Geof. Int., Vol. 28-1, 1989, pp. 103-119

A STUDY OF THE HOMOGENEITY OF THE NOAA EARTHQUAKE DATA FILE IN THE MID-AMERICA REGION BY THE MAGNITUDE SIGNATURE TECHNIQUE

F. R. ZUÑIGA* **

(Received: May 26, 1988)
(Accepted: 22 August, 1988)

RESUMEN

El archivo de datos de sismicidad compilado por la N.O.A.A. (National Oceanic and Atmospheric Administration) de los Estados Unidos conocido como "Earthquake Data File" o EDF, contiene información que puede ser de gran importancia cuando se va a estudiar la sismicidad de alguna región del mundo en particular. Sin embargo, la homogeneidad de este catálogo, el cual básicamente está formado por datos recabados por P.D.E. (Preliminary Determination of Epicenters), depende de las características de las estaciones y agencias que proporcionaron los datos originalmente.

El motivo de estudio del presente trabajo es el análisis de la homogeneidad del catálogo de la NOAA, específicamente para la región de Meso-América. Se emplea una técnica recientemente propuesta que permite simular cambios en los eventos reportados con respecto al tiempo. Estos cambios pueden presentarse tanto en el número como en las características de los eventos que constituyen el catálogo. El método empleado se describe brevemente. Los tiempos de ocurrencia de los cambios mencionados, obtenidos por la técnica de la "firma de la magnitud", y sus probables causas, se determinan, y se proponen correcciones para eliminar las variaciones que pueden afectar la definición de la sismicidad de fondo.

Los tiempos para los que se encontraron cambios, con suficiente significancia estadística, son: finales de septiembre de 1965; mediados de octubre de 1967; principios de octubre de 1969; finales de mayo de 1972, y finales de julio de 1979. Las correcciones propuestas para los eventos listados en el catálogo son: la magnitud (m_b) de los eventos de julio de 1964 a septiembre de 1965 debe ser disminuída en 0.1 unidades; los eventos de septiembre de 1965 a mayo de 1972 deben ser corregidos añadiendo 0.1 unidades a las magnitudes listadas, y por último, los eventos posteriores a mayo de 1972 no deben ser corregidos.

* *University of Colorado, Department of Geological Sciences and Cooperative Institute for Research in Environmental Sciences, Campus Box 449, Boulder, Colorado.*

** *Now at Instituto de Geofísica, U.N.A.M., Cd. Universitaria, México 04510, D. F., MEXIÇO.*

Con estos resultados se puede evitar la probable inclusión de cambios artificiales de sismicidad que pudieran de otra manera ser confundidos con variaciones reales. Así, el estudio de la sismicidad con fines de predicción en regiones de Meso-América podrá ser llevado a cabo con resultados más precisos.

ABSTRACT

The Earthquake Data File (EDF) compiled by the National Oceanic and Atmospheric Administration (NOAA) of the U. S. contains information that can be of great importance when studying seismicity of a certain region of the world. However, the homogeneity of this catalog, which comprises mostly data gathered by PDE (Preliminary Determination of Epicenters), depends on the characteristics of the stations and agencies which provided the data in the first place.

In this study, the homogeneity of the NOAA catalog, with specific focus on the Mid-America region, is investigated by means of a technique recently proposed which simulates changes in the events reported with respect to time. These changes may account for the observed variations in number and characteristics of the events listed in the catalog. The method is briefly described. Times of occurrence of such changes, as given by the magnitude signature method, are determined as well as their possible causes, and proposed corrections are given.

The times for which changes were found to be statistically significant are: end of September 1965; mid-October 1967; early October 1969, end of May 1972 and end of July 1979. The corrections proposed for events in the Mid-America region listed in the NOAA catalog are: events from July 1964 to September 1965 should be shifted −0.1 magnitude (m_b) units; events from September 1965 to May 1972 should be changed by + 0.1 magnitude units and events after May 1972 ought to be left without change.

With this results the probable inclusion of artificial changes in seismicity that could otherwise be taken as real seismicity variations can hopefully be avoided. Thus, the study of seismicity in the Mid-America region for earthquake prediction purposes can be undertaken with more accurate results.

INTRODUCTION

The hypothesis that the occurrence of large earthquakes may be predicted by studying background seismicity rates and detecting anomalous changes has been used by many authors in the search for a reliable mainshock precursor (*e.g.* Othake *et al.*, 1977; Habermann, 1981; McNally, 1981; Kanamori, 1981; Wyss *et al.*, 1981, 1984). However, before attempting to use seismicity rate anomalies as possible precursors, apparent rate changes have to be carefully examined to determine whether they are in fact inherent to the natural seismic activity.

Other possible causes for changes in seismicity rates which are not due to natural variations include, but are not limited to, the closing or opening of seismic stations, lack of reporting of particular magnitude events, or changes in the calculated magnitudes of specific events.

