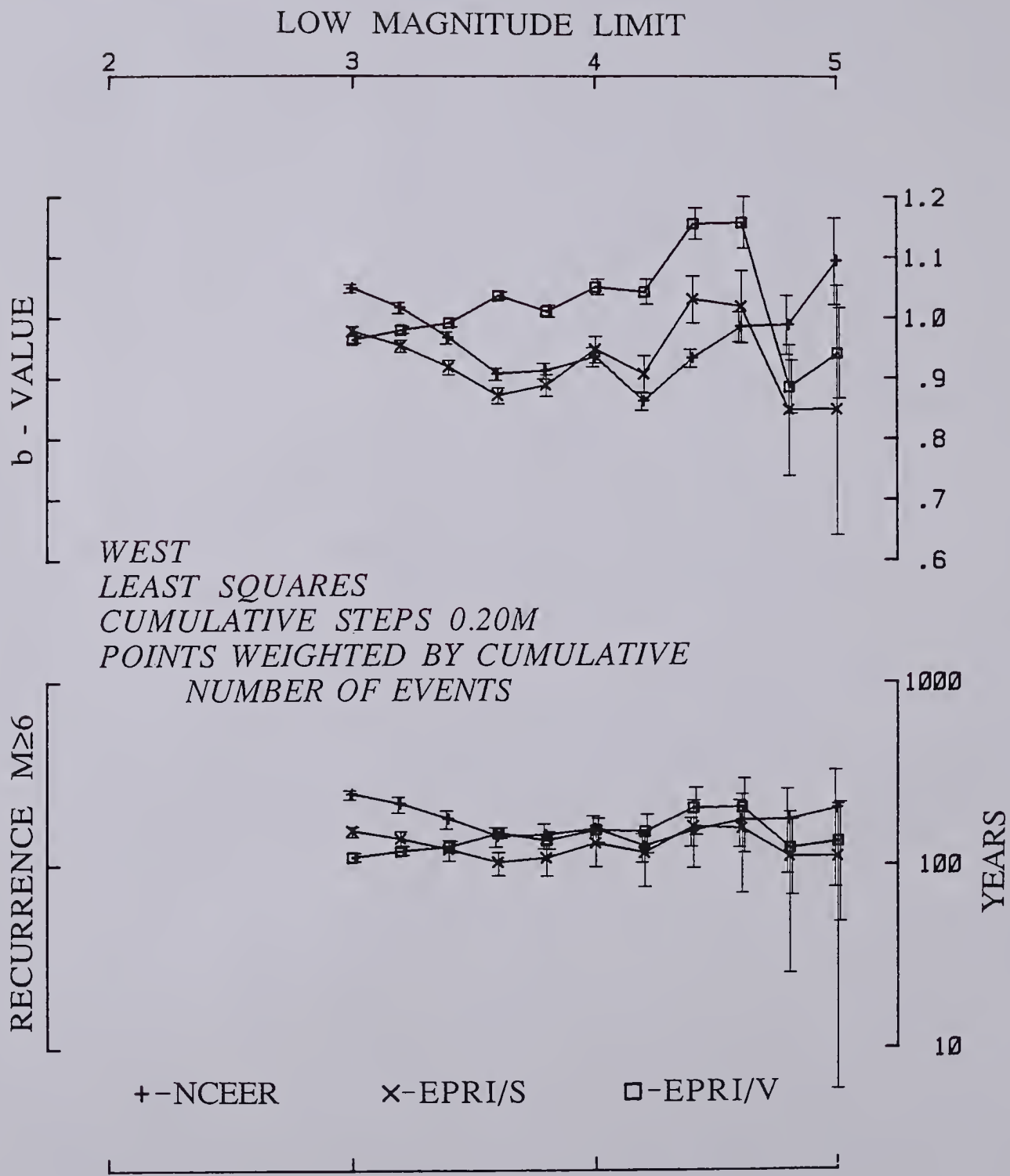


FIGURE 3-12b

FIGURE 3-13 b-value and repeat time of $M \geq 6$ as a function of low-magnitude cut off determined from a maximum likelihood fit of the magnitude distribution (Weichert, 1980) obtained from incremental magnitude steps of 0.25 magnitude units. Error bars are one standard deviation. The plots in each figure refer to the three catalogs, NCEER, EPRI/V, EPRI/S. One set of plots (A) shows results for events east of the Appalachians (EAST) and the other (B) shows results for events west of the Appalachians (WEST). Magnitude distributions are obtained from Figure 3-8.