Habermann (1987) has proposed that the recognition of each one of these situations can be done in a quantitative manner by statistically comparing the differences between the means in the seismicity rates of two time periods. Furthermore, these comparisons are performed for separate magnitude bands and the results displayed so that one is able to detect not only the time for a significant change but also the magnitudes affected by it. By analyzing the information in the magnitude domain Habermann (1987) has also proposed that particular changes show apparently unique characteristics, which he terms "magnitude signatures", which then may help to discriminate between anomalies and find appropriate corrections to homogenize the data.

In this study we employ Habermann's magnitude signature technique in order to determine seismicity rate changes, present in NOAA's earthquake data file and for the Mid-America region (lat. 5°N to 33°N; long. 80°W to 120°W), which are related to monitoring changes. By detecting such changes and evaluating their amount further studies based on this catalog, concerning seismicity rate anomalies as possible precursors to earthquakes, will be more accessible and will have more weight.

BRIEF DESCRIPTION OF THE MAGNITUDE SIGNATURE METHOD

The basis for the magnitude signature method is the statistical "z-test" (*e.g.* Meyer, 1975). This test is the most general of the statistical tests for evaluating the difference between two means. The means correspond to rates for two consecutive time periods and a particular magnitude band. In order to make the comparison between the two rates, the means M_1 and M_2 and their respective standard deviations, S_1 and S_2, and number of samples for each period, N_1 and N_2, are determined. These values can then be used to calculate a "z" score with the formula:

$$z = \frac{(M_1 - M_2)}{\left[\frac{S_1^2}{N_1} - \frac{S_1^2}{N_2}\right]^{1/2}}$$

The resulting z value can then be used to determine the confidence level of a change (*i.e.* $z = 1.64$ indicates a significance of 90 per cent, $z = 1.96$ is 95 per cent significance, and $z = 2.57$ is 99 per cent significance).

Once a change has been detected and its significance determined, an examination in the magnitude domain will provide means to identify a magnitude cutoff which excludes events affected by possible man-made changes in seismicity rates.

Magnitude signatures show the significance of an observed seismicity change as a function of upper or lower magnitude cutoffs. The magnitude signature plots have a vertical axis which indicates the z value obtained when comparing rates during two time periods. The horizontal axis of the plot shows the magnitude bands examined. A positive z value indicates a rate decrease while negative z values are obtained for rate increases (Fig. 1).

The magnitude axis is divided in two sections. Thus, the left side of the plot corresponds to magnitudes bounded by an upper limit (*e.g.* a point signaling a magnitude of 2.0 on the left side indicates those events with a magnitude lower than 2.0). Conversely, the right side of the plot shows those magnitudes bounded by a lower magnitude cutoff. Therefore, if a point indicates 2.0, for example, on the right side of the plot, then it is to be interpreted as the z value obtained for a change in events whose magnitude is larger or equal than 2.0.

The divisions on both axis generate four quadrants which provide information about the whole data set. Figure 1 shows the interpretation given to points lying on each one of the four quadrants as explained above.

CHARACTERISTIC MAGNITUDE SIGNATURES

As already mentioned, there are certain characteristics of the magnitude signature plots which indicate the particular type of change which generated them. They may be briefly summarized as follows:

Detection changes

A number of studies (*e.g.* Habermann, 1982; Wyss, Habermann and Heiniger, 1983) have demonstrated that the closing of seismic station produces a decrease in the detection of events, and that this effect shows in teleseismic data sets. A detection increase, on the other hand, can also be observed following the installation of new stations or the opening of new networks.

In both cases, however, events which are likely to be affected are those smaller than a certain magnitude. Events which generate more energy are obviously easier to detect by any existing stations than small magnitude events. It has been observed that teleseismic data sets are dominated by detection decreases (Habermann, 1987).