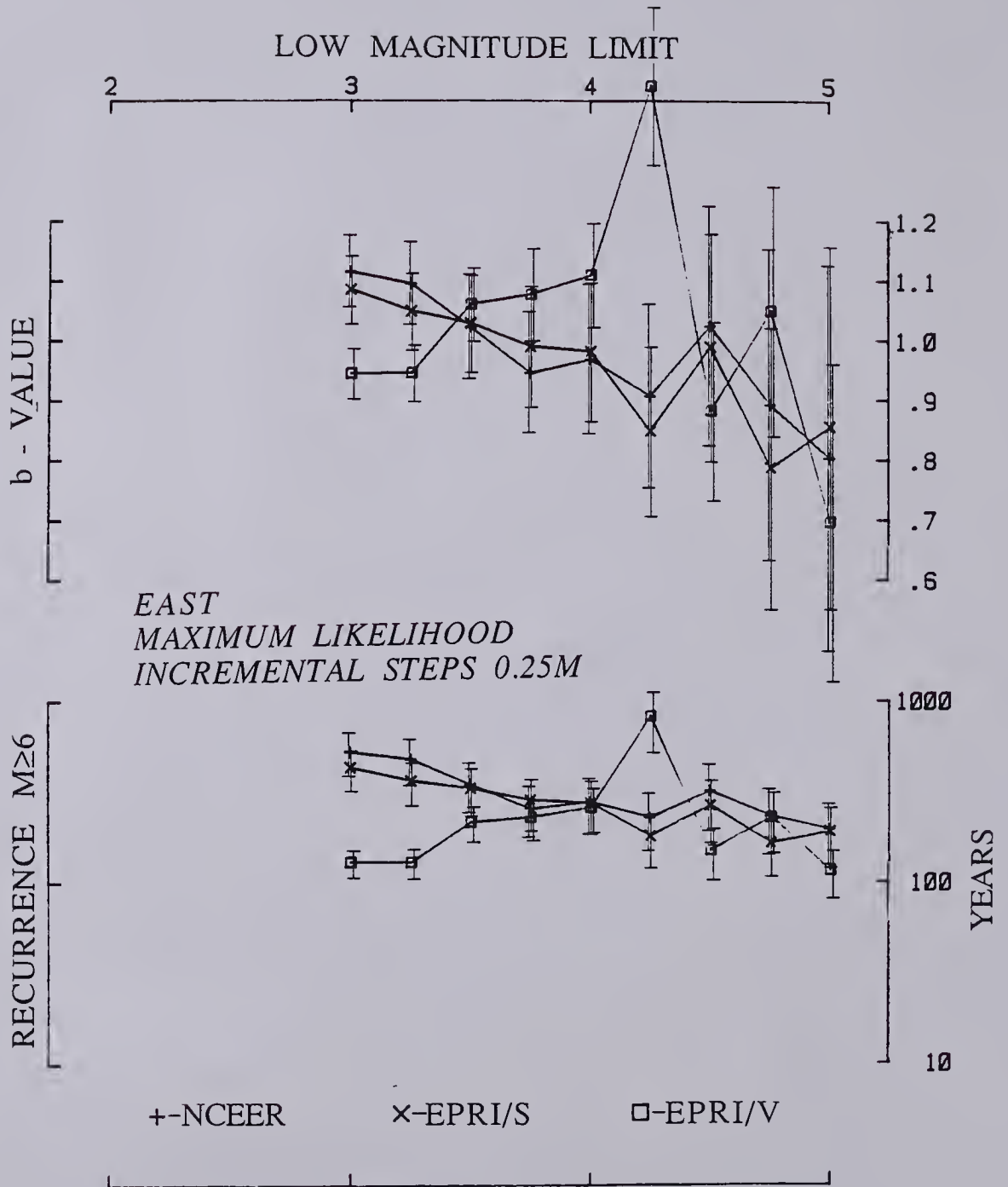


FIGURE 3-13a

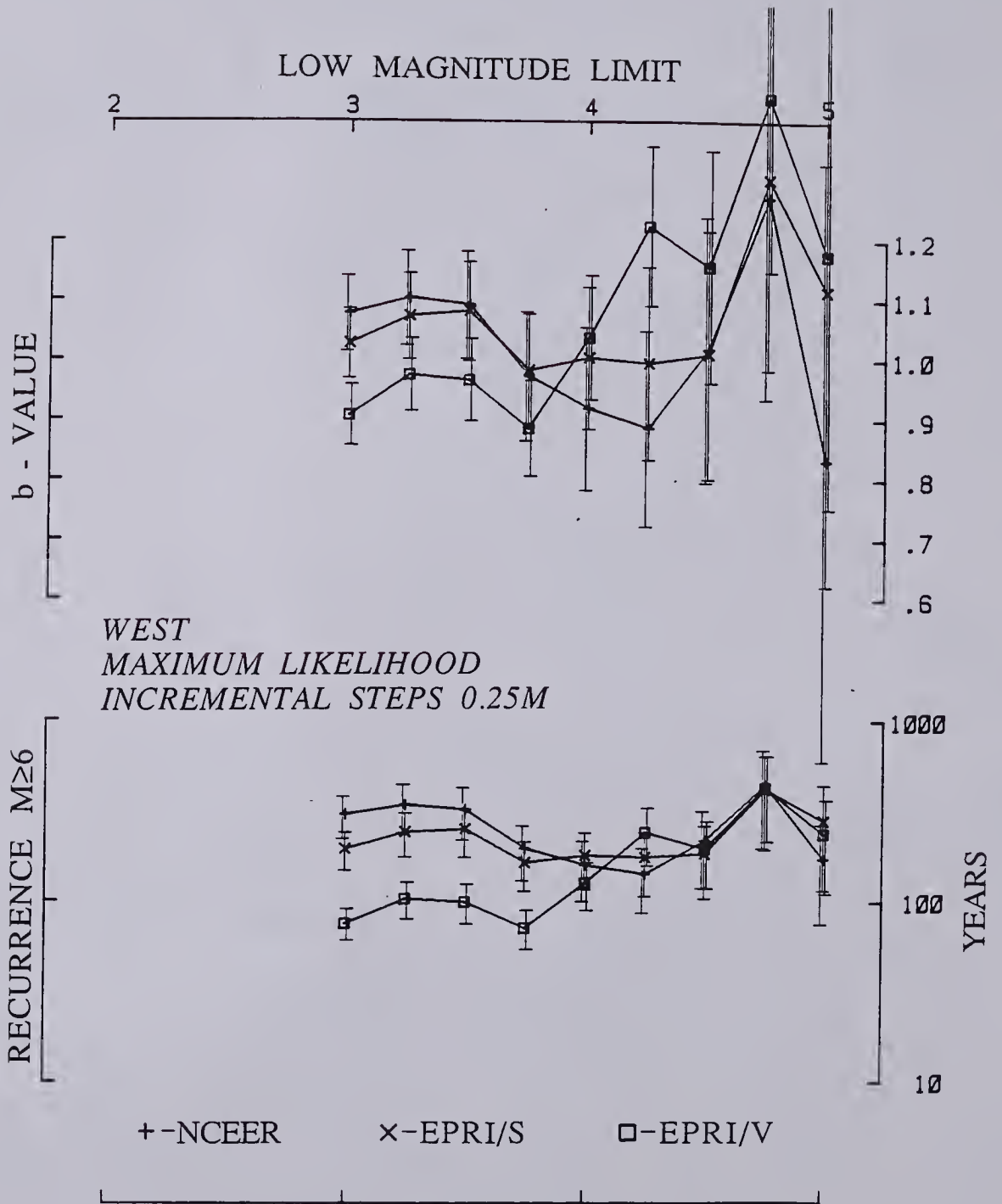


FIGURE 3-13b

FIGURE 3-14 b-values and repeat times of $M \geq 6$ determined from maximum likelihood analysis on incremental magnitude steps of 0.25 magnitude units for NCEER WEST (data from Figure 3-8). Each plot shows results for a particular choice of upper magnitude limit as a function of low-magnitude cut-off. The largest populated magnitude window in WEST is $M=6.0 \pm 1/8$. The completeness period assigned to that population is 200 years. Squares are the parameter values for a seismic zone that has the potential for events up to $M=6.0$ only. Crosses give the results for the previous case with the added possibility of $6.0 < M \leq 7.0$ and no such events in the last 275 years. Plusses are the result for the previous case with the added possibility of $7.0 < M \leq 8.0$ and no such events in the last 300 years. Note that this figure can also be compared with NCEER WEST in Figure 3-13B where the upper magnitude limit is $M=6.0$ with a completeness period of 160 years (see Figure 3-8A).

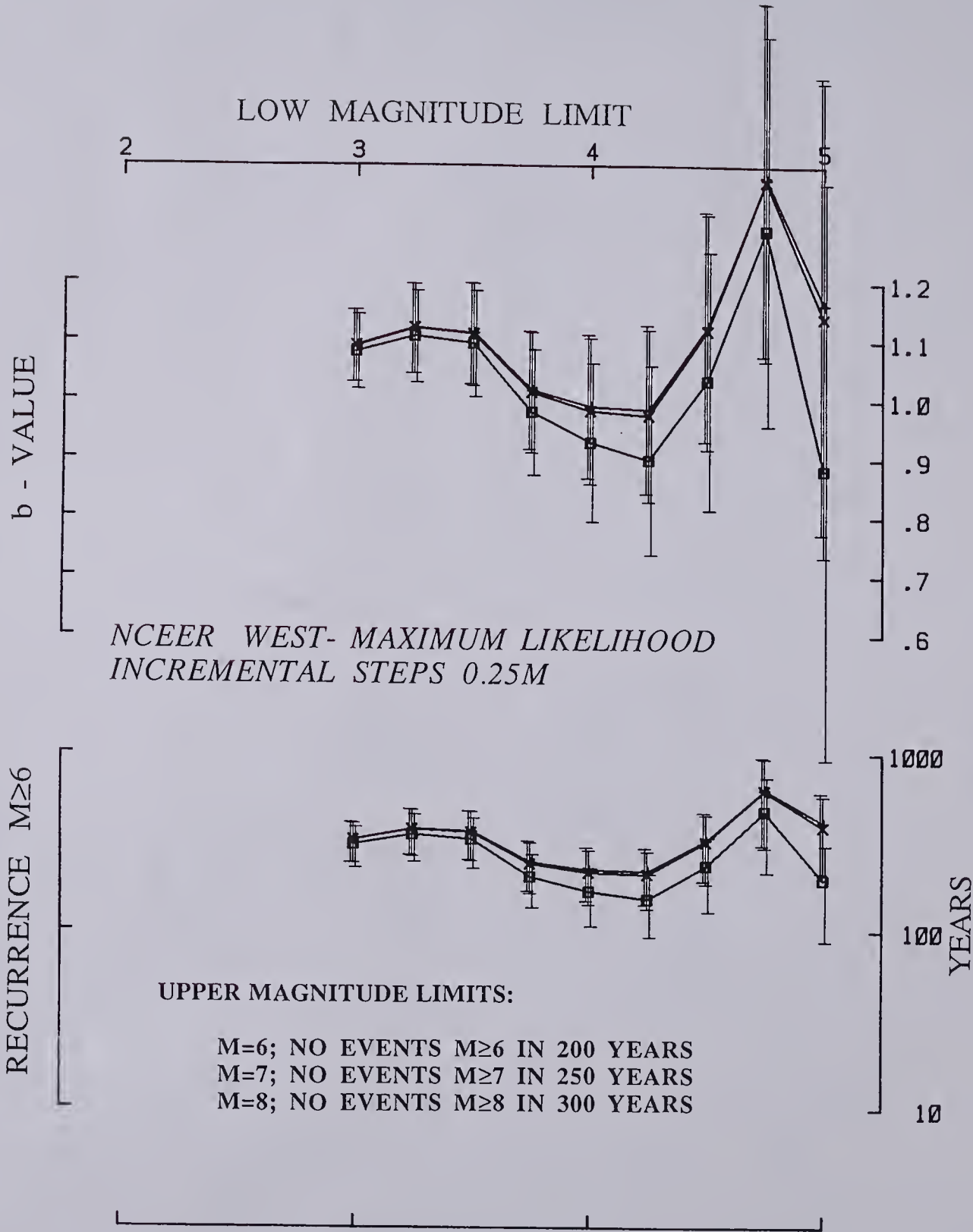


FIGURE 3-15 Effect of changing the rate for the 1886 ($M=6.8$) Charleston-like and larger earthquakes on the b-value and repeat-time of $M \geq 6$ for NCEER EAST. A: Least squares fit of the magnitude distribution obtained with cumulative $0.20M$ steps weighted according to the cumulative number of events. Results are given for three rates of $M \geq 6.75$ events: 1/200 years (plusses); 1/2,000 years (crosses); 1/20,000 years (squares). B: Least squares fit of the magnitude distribution obtained with cumulative $0.25M$ steps weighted according to cumulative number of events. Results are given for three rates of $M \geq 6.75$ as in A. C: Maximum likelihood fit of the magnitude distribution obtained with incremental $0.25M$ steps. Results are given for one event in the $M=6.75 \pm 1/8$ window each 200 years (plusses); each 2,000 years (crosses) and for an infinitely long repeat time (squares). Note that a vanishing rate differs from the case of excluding the possibility of having earthquakes in that window (i.e., a magnitude cut-off at $M=6.5+1/8$; dashed plot). Note also that the results in A, B, and C can also be compared to Figures 3-12A (1 $M=6.75$ in 160y), 3-9A (1 $M=6.75$ in 250y), and 3-13A (1 $M=6.75$ in 250; upper limit $M=6.75$), respectively.