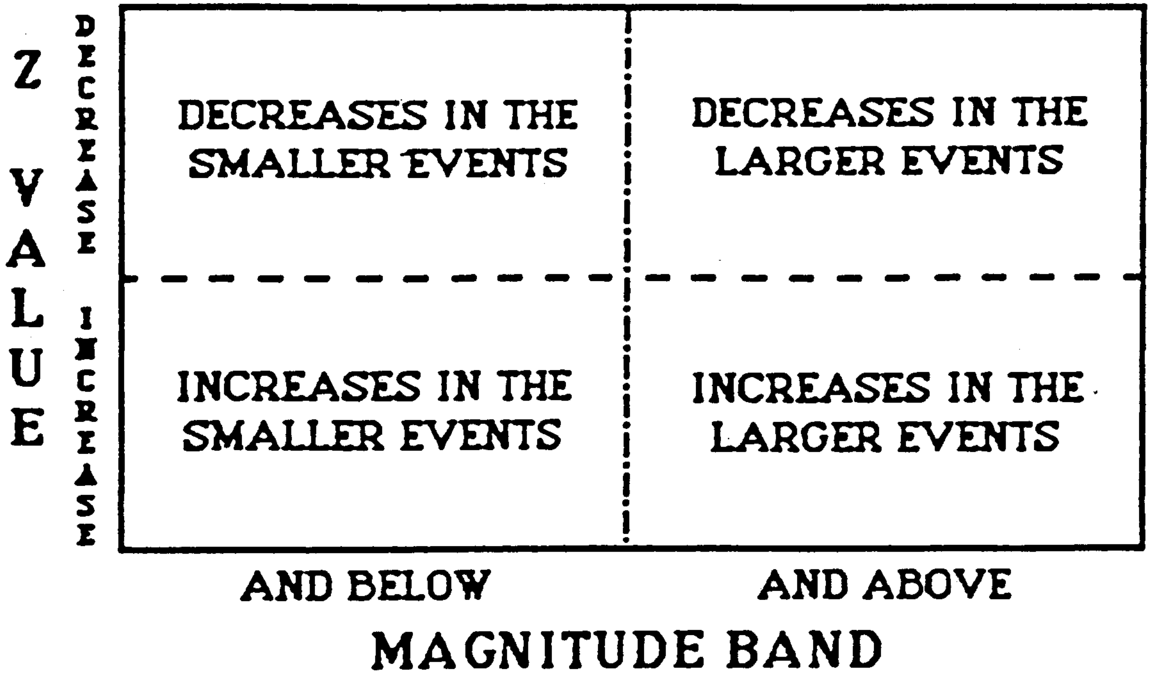


Fig. 1. Magnitude signature plots show the significance of an observed seismicity change as a function of magnitude cutoff. The vertical axis of a magnitude signature plot is the z-value which results from comparing the rates during two periods. The horizontal axis shows the magnitude bands being considered. This figure indicates schematically the meaning of the location of the points in each of the four quadrants.

Figure 2a shows the characteristic magnitude signatures of both detection decreases and increases. It can be seen that this type of change is marked by strong increases or decreases in the smaller events and no changes in the larger events.

The effects of the detection changes can be removed by eliminating events smaller than a certain magnitude cutoff. Such cutoff can be determined from the magnitude signature by the magnitude at which either a lowest cutoff appears to define a platform on the right side of the plot; or the magnitude just above the peak (or through), on the left side of the plot.

Reporting changes

This type of change occurs when events are detected and listed in the catalogs but magnitudes are not assigned. This change is usually observed in preliminary sections of local catalogs and in teleseismic catalogs. Events which should be in the larger-magnitude sets end up in the sets of smaller magnitude events because of their zero magnitude. This effect has been recognized before in a study of seismicity of the Imperial Valley (Habermann and Wyss, 1984).

Since reporting changes cause a decrease in the sets above any magnitude cutoff and an increase in the smaller magnitude sets, a change in z-value across the plot is produced which allows identification of this particular type of change (Fig. 2b).

If one is looking for real changes in seismicity rate, then the events with no magnitudes must be eliminated or a magnitude cutoff employed (also eliminating zero magnitude events from consideration).

Magnitude shifts

This type of change is similar to a reporting change. The difference is that the assigned magnitudes shift by some small amount (commonly 0.1 to 0.5 magnitude units) instead of the zero assigned when event cataloging is considered preliminary. Magnitude shifts cause changes of different signs depending on the direction of the shift and the type of magnitude band being studied.

Particularly important for seismicity quiescence studies are magnitude decreases because they cause apparent decreases in the calculated rate of activity above some

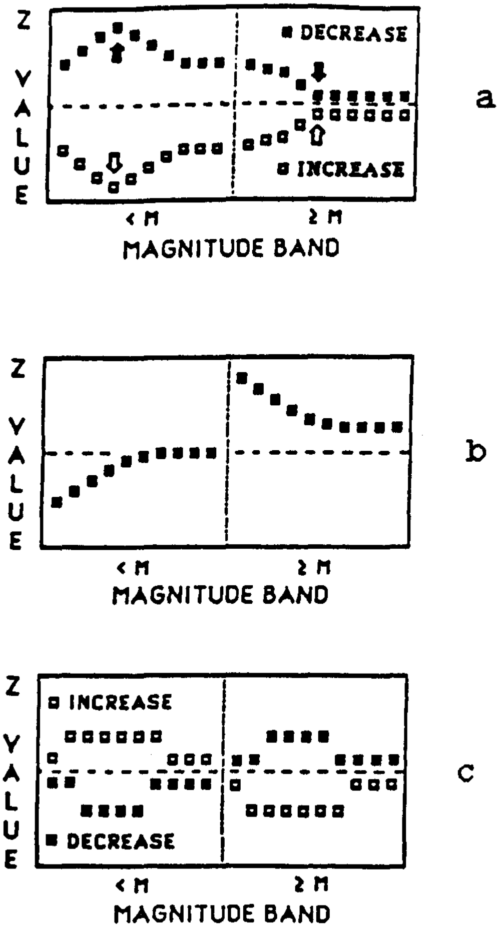


Fig.2.

A. Schematic representation of a magnitude signature which results from detection changes. Detection changes are characterized by increases or decreases in the rates of smaller events and lack of changes in the larger events. The arrows indicate the two ways of picking the correct magnitude cutoff from the plot. One is the lowest cutoff in the platform on the right side of the plot. The other is the peak on the left side of the plot.

B. Schematic representation of a magnitude signature corresponding to a reporting change. This type of change shows decreases in the larger events and increases in the smaller (events with no magnitudes are included in these data sets).

C. A schematic magnitude signature which results from magnitude changes. In these cases, if the magnitudes are lowered, the magnitude bands which include the larger events show rate decreases while the magnitude bands which include smaller events show increases. If the magnitudes increase, the opposite effect is observed.

cutoff. Thus, a change of this type could be taken as a precursor (since quiescence can sometimes be present before the occurrence of a mainshock) if the magnitude shift is not detected and accounted for.

Figure 2c shows schematic magnitude signatures for shifts which are restricted to certain sized events. As stated before, the primary characteristic of such changes is the occurrence of z-values of different signs in opposite sides of the magnitude signature.

Clear examples of magnitude shifts have been identified in California data sets (Habermann and Wyss, 1984). Magnitude shifts are also different from detection changes in that their effects cannot be eliminated using a simple magnitude cutoff. They are also more difficult to interpret. Thus, we rely on modeled magnitude changes for interpretation. After successfully modeling the change, a magnitude correction may be applied to the data set so that the effects of the shift are remedied.

It is clear that a combination of two or more of the described changes can also be possible in reality, yielding a more complex signature.

INTERPRETATION OF MAGNITUDE SIGNATURES WITH THE AID OF MODELED (SYNTHETIC) SIGNATURES

As previously pointed out, magnitude signatures show the results of comparing seismicity rates during two periods termed the background and the foreground. If we assume that the background is the normal rate for all magnitude bands, then we can shift it in time and operate on it in order to attempt reproducing the effects observed on the foreground by creating a synthetic foreground.

A magnitude signature is then determined from comparison of the original background to the synthetic foreground and then this "synthetic signature" is compared to the observed magnitude signature in order to see whether the assumptions involved are correct. The process can be repeated until the best fit between original and synthetic signatures is achieved.

A magnitude shift can be simulated by a shift in the magnitude of events belonging to a particular magnitude band in the background period while a detection change can be simulated by repeating or deleting some events in the same or another magnitude band. The technique is further described in Habermann (1987).