LOW MAGNITUDE LIMIT

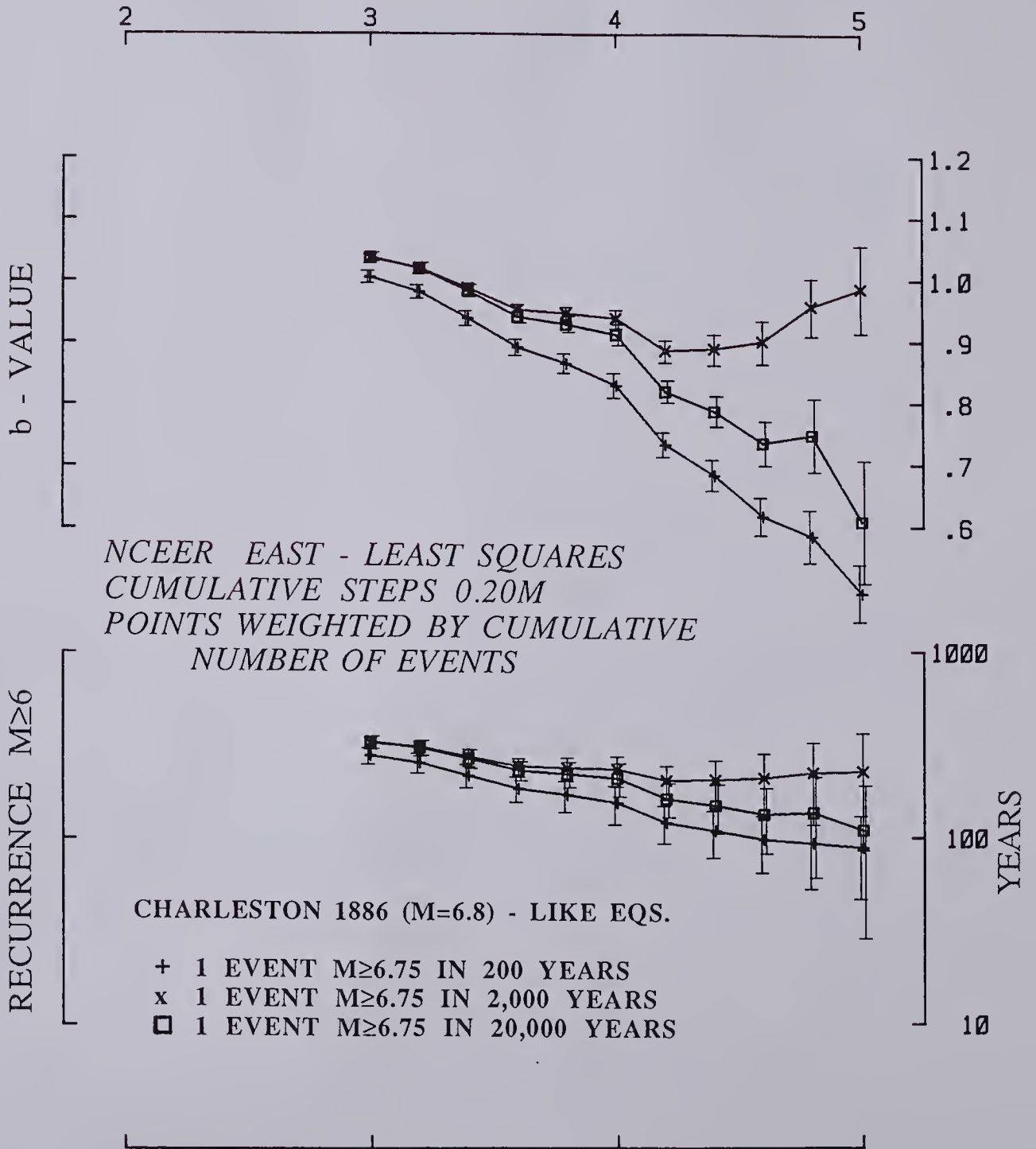


FIGURE 3-15a

LOW MAGNITUDE LIMIT

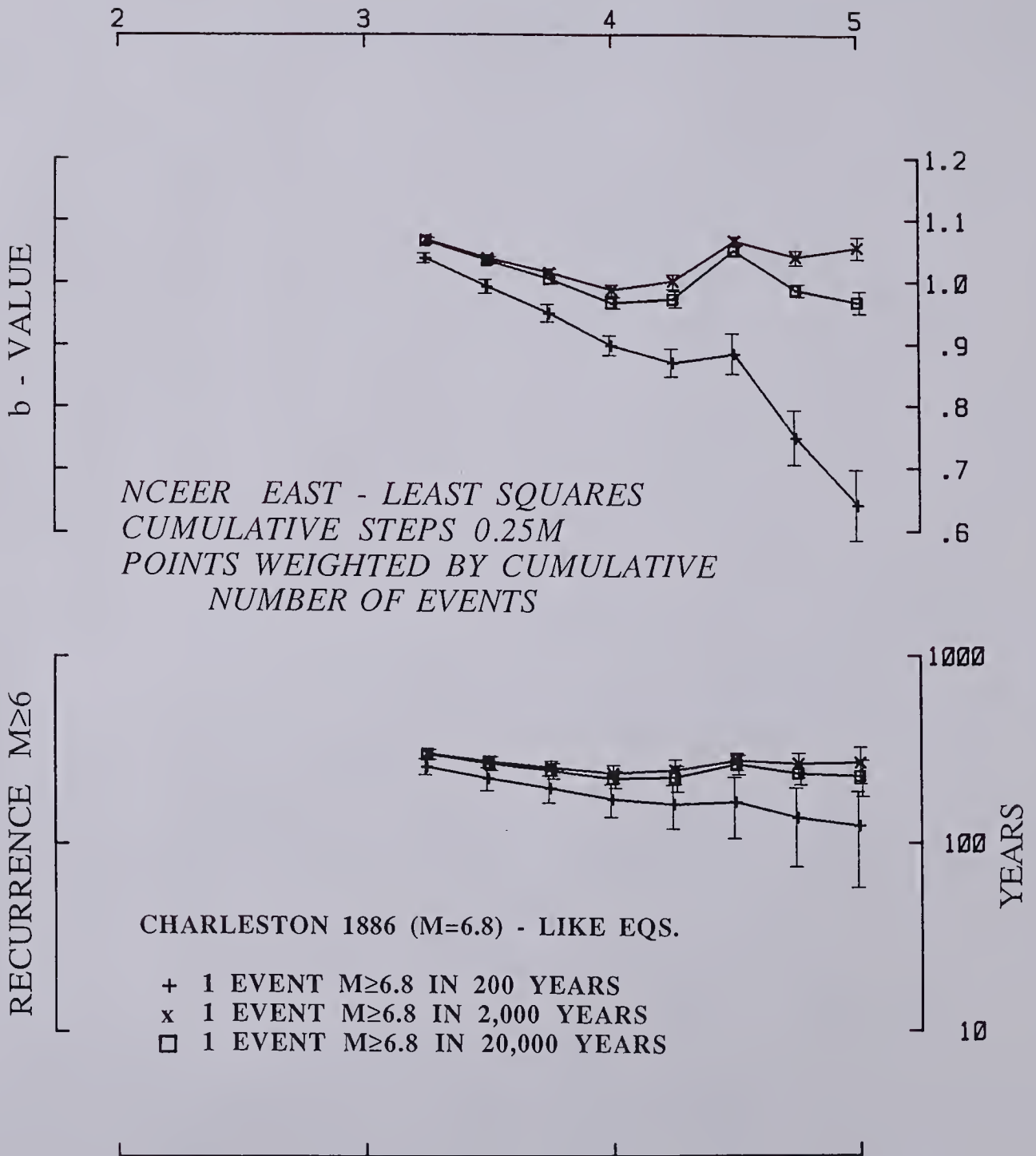


FIGURE 3-15b

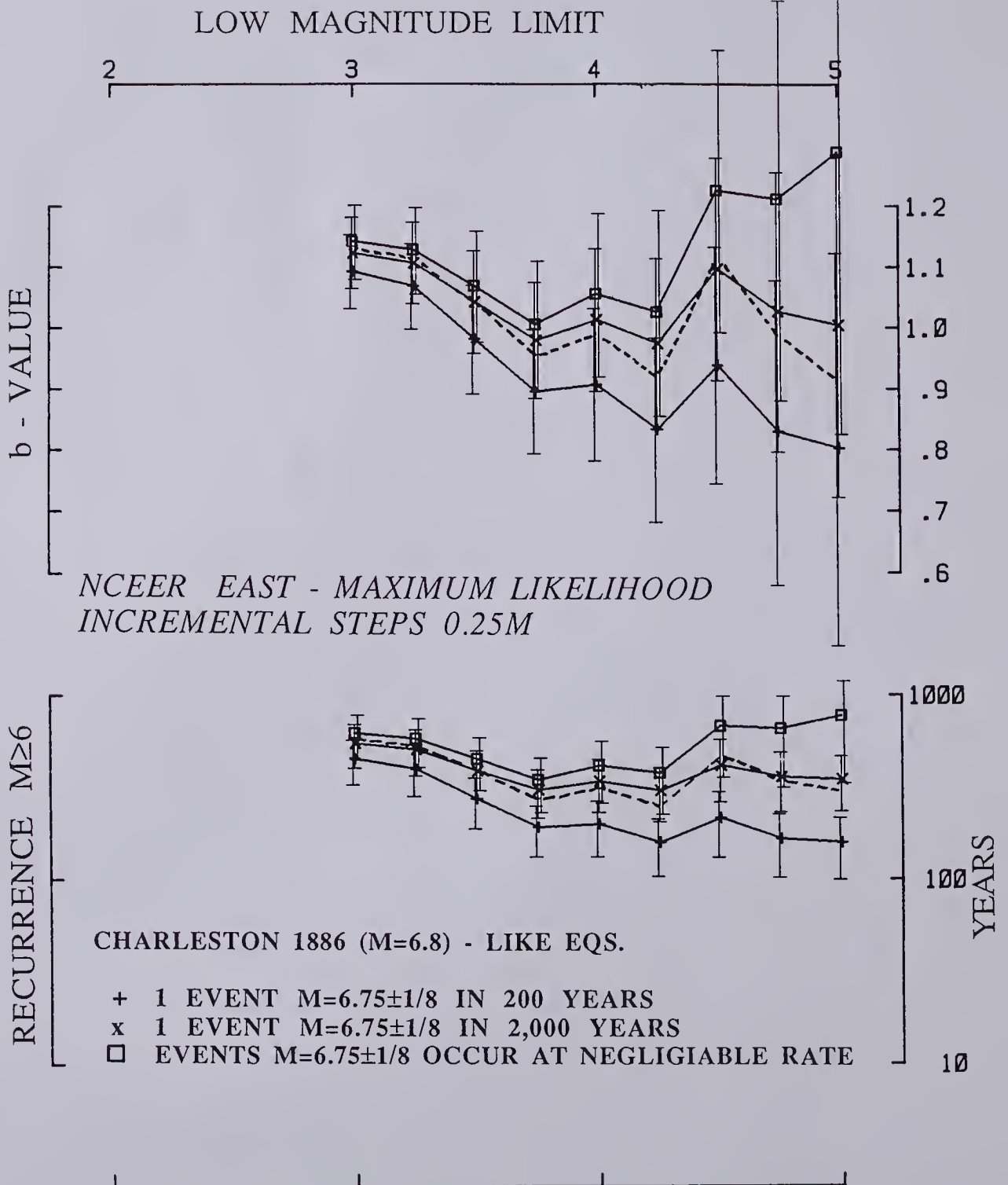


FIGURE 3-15c

SECTION 4
REVISED MAGNITUDES IN THE NCEER-91 CATALOG -- DATA
SOURCES

The first version of the revised catalog, named NCEER-91, covers the area of the eastern U.S. east of the New Madrid seismic zone (east of 85.5W and north of 30.0 N; Figure 3-1). This area contains about half of the entries in the EPRI catalog for the eastern U.S. (east of the Rockies). From the early 1800's, when railroads and telegraph spread rapidly through the eastern U.S., this area is saturated in terms of potential sources of felt reports and the data suggest a reasonably even spatial coverage. In contrast, this same area is very diversified in terms of geology, ranging from the platform/shield of cratonic North America, to the Appalachian fold and thrust belt, and to the Mesozoic rift zone and continental margin. Although seismicity patterns are related to pre-existing structure at many scales (e.g., Seeber and Armbruster, 1988), significant spatial effects on the magnitude distribution related to the tectonic zonation are not yet resolved.

The catalog produced by the Electric Power Research Institute (1987; here named EPRI) reflects the latest comprehensive effort to compile available earthquake data in the eastern U.S. and is used here as the base for the revision. All entries in EPRI were compared to corresponding entries in the compilations listed in Table 3-I. Surprisingly, a substantial portion of the improvements could be made on the basis of U.S. earthquakes and from other readily available sources. The remaining data came from compilations pertaining to relatively small areas and from John Armbruster's unpublished compilation derived from an ongoing search of archival material which has not yet uniformly covered the area of this study. As a result, the revision of the catalog is not complete nor spatially uniform.

Criteria for revising the magnitudes were diverse, but hinged on the realization that felt-area, or, more generally, the spatial distribution of intensities, is a much more reliable measure of magnitude than maximum intensity. It was also realized that instrumental magnitudes (M_{is}) as listed in EPRI can scatter considerably, depending on the assumed attenuation characteristics and on other procedural factors that cannot be readily verified. On the other hand, our compilation of macroseismic data allows for a systematic determination of felt-area magnitudes (M_{fa}). For this procedure, we have adopted the method proposed by Sibol et al. (1987) where both felt area and maximum intensity contribute to the definition of M_{fa} . Thus, M_{fa} is given priority over M_{is} , unless: 1.) M_{is} is from a special study, such as Dewey and Gordon (1984); or 2.) M_{is} represents a

consensus among several independent determinations as expected for particularly significant earthquakes; or 3.) M_{is} is determined by regional network data. In summary, reliable instrumental magnitudes are ranked at the top, followed by felt-area magnitudes, followed by instrumental magnitudes other than the ones in the above categories, and, finally, maximum intensity magnitudes (M_{mi}) which are ranked the lowest. Table 3-II compares the number of each kind of magnitude in EPRI/S and NCEER-91 and highlights the overall changes. Table 3-III lists these changes over the historic period.

Besides the reliability of the raw data, an equally important issue is the method used in the determination of magnitudes. Generally, EPRI lists several kinds of M_{is} 's and intensity data, such as felt area and maximum intensity, that can be used to calculate M_{fa} 's and M_{mi} 's. From these data we derive two catalogs, EPRI/S and EPRI/V, that list only one magnitude for each event and could be compared to NCEER-91 in terms of magnitude distribution. In both these catalogs and in NCEER-91, instrumental magnitudes are chosen as the largest listed in EPRI, but while EPRI/S and EPRI/V list instrumental magnitudes whenever available, M_{fa} takes precedence in NCEER-91 when M_{is} is not particularly reliable (see above).

EPRI/S and EPRI/V differ in the way magnitudes are derived from intensity data. In EPRI/V, M_{mi} and M_{fa} are determined with relations proposed by Veneziano and Van Dyck (1985); in EPRI/S, M_{mi} and M_{fa} are determined from relations proposed by Sibol et al. (1987). Of the algorithms to derive M_{mi} , the one proposed by Sibol et al. (1987) seems preferable in at least two respects. First, while Veneziano and Van Dyck (1985) arbitrarily assume that the relation between maximum intensity (MI) and M_{mi} is linear, Sibol et al. (1987) allow the relationship to assume the best fitting shape and show that the linear relation overestimates M_{mi} in the MI V-VI range. Secondly, the algorithm by Sibol et al. (1987) solves for a range of M_{mi} for each MI. This procedure has the advantage of smoothing over discrete intensity levels and produces a M_{mi} distribution less likely to interfere with magnitude steps in the 'b-value' analysis. Figures 3-1, 3-2, and 3-3 show the spatial distribution of magnitudes in NCEER-91, EPRI/S, and EPRI/V; Figure 3-4 compares magnitudes in EPRI/S and NCEER-91.