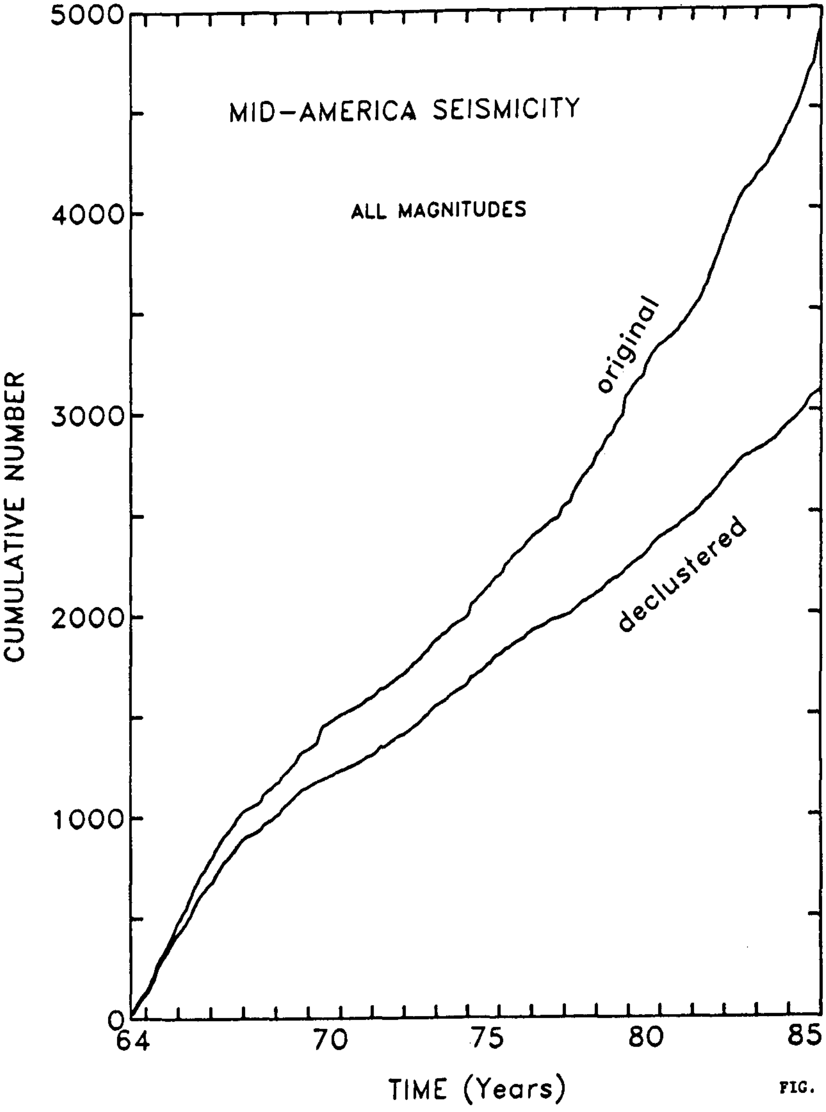


Fig. 3. Curves of cumulative number of events in the Mid-America region (lat. 5N to 33N; long. 80W to 120W) corresponding to the original (uncorrected) NOAA Earthquake Data File and the declustered data set (after removal or aftershocks).

This technique was applied to examine the NOAA Earthquake Data File, using events which occurred in the Mid-America region, to find possible alterations which may introduce artificial seismicity rate changes when employing events listed in this catalog for seismicity studies. Times of the changes as well as suitable corrections will be valuable for all those who wish to study the seismicity of the region and for making comparisons with other regions.

ANALYSIS OF THE NOAA EARTHQUAKE DATA FILE FOR EVENTS IN THE MID-AMERICA REGION

Data reduction

For the analysis of the NOAA catalog, we considered the period which includes all events between June 1964 and December 1985, the reason being that after 1964 the WWSSN (World Wide Standard Seismograph Network) is supposed to have been finally installed (Habermann, 1987) and thus the catalog can be considered more homogeneous.

A first step in the analysis of the NOAA catalog was the identification of aftershocks and clusters. Aftershock removal and declustering is important for studies of seismicity rate changes as possible precursors because an average background rate is necessary. Aftershock sequences could produce rate increases affecting seismicity backgrounds or foregrounds.

Figure 3 shows the result of declustering. It can be seen on this figure that the declustered data set is more constant than the original set with the exception of a major change in slope occurring approximately at the beginning of 1968. Aftershocks were identified visually after determining the time and location of events assumed to be mainshocks (those with $m_b \geqslant 5.5$). Events occurring in the near vicinity of the mainshock (within a radius of approximately 50 km) less than a month after the time of the mainshock were marked as aftershocks and later deleted.

After declustering, the data set was then ready to be analyzed with the aid of magnitude signatures, so the following step involved identifying the major changes in seismicity rates which occurred during the total period of study (July 1964 to December 1985). This was done by comparing a running background (*i.e.* the background increased as we progressed forward in time) to a pre-defined length of foreground. Whenever a statistically significant change took place, the time of this change would be recorded.

The analysis was repeated for a sequence of magnitude bands (26 total, figure 4) and changes which appeared significant in most magnitude bands were then considered for further study. The times for which changes were found to be statistically significant are: end of September 1965 (around the 24th); mid-October 1967 (approximately the 13th); early October 1969 (around the first); end of May 1972 (near the 20th); and end of July 1979 (approximately the 28th).

Magnitude signatures for the periods of interest

Figure 4 shows the magnitude signatures determined for the times mentioned above. In each case a background was defined starting from the time of the previous major rate change and the foreground was taken up to the time of the following occurrence of a change. The figure also shows the synthetic magnitude signature which best fitted the observed one.

The signature calculated for the period July 64-September 65 to September 65-October 67 (Fig. 4a) shows the effect that a combination of variations in reporting and magnitude shifts can produce. This signature was modeled by assuming a detection decrease of 0.6 times the number of small earthquakes in the background (*i.e.* events with a magnitude $m_b < 4.1$). Additionally, a negative magnitude shift (-0.2) in earthquakes of all magnitudes was required in order to attain the lowest residuals.

Simple detection changes are assumed to have taken place during both the middle of October 1967 and early October of 1969. Both cases were modeled by detection decreases and no magnitude shifts were necessary. In the case of October 1967 (Fig. 4b), the change was defined in terms of the seismicity rate between September 1965 and October 1967 and that between October 1967 and October 1969. A decrease of half the number of events in the background for all earthquakes with magnitude $m_b < 4.4$ produced a reasonable fit.

The same amount in the decrease was used to model the following period, but the magnitude cutoff appeared higher ($m_b < 4.7$). In this case (Fig. 4c) the change was defined considering the seismicity rate between October 1967 and October 1969 and that between October 1969 and May 1972.