M_{fa} 's and M_{mi} 's in NCEER-91 are derived from the relations proposed by Sibol et al. (1987). M_{fa} in these relations depends on maximum intensity as well as felt area. Our results confirm that these relations represent a substantial improvement over relations previously adopted in EPRI. They do also suggest that systematic bias still effects the

determination of M_{mi} , as discussed above. This bias, however, is not problematic for the b -value determinations of NCEER-91 because in the time windows used in that analysis, almost all the magnitudes are either M_{is} or M_{fa} with only few remaining M_{mi} (Figures 3-6, 3-7, and 3-8).

EPRI includes only events with magnitudes (of any kind) $M \geq 3$ or $M_I \geq III$, which corresponds approximately with a $M_{mi} \geq 3$. NCEER-91 is also limited to events with $M \geq 3$, however the sets of events in the two catalogs are significantly different because of the magnitude revisions. In general, an event in EPRI may have been eliminated from NCEER-91 because the magnitude is decreased below the threshold or because it is found to be a non-earthquake; an event not in EPRI may be in NCEER-91 because its magnitude is raised above the threshold or because it is a newly discovered earthquake (Table 3-II).

Generally, reports on earthquake effects in the eastern U.S. become more informative and complete with time. Some particularly significant transitions can, however, be identified. Before 1800, easily accessible archival material, such as newspapers, become scarce and search efforts tend to be poorly rewarded. The rapid growth of telegraph and railroads through the eastern U.S. during the early 1880's made fresh news from afar readily available and greatly stimulated the newspaper industry. Since that time, newspapers are a good and easily available source of macroseismic data. The Civil War corresponds with a major gap in newspaper reporting of events which are not directly related to the war. The compilation by Rockwood provides exceptionally good coverage from 1871 to 1886. The decision by the Federal Government in 1928 to begin collecting macroseismic data and publishing these data in U.S. Earthquakes marks a fundamental transition in the level of uniformity and completeness in the available macroseismic coverage. The archival reexamination by Armbruster and Seeber (1987 and unpublished data) provide exceptionally good coverage for southeastern U.S. seismicity in the 19th century up to 1889, before and after the 1886 Charleston, SC earthquake. A substantial fraction of the changes introduced by unpublished results of our incomplete re-examination are for events after 1928 and very few are from before 1800. Table 3-III compares the temporal distribution of magnitudes, both size and kind, in NCEER-91 and EPRI/S.

In cases where new archival data are used to revise source parameters, both positive and negative evidence on felt area is considered (i.e., if an event is described as being felt in a particular suburb of a town, implicitly it was not widely felt and it probably had a small felt area). Many of the revisions involve earthquakes with high maximum intensity relative

to the felt area, presumably shallow events. Some of these felt areas could be measured. Other felt areas could be constrained to be less than 100 km^2 , corresponding to $M_{fa} \leq 3$; these events were excluded from the catalog. Some events were determined to be explosions. New events were discovered and some were found to be felt over a much larger area than originally thought. Table 3-II lists quantitatively the differences between NCEER-91 and EPRI/S. Figure 3-4 highlights these differences in an epicentral map.

This effort has resulted in substantial changes to the catalog. That these changes represent an improvement can be inferred from Figures 3-6 – 3-8 where NCEER-91 exhibits the smoothest temporal seismicity distributions, without the large apparent seismicity highs in the late pre-instrumental period that characterize the EPRI catalogs, and particularly EPRI/V. Moreover, the spatial distribution of seismicity from NCEER-91 is more clustered than seismicity from EPRI (Figures 3-1 – 3-3). This difference is not a result of improved locations (that aspect of this study will be presented elsewhere), but stems from cleaning the catalog of non-earthquakes and, more importantly, tectonically insignificant small shallow earthquakes. The statistical characteristics of NCEER-91 and EPRI, however, are similar. This surprising outcome stems from the coincidental similarity between the number of earthquakes in the $M=3-4$ range taken out of the catalog, because they were reduced in magnitude or not earthquakes, and the number of new events brought into this range.

NCEER-91's event listing is keyed to EPRI's listing. It contains information on changes relative to EPRI and lists the sources which the changes are based on. The NCEER-91 catalog as described in this paper is available upon request from the National Center of Earthquake Engineering Research in Buffalo, N.Y..

SECTION 5
**RATES OF SEISMICITY AND COMPLETENESS FOR DIFFERENT
MAGNITUDES**

The magnitude distribution is obtained from the rate of occurrence of events in a set of magnitude windows. Under the assumption of stationarity, each of these windows can cover different time periods. Generally, the larger the magnitude, the rarer the earthquakes, but also, conveniently, the further back in time the catalog will be complete for those magnitudes. Thus, in the most effective use of the data, rates for each magnitude window are computed over completeness periods appropriate to that window (Stepp, 1972)

Figures 3-6, 3-7, and 3-8 display plots of the temporal distribution of seismicity in sets of magnitude ranges for several combinations of catalogs, areas, and magnitude windows; they will be referred to as magnitude-time plots. In Figure 3-6 seismicity rates cumulative over magnitude are given for steps of 0.25 magnitude units (M). In Figure 3-7 the steps are 0.2M and cumulative seismicity rates are given as a range between a minimum and a maximum value. In Figure 3-8 the steps are 0.25M, but the rates are incremental rather than cumulative. In a stationary seismic regime, all these plots should show a constant rate over the periods of completeness. These periods should start progressively earlier for larger magnitudes. Backward in time from the completeness periods, the apparent rates are expected to drop, reflecting incomplete reporting. This pattern can be perturbed by random scatter in a finite sample, by uneven coverage in the catalogs and by deviations from stationarity.

After assuming that the seismicity sampled is stationary, rates and completeness periods are estimated taking into consideration the shape of the magnitude-time plot as can be visually filtered out of the statistical perturbations. The overall tendency for completeness periods to increase with magnitude and our understanding of the history of seismology in the eastern U.S. play also an important role in this subjective judgement. We hope to provide the most reliable data in NCEER-91. Thus, rates and completeness estimates are based on our best overall judgement in the case of NCEER-91; in the case of EPRI, these estimates rely primarily on the shape of the magnitude-time plots.

Minimum and maximum values of both seismicity rates and completeness periods are estimated for the sets of plots in Figure 3-7 (cumulative; steps 0.20M). The range of estimated rates at a particular magnitude is carried through the least squares fitting

procedure by linearly distributing as many points between the extreme values as the cumulative number of events in the magnitude window. The rates shape the magnitude-distribution plots and are the main parameters recovered from the magnitude-time plots. Rates and completeness are interdependent, but often weakly so and rates can be picked from the magnitude-time plots with greater confidence. The completeness periods control the number of events determining a rate, but they primarily affect the statistical significance rather than the value of that rate. Thus, the parameter values were picked in the plots to reflect the most appropriate choice of maximum and minimum rates rather than completeness. Considering the large effect on the shape of magnitude-time plots stemming from problems in the catalogs which are discussed below, we feel that the real uncertainties in completeness are generally larger than the range in values given in Figure 3-7.

The shape of the magnitude-time plots for EPRI/V and, to a lesser extent, for EPRI/S deviate systematically from the expected flat distribution during completeness and a gradual drop off before that. Rather, they tend to show apparent rates of seismicity peaking in the early part of the historic period and then decreasing to lower levels in the late instrumental period (Figures 3-6, 3-7, and 3-8). In the low magnitude windows, the highest rates occur earlier than independently estimated completeness limits. This persistent broad early peak in the magnitude-distribution plots from EPRI coincides with the time during which many of the magnitudes in the EPRI catalogs are Mmi. We believe that these high levels of seismicity are not real and stem from a systematic over-estimation of Mmi's. Probable reasons for this systematic bias are discussed above. In light of the inherently ambiguous measure of magnitude given by Mmi's, rather than developing a new Mmi algorithm, NCEER-91 improves the intensity data base so that Mfa's can be substituted for most of the Mmi's during the periods of completeness (Table 3-II compares the number of each magnitude type in the two catalogs; the contribution to the rates from each of the three types of magnitude is shown in Figures 3-6, 3-7, and 3-8). Most of the Mfa's in NCEER-91 are lower than corresponding Mmi's in EPRI. The resulting improvement in the catalog is manifested by a substantial decrease or disappearance of the suspect pre-instrumental peak in seismicity (Figures 3-6, 3-7, and 3-8).

SECTION 6

MAGNITUDE DISTRIBUTION AND RECURRENCE TIMES

The rates of seismicity in discrete magnitude windows from the magnitude-time plots in Figures 3-6, 3-7, and 3-8 provide the basis for formulating magnitude distribution and recurrence relationships. This paper follows the standard procedure of assuming a linear distribution for the logarithm of the number of events, or the rate of seismicity (N), plotted versus magnitude (M) (Gutenberg and Richter, 1954). We calculate the constants b and a in the relationship $\text{Log}N = a - bM$ in a number of alternative standard procedures to investigate the procedural dependency of the results. The rates are determined from Figures 3-6, 3-7, and 3-8 for cumulative magnitude windows at steps of 0.25M (Figure 3-6) and 0.20M (Figure 3-7), and for incremental magnitude windows, at steps of 0.25M (Figure 3-8).

The linear fit to the rates from cumulative magnitude windows is obtained by the least-squares method (Figures 3-9 – 3-12) and the fit to the rates from incremental magnitude windows is accomplished by the maximum-likelihood method (Figure 3-13). Values of b and the recurrence rates of $M \geq 6$ (which are more significant for earthquake hazard than the rates for earthquakes of $M \geq 0$ -- the a -values) are calculated for sets of distribution plots (not shown) with low magnitude cut-offs at each of the magnitude steps. These parameters are displayed in Figures 3-9 – 3-13 as functions of low-magnitude cut off. Figure 3-14 displays the effect of high-magnitude cut-off on the maximum likelihood fit of incrementally counted magnitude distribution for the NCEER-91 catalog. Figure 3-15 shows the effect of different rates (or completeness periods) assigned to the magnitude window with the Charleston 1886 event as the sole entry (see Figures 3-6 – 3-8). The data considered in Figure 3-15 are from NCEER-91 east of the Appalachians. This event is assigned a magnitude $M=6.8$ following EPRI. This may be a realistic value for an M_b , but is probably an underestimate for an M_s (Seeber and Armbruster, 1981; Johnston and Hanks, 1991).

SECTION 7 DISCUSSION

7.1 NCEER vs EPRI: Data Comparison

The NCEER catalog is directly compared to the EPRI/S catalog in Figure 3-4 and Tables 3-II and 3-III. Magnitude differences in Figure 3-4 seem to be location dependent. This spatial dependency is partly ascribed to the contribution from special studies, such as the systematic re-examination of archival data for the southeastern U.S. during the early historic period up to and including the 1886-89 aftershock zone of the Charleston, S. C. earthquake (Seeber and Armbruster, 1987). An ongoing re-examination of archival data for the New York City area has also produced much new data and substantial changes to the catalog. The concentration of increased magnitudes in these areas (plusses) is primarily the result of new earthquakes (i.e., earthquakes which were either newly discovered, or were raised in magnitude from below to above the $M=3$ threshold). Following earlier catalogs, EPRI locates almost all of the 1886-89 aftershocks at the presumed main shock epicenter near Charleston and assigns them a magnitude consistent with their maximum intensity in that area. Since many of the aftershocks were actually centered at considerable distance from Charleston, their magnitude was generally underestimated.

In contrast, decreased NCEER-91 magnitudes (circles) predominate in New England and in eastern Tennessee. We have not carried out archival searches in these areas and changes derive mostly from published intensity data which we used to compute new felt-area magnitudes (M_{fa}). These magnitudes tend to be lower than corresponding maximum-intensity magnitudes (M_{mi}), as discussed above. A few large changes at remote locations represent large events in EPRI which turned out to be mistakes or non-earthquake sources. In summary, the overall pattern of changes is mixed: over-estimated M_{mi} are revised to lower M_{fa} removing many events from the catalog; conversely, a similar number of events is either newly discovered or their magnitude is revised upward with the addition of macroseismic data (see Table 3-II). These two-way changes tend to cancel each other out in terms of the magnitude distribution, as discussed below.

The magnitude-time plots in Figures 3-6, 3-7, and 3-8 offer the clearest indication that NCEER-91 represents an improvement. Both EPRI/S and EPRI/V, but particularly the latter, show rates of seismicity in the $M=3-4$ range that are consistently higher in the the

pre-instrumental period than in the instrumental period (these periods are here identified as the time when most of the magnitudes are determined from either macroseismic or instrumental data, respectively; the transition between these periods occurs substantially after the introduction of seismographs). The pre-instrumental high rate of seismicity is suspect and it is probably the result of systematically overestimated magnitudes, particularly Mmi's, as discussed above. This bias has been eliminated from NCEER-91 to the extent that Mfa's have been substituted for Mmi's. Table 3-III and Figures 3-6, 3-7, and 3-8 show that a large part of the Mmi's have indeed been converted to Mfa's and that the rates in NCEER-91 are fairly continuous through the transition from pre-instrumental to instrumental magnitudes.

The statistical results shown in Figures 3-9 to 3-15 are also indicative of data quality. Generally, differences between NCEER-91 and EPRI/S are small, while both these differ significantly from EPRI/V. Moreover, NCEER-91 and EPRI/S are reasonably well behaved, i.e., b-values are either constant or vary smoothly for different low-magnitude cut-offs. In contrast, b-values from EPRI/V are more variable and show significant changes with different low-magnitude cut-offs for a maximum likelihood fit. This instability in the b-values from EPRI/V is not surprising, given the instability of the rates from this catalog (Figures 3-6, 3-7, and 3-8) which is mostly a consequence of systematic bias in Mmi's (see above).