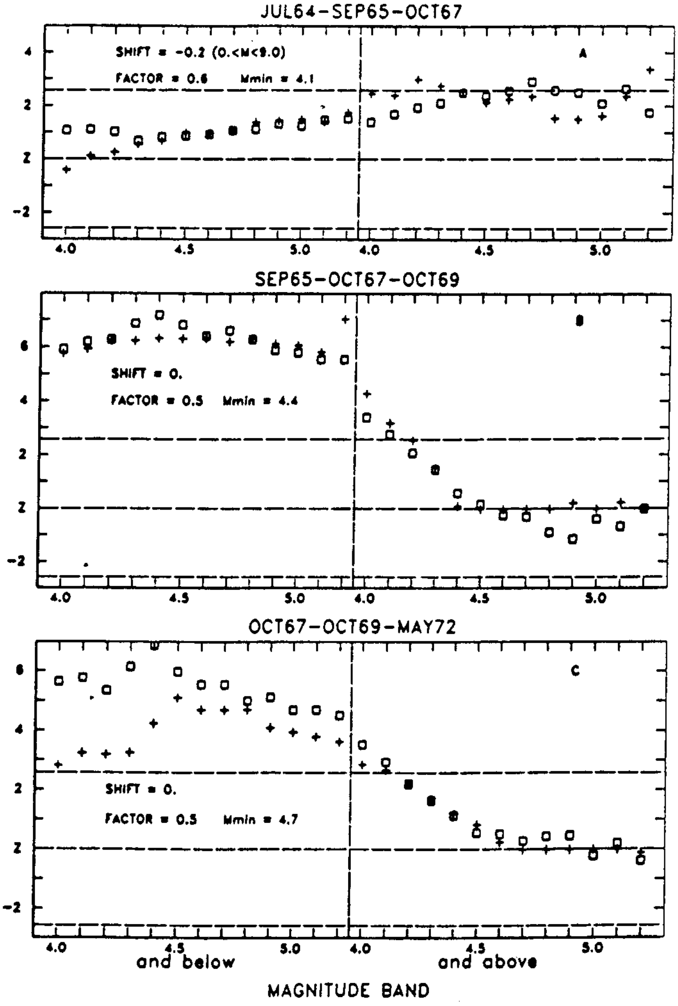
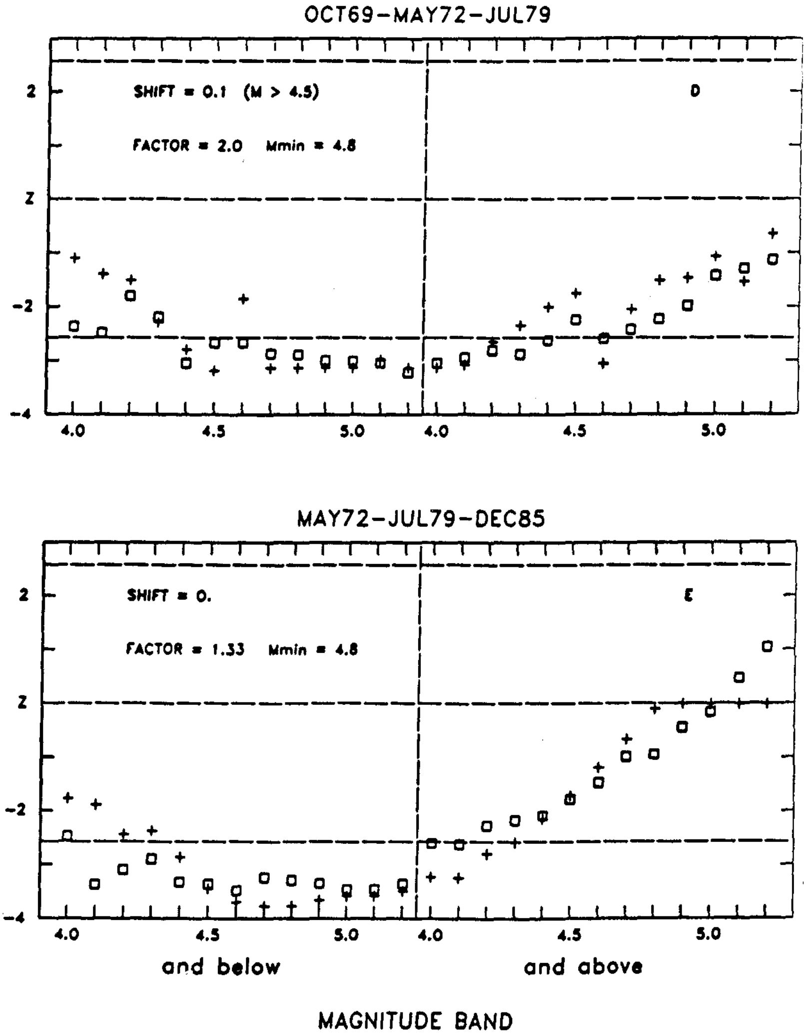


Fig. 4. Observed magnitude signatures and synthetic signatures for events in the Mid-America area and for periods defined by the occurrence of major changes in seismicity rates (see text). In each case, the period considered for background and foreground is shown on the top. Observed signatures are marked by boxes while synthetic signatures are shown with pluses. The assumed corrections (amount and sign of magnitude shift and magnitudes affected, and rate factor employed to simulate increases or decreases in the number of events with a certain magnitude) which produce the synthetic magnitude signature are also shown.