Given the substantial changes accomplished in NCEER-91 relative to EPRI (e.g., Tables 3-II and 3-III), the small difference between the magnitude distributions from NCEER-91 and EPRI/S is surprising. Two factors need to be mentioned. First, differences derived from applying different procedures are large (see below) and could mask differences derived from using different catalogs. Secondly, changes in the catalog include both additions and removals (Figure 3-4) which appear to add up to little change in terms of statistical characteristics sampled over large areas. At this stage of the revision, NCEER-91 improves the catalog, but does not affect significantly statistical parameters over the eastern U.S. or major subdivisions thereof. Both NCEER-91 and EPRI/S yield b-values clustered between 1.0 to 1.1 and repeat times for a magnitude $M \geq 6$ clustered between 100 to 300 years in either area, east and west of the Appalachian front. Parameter values for the Appalachian province (EAST) are similar to values for the central platform east of the New Madrid seismic zone (WEST). The total scatter in parameter values is substantially wider than the statistical confidence limits, but none of these variations, either in space or over the magnitude range, can be declared significant.

Locally, NCEER-91 provides much improved coverage and statistical parameters important for quantitative hazard assessment have been significantly affected (e.g., Armbruster and Seeber, 1987; Jacob et al., 1990). Finally, new constraints on earthquake locations are expected to improve the basis for characterizing source zones and the space-time distribution of seismicity. These topics, however, are beyond the scope of this paper.

7.2 Is the Magnitude Distribution Linear ?

Magnitude distributions are often assumed to be linear in a log-linear representation (e.g., Johnston and Nava, 1985; Bollinger et al., 1989; Veneziano and Van Dyck, 1985). This assumption is justifiable on the grounds that it produces parameters, the 'a' and 'b' values, that can be compared among seismic zones and can be monitored in time. Considering the large uncertainties and the scarcity of data for the eastern U.S., however, deviations from a linear distribution are possible, particularly in the portion of the magnitude range which is of concern for engineering applications and where data are often lacking.

Early workers (e.g., Utsu, 1971) did not expect magnitude distributions to be linear since data sets would generally contain events from unresolved, but distinct sources with different distributions. Moreover, a case can be made for a complex magnitude distribution for a single source zone in southern New York on the basis of structural geology data. Detailed geologic investigations in the Manhattan Prong have characterized a set of seismogenic faults in that intraplate source zone. These faults are segmented with segment dimensions structurally constrained in a narrow band (Seeber and Dawers, 1989). This observation led to the concept of a typical earthquake rupture whereby a region is characterized by a set of similarly segmented faults inferred to produce earthquakes of sizes within a preferred or 'typical' range (Hough and Seeber, 1991). A systematic segmentation of seismogenic faults implies that the size distribution of potential ruptures is scale-dependent. Thus, the magnitude distribution in the Manhattan Prong can be expected to be peaked at magnitudes corresponding to the typical ruptures for that zone. The concept of 'typical earthquake' is based primarily on structural geology data and is thought to be especially applicable to source zones in intraplate environments. This concept is parallel to, but distinct from the concept of 'characteristic earthquake' which is based on deformation data and is applicable to single faults in high strain-rate environments, generally associated with plate boundaries.

Even if magnitude distributions for individual source zones are generally non-linear, the distribution for the combined data from many zones with different structural characteristics may approximate a linear distribution. Thus, the linear behavior widely reported or assumed in magnitude distributions for the eastern U.S. (e.g., Bollinger et al., 1989; Johnston and Nava, 1985) may reflect on the paucity of data, since many source zones need to be combined in order to achieve statistical significance in the analysis. Whatever the reason for an apparently linear distribution, uniformitarianism would make this relationship useful for predictive purposes. The following argument, however, cautions on the extrapolation of magnitude distributions beyond the range of statistically significant observations. Extrapolations are generally required in order to reach the magnitude range of practical interest for earthquake hazard (typically $M \geq 5.5$).

Areas within which the size distribution of fault segments are likely to be clustered are expected to scale with the size of these segments. For instance, within the Manhattan Prong, an area a few tens of km across, brittle structure at the km scale is controlled by the intersection between a set of northwest striking and regularly spaced faults with a set of pre-existing regularly spaced isoclinal folds striking northeast. Interference between these structures results in a regular pattern of fault segmentation with typical lengths of 0.5-2.0 km (Dawers and Seeber, 1991). Since these faults are seismogenic, typical magnitudes are expected to be clustered in the range $M=4$ to 5 (Hough and Seeber, 1991). Such clustering is noted in the available earthquake data, but it may not yet be statistically significant.

The inferred clustering in the size of ruptures and corresponding magnitudes in the $M=4-5$ range pertains to the Manhattan Prong, an area about an order of magnitude larger than the dimension of these typical ruptures. The distribution of rupture sizes and magnitudes below the $M=4-5$ range may be similarly effected by systematics in the structure, but the size of the area where this systematics apply is expected to scale down accordingly. Generally, a given area may be occupied by a single family of fault segments of a particular size, but by more than one family of smaller segments. Thus, the magnitude distribution for the Manhattan Prong may appear to be linear for magnitudes $M < 4$, but may deviate from linearity above that.

The geologic observations outlined above suggest a model where the size distribution of seismogenic structures and earthquakes is clustered. It seems reasonable to assume that structures of horizontal dimensions clustered around R_i and producing earthquakes with

ruptures clustered around R_j^2 will be distributed over a source area A_j , much larger than R_j^2 , but proportional to it:

$$A_j/R_j^2 = \text{constant} \quad (1)$$

(An equivalent statement can be made in terms of a source volume). Thus, a given area will contain distinct source zones, each characterized by their typical rupture. The number of these ruptures and corresponding source zones decreases with increasing rupture size, up to the single typical rupture that is characteristic of the entire area considered. Below the corresponding typical magnitude, the magnitude distribution will reflect a superposition of distinct sources and may appear to be linear; at and above this magnitude, the distribution may reflect clustering in size and be far from linear. Are earthquake data sufficient to test this hypothesis?

If $\text{Log}(M_0) \sim 1.5M$ and $M_0 \sim R^3$ (i.e., stress drop is constant over the magnitude range; M_0 =seismic moment; R =rupture radius), then Gutenberg-Richter's law relating the number of earthquakes N and the Magnitude M , $\text{Log}N = a - bM$, can be expressed in terms of rupture radius,

$$\text{Log}N + b\text{Log}R^2 = \text{constant.}$$

If $b=1$ is acceptable as tending to give the best approximation to the magnitude distribution,

$$NR^2 = \text{constant} \quad (2)$$

By combining (1) and (2),

$$N_j A_j = \text{constant.}$$

Thus, the number of earthquakes N_j with typical rupture dimension R_j that characterize a source zone with spatial dimension A_j is independent of the size of this zone. By increasing the area considered in the analysis, the data available increase but the typical rupture size in that area would generally also increase, yielding fewer typical events per unit area. The chance of obtaining a statistically significant sample of a typical earthquake uncontaminated by different rupture families in the same size range improves by increasing the rate of

seismicity, but does not depend on the size of the area considered. We conclude that the available earthquake data may not be sufficient to resolve non-linearity at any source-area size and that the hypothesis of typical earthquakes leading to a punctuated magnitude distribution is not likely to be tested by earthquake data discussed here. Thus, the data do not rule out the possibility of a clustered magnitude distribution.

7.3 Magnitude Distributions in Cratonic Eastern U.S. (WEST) Versus Appalachians (EAST)

The demarcation line between the Appalachian province and the cratonic part of the eastern U.S. (labeled 'EAST' and 'WEST' in Figures 3-6 to 3-15, respectively) is shown in Figures 3-1 – 3-4. This line traces the foreland limit of the exposed allochthonous crystalline sheets of the Appalachian orogen and is named the crystalline front. Seeber and Armbruster (1988) showed that the depth-distribution of seismicity tends to be discontinuous across that limit; both the upper and lower limit of the seismicity tends to be deeper on the west than on the east side of that structural boundary. Seismicity on the cratonic (western) side originates in the Precambrian autochthonous rocks. Except for the Adirondacks where cratonic basement is exposed, these rocks lie below the Appalachian foredeep and platform sedimentary rocks which appear to be generally aseismic. On the eastern side of the line, seismicity is thought to originate from the allochthonous crystalline slab which was thrust above the North American basement and platform rocks during the Appalachian orogeny. This slab tends to be exposed at the surface, is thin near the forward edge and increases in thickness toward the continental margin (e.g., Cook et al., 1981; Brown et al., 1983). Accordingly, Appalachian seismicity is characterized by shallow small earthquakes near the front and by an increase in maximum depth and maximum size eastward.

The contrast in tectonic setting and depth distribution for seismogenesis on either side of the Appalachian front suggests the hypothesis that this boundary separates source zones with distinct magnitude distributions. Figures 3-9, 3-10, 3-11, and 3-12 show systematic differences in the statistical parameters from all three catalogs for these two zones. These differences are most apparent in the shape of the plots of b-value versus low-magnitude limit. As detailed in the following discussion, however, these differences may all derive from procedural and statistical factors and are not significant.

In accordance with the assumption of linearity, b-values from WEST are relatively stable; b-values for NCEER WEST hover about 1.0 (± 0.07) for low-magnitude limits between $M=3.0$ and $M=4.5$ (Figures 3-9, 3-10, 3-11, and 3-12). In contrast, corresponding plots for EAST show greater procedure-dependent variations and a tendency for b-values to drop with rising low-magnitude limit. At a lower limit of $M=3$, b-values for NCEER EAST are also near 1, but they range from 0.88 to 1.11. At a lower limit of $M=4.5$, the b-values have dropped considerably with most of the values between 1.0 and 0.6. This drop is about twice as large for $M \geq 5.0$.

In the three least-squares fits of EAST with steps of $0.25M$ (Figures 3-9A, 3-10A, and 3-11A) the drop in b-value from $M \geq 3.0$ to $M \geq 5.0$ is similar, about 0.3, but the overall level of these plots differ. The difference in these b-value plots seems to be related to the weighting procedure: the more the fitting procedure weights points at low relative to high magnitudes, the higher are the overall b-values. The drop in b-value with rising low magnitude limit indicates an upwardly concave magnitude distribution. The similarity of the drop in b-value among the different weighting procedures suggest that the non-linearity is primarily the result of unusually high rates in the magnitude range $M \geq 5$. If, instead, the non-linearity were the result of high values at low magnitudes, the drop in b-value would be systematically larger the more weight was given to points near the low-magnitude limit.

The drop in b-value with rising low-magnitude limit is very sensitive to the rate assigned to the 1886, $M=6.8$ Charleston earthquake. The drop in the maximum likelihood fit disappears when the repeat time for that event is an order of magnitude larger than historic time (Figure 3-15). This surprising result is ascribed to the leverage exerted on the b-value by a data point far from the expected value (Weichert, 1980).

The drop in b-value with rising low-magnitude limit is large, characterizes all three catalogs and is detected by each of the procedures adopted (Figures 3-9 – 3-13), but it is not significant in all of them. The maximum-likelihood fits of NCEER-91 and EPRI/S show no significant change with shifting low-magnitude limits (Figures 3-13A,B). This result would not be effected by conceivable repeat-times variations for the Charleston-size events in EAST (Figure 3-15C), or by changes in the expected upper limit to magnitudes in WEST. Only EPRI/V shows deviations from a linear magnitude distribution which reach beyond the standard confidence limits. The erratic behavior of EPRI/V can be ascribed to noise rather than a systematic change in b-value (see above).

The rate of one in 200 years, the length of well-documented historic time, is assigned to magnitude windows in EAST containing the $M=6.8$ Charleston 1886 event (Figures 3-6, 3-7, and 3-8). This is a reasonable choice if these earthquakes could occur anywhere in EAST; such an earthquake in remote parts of the Appalachians in the 1700's might have been miss-identified as a much smaller event near a settlement. If, at the opposite extreme, the Charleston area is the only source of this kind of event in EAST, then paleoseismic data from that area (e.g., Obermeier et al., 1985) would suggest an order of magnitude longer repeat time, or about 2,000 years, for this event. Thus, according to the historic and paleoseismic data, 200 and 2000 years are the minimum and maximum values of repeat time for the Charleston earthquake. These limits are consistent with a linear magnitude distribution according to the maximum-likelihood fit of NCEER-91 (Figure 3-15C). In this fit, as the lower magnitude limit rises, b-value drops for a repeat time of 200 years, rises for an infinite repeat time (i.e., no event) and remains constant for a repeat time of 2000 years (Figure 3-15C).

7.4 Effects on the Statistical Parameters (a- and b-Values) by Differences in Curve-Fitting and other Procedures

The most surprising result of this study is the strong dependence of the statistical results on some of the procedural factors. Our statistical analysis involves basically two steps, first determining rates for a set of magnitude windows, and then fitting the results with a line in a log-linear space. The determination of the rates (Figures 3-6 – 3-8) is a relatively subjective procedure involving an eyeball fit of a constant rate to a temporal distribution, the magnitude-time plot, that reflects both a scatter and a decay of the apparent rate beyond the period of completeness. The completeness periods are chosen according to the shape of the magnitude-time plots, but are also chosen in light of the history of seismological reporting and recording in the eastern U.S. (see above). Linear fits to the magnitude-distributions of rates are accomplished by one of several algorithms available in the literature, in this case either by a maximum likelihood (e.g., Bollinger et al., 1989) or by a least squares method (e.g., Johnston and Nava, 1985). These as well as other more detailed aspects of the procedure are found to have a large effect on the results. In many cases, differences between alternative curve-fitting procedures exceeded standard confidence limits obtained by some of the procedures. Thus, realistic confidence limits are obtained by comparing results from alternative procedures, rather than by relying exclusively on statistical error bars.

Subjective Picks of Seismicity Rates

Three sets of magnitude-dependent seismicity rates for different magnitude thresholds are derived from each of the catalogs NCEER-91, EPRI/S and EPRI/V. In one case magnitude windows are cumulative (all events from the magnitude characterizing the window to the highest magnitude) with steps of 0.25M (Figure 3-6), in another they are also cumulative but with steps of 0.20M (Figure 7). Differences in statistical parameters between Figures 3-9 and 3-12 are solely the result of these two rate estimates. In the third rate estimate, magnitude windows are incremental with steps of 0.25M (Figure 3-8). The effect of shifting from incremental to cumulative magnitude windows can be seen by comparing Figures 3-9 and 3-13. In this comparison, however, the shift from a least squares to a maximum likelihood fit also plays a role. In both these comparisons, the overall shape of the b-value distributions are encouragingly similar, but systematic differences can be seen, particularly for the curves in Figures 3-9 and 3-12.

The plots of statistical parameters (recurrence times and b-values) resulting from the two cumulative rate estimates with magnitude windows of 0.25M and 0.20M, respectively (Figures 3-9 and 3-12), are surprisingly different, particularly the plots for EAST. Both these plots (Figures 3-9A and 3-12A) are characterized by a drastic drop in b-value with increasing low-magnitude cut-off, but the plot for steps of 0.20M is steeper and lower than the plot for 0.25M. This difference is ascribed in part to systematically higher rates assigned to the 0.20M than to the 0.25M magnitude windows at high magnitude. The set of rates for 0.20M are smeared between extreme values that diverge at large magnitudes where the data are scanty (Figure 3-7). The drop in b-value is also the result of high rates assigned to the Charleston-like events; this drop decreases substantially for both the 0.20M and the 0.25M rates by removing the 1886 Charleston earthquake (Figure 3-15).

Least Squares vs Maximum Likelihood

Generally, b-values from maximum likelihood fits (Figure 3-13) show larger variations as a function of magnitude threshold than least squares fits (Figure 3-9). This characteristic is expected from the higher weight given to rates near the low-magnitude limit in the maximum likelihood fit than in the least squares fit. The same effect combined with a relatively high level of $M \geq 5$ events in both EAST and WEST causes repeat times for $M \geq 6$

events to be higher for maximum likelihood than for least squares fits. The over-abundance of $M \geq 5$ in both EAST and WEST is indicated by abnormally low b-values for $M \geq 5$.

One of the obvious differences between the least squares and the maximum likelihood methods is in the size of the confidence limits. While standard confidence limits from maximum likelihood encompass most of the procedure-dependent differences in results, standard limits from least squares are often much less than these differences, if the points are weighted to reflect the amount of contributing data (Figures 3-9 and 3-10). If the points are all weighted uniformly (Figure 3-11), error bars are similar to procedure-related changes, but the results may be misleading because the fit weights the points incorrectly. The huge effect of varying the frequency of the Charleston event on the results from the least squares fit with all points weighted the same amount (Figure 3-15) is symptomatic of the behavior of this fit.

In summary, our results confirm previous assertions that maximum likelihood is the desirable fitting procedure (e.g., Bollinger et al., 1989), provided the linearity of the magnitude distribution is not questioned. If systematic deviations from a linear magnitude distribution are considered possible (see above), than a procedure that balances amount of data with the length of extrapolation seems desirable. The goal for hazard purposes may be to estimate the recurrence time of $M \geq 6$; if linearity is questioned, the rates of $M \geq 5$ should play a greater role in determining this recurrence than the rates of $M \geq 3$ and a least-squares approach may be preferable.

The results indicate a strong dependence on procedural factors, such as the choice of algorithm to calculate M_{mi} , of magnitude steps, and of the method used to accomplish the straight line fit to the data. Thus, the significance of deviations from a linear magnitude distribution and differences in b-value between different areas remain unclear. The b-value and recurrence analysis is useful because it offers a means of comparison of results using different procedures on different data sets and because it provides lower bounds on the uncertainties that can be used for hazard estimates. These uncertainties are found to be large even if the linear extrapolation of the magnitude distribution to large magnitudes are justified. These uncertainties would be substantially larger if the magnitude distribution is not linear at large magnitudes (see above). Further uncertainties would arise from violations of the stationarity assumption.

SECTION 8 CONCLUSIONS

1. The NCEER-91 earthquake catalog represents a substantial improvement over the EPRI catalog in terms of magnitude constraints. Many of the maximum-intensity magnitudes were upgraded to felt-area magnitudes and found to be systematically over-estimated. The main manifestation of this improvement is more uniform rates of seismicity at all magnitude levels over the historic period and a sharper definition of the seismic zones by eliminating epicenters of insignificant events and non-earthquakes.
2. The EPRI/V catalog with magnitudes according to Veneziano and Van Dyck (1985) produces b-values and repeat times which are much less stable than either EPRI/S, the EPRI catalog with magnitudes according to Sibol et al. (1987), or NCEER-91. The erratic behavior of b-values from EPRI/V is primarily ascribed to systematic bias in assigning magnitudes from maximum intensity. The b-values and repeat times from NCEER-91 and EPRI/S are insignificantly different, although about one third of the magnitudes determining these parameters have been modified. This inability to resolve differences reflects on both large uncertainties and on the null overall effect on statistical parameters of upward magnitude revisions or additions and downward revisions or deletions.
3. Procedural factors can greatly affect results. They include the determination of seismicity rates, periods of completeness and curve-fitting magnitude distributions. The most subjective step in the analysis, the choice of rates and completeness periods, and the most debatable step, the choice of a rate for the Charleston, S.C. 1886 M=6.8 (EPRI) event, introduce large uncertainties in the results. Differences arising from alternative fitting methods, i.e., least squares fit versus maximum likelihood fit, can also be large, particularly if the magnitude distribution deviates substantially from linearity. Error estimates from maximum likelihood fits seem commensurate with procedure-dependent variations in the results. Error estimates from least squares fits are unrealistically small, particularly when rates are weighted according to the number of events contributing to those values.
4. In both areas studied, east and west of the Appalachian front, b-values from maximum likelihood fits are concentrated at $b=1.05\pm 0.05$; repeat times of $M\geq 6$ are concentrated at 200 ± 100 years. Both sets of values obtained from least-squares fits tend to be somewhat lower. Generally, b-values obtained from the data truncated at different low-magnitude

limits are different, but insignificantly so and any deviation from a linear magnitude distribution in a log-linear representation remain unresolved. Differences in statistical parameters between the areas east and west of the Appalachian front are also insignificant.