FIG. 4 (cont.)


OCT69-MAY72-JUL79
SHIFT = 0.1 (M > 4.5)
FACTOR = 2.0 Mmin = 4.8
D
2
Z
-2
-4
4.0
4.5
5.0
4.0
4.5
5.0
MAY72-JUL79-DEC85
SHIFT = 0.
FACTOR = 1.33 Mmin = 4.8
E
2
Z
-2
-4
4.0
4.5
5.0
4.0
4.5
5.0
and below
and above
MAGNITUDE BAND

The following change (Fig. 4d), defined by seismicity rates between October 1969 to May 1972 and between May 1972 and July 1979, was again assumed to be product of a combination of changes. In order to obtain the closest fit, a magnitude shift of 0.1 in the positive direction and in the larger magnitude bands ($m_b \geqslant 4.5$), and an increase of twice the number of events of the background for earthquakes with $m_b < 4.8$ were necessary.

Finally, another simple detection change was assumed to have taken place sometime towards the end of July in 1979 (Fig. 4e). For this last case, an increase in the number of earthquakes with magnitudes $m_b < 4.8$, equal to 1.33 times the background events, was used. No magnitude shift was required.

Corrections to the catalog

Based on the results of this analysis we proceeded to correct the catalog. This was done by applying magnitude shifts which would eliminate those shifts observed through the signature interpretations. Rate increases or decreases are not reproduced, since that would imply artificially repeating events.

In Figure 5, cumulative curves are displayed corresponding to the original data set, the declustered data set (after removal of aftershocks) and the declustered and magnitude corrected data set. All curves shown are for events with $m_b \geqslant 4.8$, considering the most conservative magnitude cutoff obtained through the synthetic signatures (highest magnitude cutoff necessary to model rate changes).

The magnitude shifts, assumed to correct for the apparent shifts in magnitude in the catalog, were:

- Events from the 1st of July 1964 to the 24th September 1965: shifted −0.1 magnitude units.
- Events from September 25th, 1965 to May 20th, 1972: shifted +0.1 units.
- Events from the 21st of May, 1972 to the 31st of December 1985: no magnitude shift applied.

In obtaining these corrections a further consideration was that events belonging to the more recent period are to be left without change, so that new events can just be

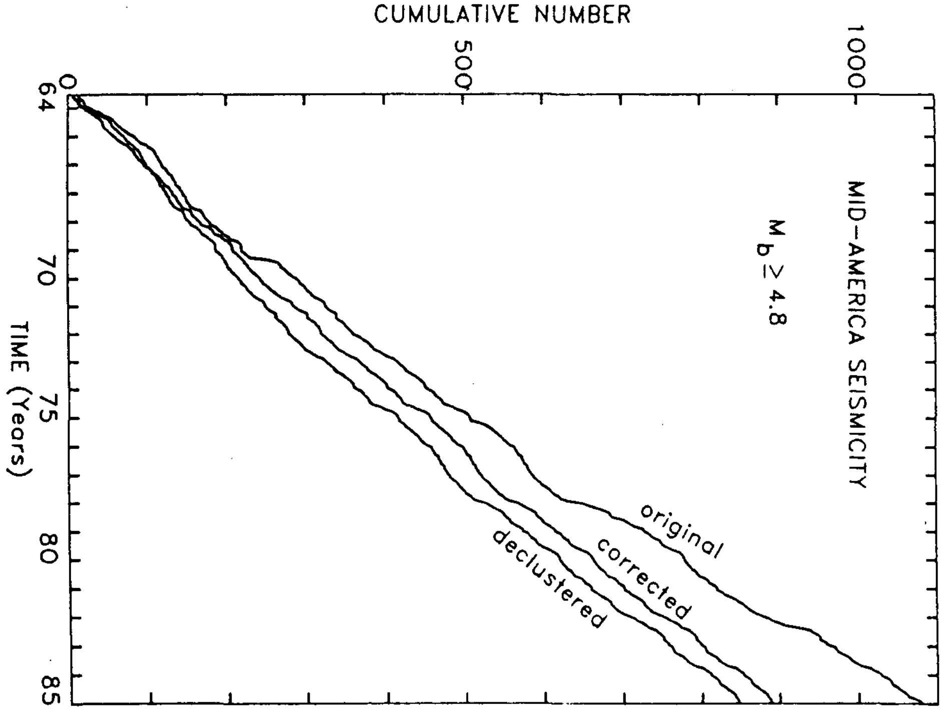


Fig. 5. Cumulative number of events calculated using the original data set, the declustered data set and the data set which was corrected for the apparent magnitude shifts. All cases shown include events with a magnitude $m_b \geqslant 4.8$, taking into account the magnitude cutoff determined.

added to the existing information without the need of performing additional corrections.