5. Assuming a linear magnitude distribution, the Charleston, S.C. 1886 earthquake ($M=6.8$; EPRI) fits better the magnitude distribution east of the Appalachian front with a repeat time of 1000-2000 years. This repeat time is almost an order of magnitude longer than the historic record (which happens to include an event of this magnitude), but is similar to the recurrence time for large liquefaction events in the Coastal Plain of South Carolina.

6. Arguments based on structural geology data and detailed earthquake data from the Manhattan Prong suggest the possibility of a typical magnitude in the range $M=4-5$, where the magnitudes of the largest historic earthquakes in this region are clustered. In this case, the magnitude distribution would be complex and the log-linear extrapolation of this distribution to magnitudes larger than the observed ones may lead to unrealistic results.

SECTION 9
REFERENCES

- Algermissen, S. T., D. M. Perkins, P. C. Thenhaus, S. L. Hanson and B. L. Bender (1990), Probabilistic earthquake acceleration and velocity maps for the United States and Puerto Rico. 1:7,500,000; USGS Map MF-2120
- Armbruster, J.G., and L. Seeber, Seismicity 1986-89 in the Southeastern United States: The aftershock sequence of the Charleston, S.C. earthquake, U.S. Nuclear Regulatory Commission, Washington, D.C., NUREG/CR-4851, 153 pp., 1987.
- Armbruster, J.G., (b) and L. Seeber, Seismicity and Seismic Zonation along the Appalachians and the Atlantic Seaboard from Intensity Data, in Technical Report NCEER-87-0025, p. 163-177, 1987.
- Barstow, N.L., Brill, K.G., Jr., Nuttli, O.W. and Pomeroy, P.W., An approach to seismic zonation for siting nuclear electric power generating facilities in the eastern United States, NUREG CR1577, 143 pp., 1981.
- Bernreuter, D.L., J.B. Savy, R.W. Mensing, D.H. Chung (1984). Seismic hazard characterization of the eastern United States; Methodology and interim results for ten sites. Lawrence Livermore National Laboratory, NUREG/CR-3756, UCRL-53527, 435p.
- Bollinger, G.A. and M.G. Hopper, The Earthquake History of Virginia 1900-1970, Virginia Polytechnic Institute and State University, 85 p., 1972.
- Bollinger, G.A., F.C. Davison, Jr., M.S. Sibol, and J.B. Birch, Magnitude recurrence relations for the southeastern United States and its subdivisions, J. Geophys. Res., 94, 2857-2873, 1989.
- Boston Edison Company, Historical seismicity of New England, U.S. Nuclear Regulatory Commission, Docket No. 50-471, 641 pp., 1976.
- Brown, L., C. Ando, S. Klemperer, J. Oliver, S. Kaufman, B. Czuchra, T. Walsh, and Y.W. Isachsen, Adirondack-Appalachian crustal structure; The COCORP Northeast Transverse: Geol. Soc. Amer. Bull., v. 94, p. 1173-1184, 1983.
- Coffman, J.L. and Angel, C., Summary of earthquake intensity file, Key to geophysical records documentation No. 19, NOAA, Boulder, Co., 1983.
- Cook, F.A., L.D. Brown, S. Kaufman, J.E. Oliver and T.A. Petersen, COCORP seismic profiling of the Appalachian orogen beneath the crustal plane of Georgia: Geol. Soc. Amer. Bull., v. 92, p. 738-748, 1981.
- Coppersmith, K.J., A.C. Johnston and W.J. Arabasz, Estimating maximum earthquakes in the central and eastern United States: A Progress Report, in Technical Report NCEER-87-0025, p. 217-232, 1987.
- Dawers, N. and L. Seeber, Intraplate faults revealed in crystalline bedrock in the 1983 Goodnow and 1985 Ardsley epicentral areas, New York, Tectonophysics, v. 186, p. 115-131, 1991.
- Dewey, J.W. and D.W. Gordon, Map showing recomputed hypocenters of earthquakes in the eastern and central United States and Canada, 1925-1980: U.S. Geological Survey Miscellaneous Field Studies Map. MF1699, scale 1:2,500,000, 1984.

- Electric Power Research Institute, Catalog of Central and eastern North American earthquakes to 1985, written communication, J.C. Stepp, 1987.
- Electric Power Research Institute, Seismicity Owners Group (1985). Seismic hazards methodology for nuclear facilities in the eastern United States. v. 1: Development and application of methodology, EPRI/SOG Draft 85-1.
- Gutenberg, B. and C.F. Richter, Seismicity of the Earth and associated phenomena, Princeton University Press, Princeton, N.J., 273 p., 1954.
- Hopper, M.G., and G.A. Bollinger, The earthquake history of Virginia 1774 to 1900, Virginia Polytechnic Institute and State University, 87 p., 1971.
- Hough, S.E. and L. Seeber, Seismological constraints on source properties of the m=4.0 Ardsley, New York earthquake: A characteristic rupture? Submitted to Jour. Geophys. Res., 1991.
- Jacob, K.H., J.-C. Gariel, J. Armbruster, S. Hough, P. Friberg and M. Tuttle, Site-specific ground motion estimates for New York City, Proceedings, Fourth U.S. Conference on Earthquake Engineering, Palm Springs, Ca., v. 1, p. 587-596, published by EERI, ~1990.
- Johnston, A. C. and T. C. Hanks, Similarities and differences in earthquake felt and damage areas between eastern and western North America (abstract), EOS, v. 72, N. 17, p. 195, 1991.
- Johnston, A.C. and S.J. Nava, Recurrence rates and probability estimates for the New Madrid seismic zone, J. Geophys. Res., v. 90, 6737-6753, 1985.
- Leffler, L. M. and S. G. Wesnowsky, Paleoseismic study of the New Madrid seismic zone by using liquefaction features as seismic indicators (abstract), EOS v.72, N. 17, p. 196, 1991.
- Obermeier, S.F., G.S. Gohn, R.E. Weems, R.L. Gelinas and M. Rubin, Geologic evidence for recurrent moderate to large earthquakes near Charleston, South Carolina, Science, 227, 408-411, 1985.
- Reinbold, D.J. and A.C. Johnston, Historical seismicity in the southern Appalachian seismic zone, USGS Open File Report 87-433, 1987.
- Richter, C.F., An instrumental earthquake scale, Bull. Seismol. Soc. Amer., v. 25, p. 1-32, 1935.
- Rockwood, C.G., Jr., Notes on Recent American Earthquakes, Amer. Jour. of Sciences, v. 4, p. 1-4, 1872 thru v. 32, p. 7-19, 1986.
- Seeber, L. and J.G. Armbruster, The 1886 Charleston, South Carolina, earthquake and the Appalachian detachment: Jour. Geophys. Res., v. 86, p. 7874-7894, 1981.
- Seeber, L. and J.G. Armbruster, The 1886-1889 Aftershocks of the Charleston, South Carolina, Earthquake: A Widespread Burst of Seismicity, Jour. Geophys. Res., v. 92, p. 2663-2696, 1987.
- Seeber, L. and J.G. Armbruster, Seismicity along the Atlantic seaboard of the U.S.: Intraplate neotectonics and earthquake hazard, in The Atlantic Continental Margin, U.S., R.E. Sheridan & J.A. Grow, Eds. Geological Society of America, The Geology of North America, Vol. I-2: 565-582, 1988.

- Seeber, L. and N. Dawers, Characterization of an Intraplate Seismogenic Fault in the Manhattan Prong, Westchester Co., N.Y., Seismol. Res. Lett., v. 60, p. 71-78, 1989.
- Sibol, M.S., G.A. Bollinger and J.B. Birch, Estimation of magnitudes in central and eastern North America using intensity and felt area, Bull. Seismol. Soc. Amer., v. 77, p. 1635-1654, 1987.
- Stepp, J.C., Analysis of the completeness of the earthquake hazard sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard, Proc. Internat. Conf. on Microzonation for Safer Constr. Res. Appl., 2, Seattle, Wa., p. 897-909, 1972.
- Stover, C.G., Catalog of Central and Eastern United States earthquakes, written communication, 1982.
- Taber, S., Seismic activity in the Atlantic Coastal Plain near Charleston, South Carolina, Bull. Seismol. Soc. Amer., v. 4, p. 108-160, 1914.
- Taber, S., Earthquakes in South Carolina during 1914, Bull. Seismol. Soc. Amer., v. 5, p. 96-99, 1915.
- United States Earthquakes, U.S. Department of Commerce, U.S.G.S., Annual 1928-1985.
- Utsu, T., Aftershocks and earthquake statistics, Jour. Fac. Sci. Hokkaido Univ. Ser. VII, Geophysics, III, 379-441, 1971.
- Veneziano, D. and J. Van Dyck, Analysis of earthquake catalogs for incompleteness and recurrence rates, in Seismic Hazard Methodology for Nuclear Facilities in the Eastern United States (April 30, 1985), v. 2, Appendix 6, Appendix A, p. A220-A297, Electric Power Research Institute, Palo Alto, Calif., 1985.
- Visvanathan, T.R., Earthquakes in South Carolina, 1698-1975, Bulletin 40, South Carolina Geological Survey, 61 p., 1980.
- Weichert, D.H., Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, Bull. Seismol. Soc. Amer., v. 70, p. 1337-1346, 1980.