The curves in the figure demonstrate that the average seismicity rate (defined by the slope) follow a more constant trend after declustering and even more so after magnitude corrections have been made. Since these corrections were performed by "blind" simulations (not involving any predefined values), we can take them as valid for the studied region.

DISCUSSION AND CONCLUSIONS

In other studies it had been established that the closure or opening of seismic stations, particularly in the U. S., was responsible for most of the changes in detection and reporting of events in Mexico and Central America (Habermann, 1982; Habermann, 1987).

In the case of the changes observed for 1965, 1967 and 1969, which involve detection decreases for the three time periods, one possibility is that closure of stations (like those belonging to the VELA array, *e.g.* Habermann 1987), was responsible for the apparent differences in reporting and detection, since stations ceased to operate around those times.

For the 1979 change, whose signature indicate that the number of events detected increased around that time, a distinct possibility is that new stations in Mexico (like those belonging to the SISMEX and RESMAC arrays, which started operating in the mid and late-seventies respectively) and changes in monitoring techniques may have affected the reporting level. Something similar may have occurred in the case of the 1972 change, since it also involved a detection increase, although new stations that may have been involved are not known to present time.

Nevertheless, the fact that no change was detected for events reported after July 1979 and till the end of the period of study (December 1985) is an indication that even though new stations have been added to the networks and other have closed, the consistency of the catalog has not been altered. This may also show that the coverage of the current networks is already quite reasonable for events larger than the cutoff ($m_b \geqslant 4.8$).

In summary, if the proposed corrections and suggested magnitude cutoff are employed, a more stable average seismicity rate is obtained. When studies of seismicity rates in the Mid-America region are to be undertaken, these corrections are necessary in order to define a background against which localized rate changes may be detected.

ACKNOWLEDGMENTS

I would like to thank M. Wyss and R. E. Habermann for their help and insight. R. E. Habermann also provided computer codes for the analysis. W. Rinehart helped in gathering the data. The comments of two anonymous reviewers improved the presentation of this paper. This work was partially supported by a grant from the National Oceanic and Atmospheric Administration of the U.S.

BIBLIOGRAPHY

HABERMANN, R. E., 1987. Man-made changes of seismicity rates. *Bull. Seism. Soc. Am., 77,* 141-159.

HABERMANN, R. E. and M. WYSS, 1984. Background seismicity rates and precursory quiescence: Imperial Valley, California. *Bull. Seism. Soc. Am., 74,* 1743-1756.

HABERMANN, R. E., 1982. Seismicity rates in the Kuriles Island Arc, 1973-1979. *Earthquake Pred. Res. 1,* 73-94.

HABERMANN, R. E., 1981. Precursory seismicity patterns: stalking the mature seismic gap. *In:* Earthquake Prediction, D. W. Simpson and P. G. Richards, (Eds.), American Geophysical Union, Washington, D. C., 29-42.

KANAMORI, H., 1981. The nature of seismicity patterns before large earthquakes. *In:* Earthquake Prediction, D. W. Simpson and P. G. Richards, (Eds.), American Geophysical Union, Washington, D. C., 1-19.

McNALLY, K., 1981. Plate subduction and prediction of earthquakes along the Middle America trench. *In:* Earthquake Prediction, D. W. Simpson and P. G. Richards, (Eds.), American Geophysical Union, Washington, D. C., 63-72.

MEYER, S. L., 1975. Data analysis for scientists and engineers, John Wiley and Sons, New York.

OTHAKE, M., T. MATUMOTO and G. LATHAM, 1977. Temporal changes in seismicity preceding some shallow earthquakes in Mexico and Central America. *Bull. Inter. Inst. Seismol. Earthquake Eng. 15,* 105-123.

WYSS, M., R. E. HABERMANN and J.-C. GRIESSER, 1984. Seismic Quiescence and Asperities in the Tonga-Karmadec arc. *J. Geophys. Res., 89,* 9293-9304.

WYSS, M., R. E. HABERMANN and C. HEINIGER, 1983. Seismic quiescence, stress-drops and asperities in the New Hebrides arc. *Bull. Seism. Soc. Am., 73,* 219-236.

WYSS, M., F. W. KLEIN and A. C. JOHNSTON, 1981. Precursors to the Kalapana $M_s = 7.2$ earthquake. *J. Geophys. Res., 86,* 3881-3900.