Processing Earthquake Catalog Data for Seismic Hazard Analysis

A. Deniz ¹, M. S. Yucemen²

¹Technological Engineering Services Co. Ltd., Antalya, Turkey
²Department of Civil Engineering, METU, Ankara, Turkey

Abstract

Earthquake catalogs are the most important sources of information in forming the seismic database to be used in seismic hazard analysis. However, the information presented in the catalogs cannot be used directly and has to be processed. Generally, earthquake magnitudes are reported in different magnitude scales and it is desirable to form a unified catalog by converting the different magnitude scales into a single one. The implementation of the Poisson model requires the elimination of the spatial and temporal dependencies created by fore and aftershocks. Also the earthquake catalogs are often biased due to incomplete reporting for small magnitude earthquakes. The paper aims at presenting procedures for overcoming these problems. As an output of the study a unified catalog for the earthquakes that occurred in Turkey and the neighboring regions is compiled in terms of the moment magnitude scale.

Keywords: earthquake catalog, earthquake cluster, incompleteness, magnitude conversion, seismic hazard.

1 Introduction

One of the most important and fundamental inputs to a seismic hazard analysis is the past seismic activity data, which describes the spatial and temporal distribution of earthquakes in the region of interest. For this purpose, a seismic database has to be compiled based on available earthquake catalogs. However, the information presented in the catalogs cannot be used directly because of non-uniformity in magnitude scales, dependency and incompleteness problems.
Earthquake magnitudes reported in different magnitude scales require the conversion of different magnitude scales into a single scale, in order to unify the earthquake catalog. Besides, the classical probabilistic seismic hazard analysis (PSHA), based on the widely used Poisson model, assumes the occurrence of earthquakes as spatially and temporally independent. This requires the elimination of the spatial and temporal dependencies created by fore and aftershocks. Another problem is due to the fact that earthquake catalogs are often biased because of incomplete reporting for small magnitude earthquakes as well as for large magnitude earthquakes having long return periods. In the following sections, these problems are discussed and alternative methodologies are presented for handling them.

2 Development of a Single Scale Earthquake Database

In converting the different magnitude scales contained in earthquake catalogs to a single scale, because of various reasons, it is more appropriate to take the moment magnitude (Mw) as the basis. The conversion of earthquake magnitudes reported in different scales, namely, surface wave magnitude, Ms, local magnitude, ML, body wave magnitude, Mb and duration magnitude, Md to Mw is quite an important task. A number of relationships have been developed empirically at the international level between moment magnitude and other magnitude scales (e.g. Boore and Joyner, 1982). On the other hand, Ulusay et al. (2004) have conducted a similar study and developed a set of relationships based on the earthquakes that occurred in Turkey. However these relationships are obtained by using the standard least squares regression analysis in which it is assumed that the independent variable (each magnitude scale to be converted) is free from measurement error and consider the random error on the dependent variable (Mw) only. Since, a certain degree of error is present also on the independent variables, the assumptions of the classical least squares regression fail. In this case, the orthogonal regression method suits well to the needs of the regression analysis. In a recent study, Castellaro et al. (2004) investigated the bias introduced if the analysis is based on the standard regression procedure and found out that large errors (up to 0.4 in some magnitude scales) are possible for the Unified Italian Catalog that they have developed.

Employing the orthogonal regression, Deniz (2006) developed the following conversion equations based on earthquakes that occurred in the last 100 years in Turkey, as given in Equations 1.a – 1.d.

\[
\begin{align*}
M_w &= 2.25 \times M_s - 6.14 \\
M_s &= 1.27 \times M_b - 1.12 \\
M_L &= 1.57 \times M_s - 2.66 \\
M_d &= 0.54 \times M_s + 2.81
\end{align*}
\]

Orthogonal regression estimates the slope of the conversion equation greater than that obtained from the standard least squares regression. Accordingly, moment
magnitude values obtained from the orthogonal regression based conversion equations will be larger compared to those obtained from the standard least squares equations for large magnitude values. The opposite is valid for small magnitudes. Considering the fact that the contribution of small magnitude earthquakes to seismic hazard is low, the use of orthogonal regression based conversion equations will yield more conservative seismic hazard values.

In our study, an earthquake database expressed entirely in terms of the moment magnitude scale is obtained by using the expressions given by Deniz (2006).

### 3 Elimination of Fore and Aftershocks

The computational approach to PSHA, based on the widely used Poisson model, assumes the occurrence of earthquakes as spatially and temporally independent. In order to satisfy the assumptions of the Poisson process, it is necessary that earthquake clusters should be identified and dependent events (fore and aftershocks) be eliminated from the earthquake catalog.

A number of studies have been conducted in the past to develop procedures for the identification of fore and aftershocks from spatial and temporal points of view, e.g. (Omori, 1894; Gardner and Knopoff, 1974; Prozorov and Dziewonski, 1982; Van Dyck, 1985; Utsu et al., 1995; Savage and Rupp, 2000 and Kagan, 2002). It is observed that fore and aftershocks exhibit similar behavior spatially and temporally around the main shock. Accordingly, procedures for the identification of the fore and aftershocks do not differ. In these studies, it is assumed that earthquakes of a specified magnitude cause the same secondary seismic activity irrespective of the seismic source, region, type and length of the fault. For each earthquake magnitude level, all earthquakes within a specified space and time window after a main shock, having lower levels of magnitudes are assumed to be the aftershocks of the initial main shock. For an earthquake to be identified as a foreshock, it is assumed that there should be another earthquake having a greater magnitude in its vicinity (within the time and space window associated with its magnitude level). In such cases the earthquake having the greater magnitude is assumed to be the main shock. However, all earthquakes with magnitudes larger than 6.0 are treated as main shocks, exceptionally.

Deniz (2006) developed space and time windows based on the studies conducted by Van Dyck (1985), Utsu et al. (1995), Savage and Rupp (2000) and Kagan (2002). He used the enveloping space windows of the four studies mentioned above for specifying the dimensions of the space windows for aftershocks. For the dimensions of the time windows corresponding to different magnitude levels, he took the average of the values given by Gardner and Knopoff (1974) and Savage and Rupp (2000). The resulting dimensions of the space and time windows are given in Table 1. For magnitude values that are not listed in this table linear and log-linear interpolation is to be used for determining the corresponding time and space windows, respectively. The values given in Table 1 are recommended values that are obtained based on a limited study, which should be taken into consideration when they are implemented.
Table 1. The dimensions of the space and time windows for the identification of fore and aftershocks (after Deniz, 2006)

<table>
<thead>
<tr>
<th>Magnitude</th>
<th>Space window (km)</th>
<th>Time window (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>35.5</td>
<td>42</td>
</tr>
<tr>
<td>5.0</td>
<td>44.5</td>
<td>83</td>
</tr>
<tr>
<td>5.5</td>
<td>52.5</td>
<td>155</td>
</tr>
<tr>
<td>6.0</td>
<td>63.0</td>
<td>290</td>
</tr>
<tr>
<td>6.5</td>
<td>79.4</td>
<td>510</td>
</tr>
<tr>
<td>7.0</td>
<td>100.0</td>
<td>790</td>
</tr>
<tr>
<td>7.5</td>
<td>125.9</td>
<td>1326</td>
</tr>
<tr>
<td>8.0</td>
<td>151.4</td>
<td>2471</td>
</tr>
</tbody>
</table>

4 Adjustment for the Incompleteness in Earthquake Catalogs

The earthquake catalog data to be used in the seismic hazard analysis must be complete at all magnitude levels. The quality and the amount of the recorded data become much poorer as recordings become ancient. In recent years, all earthquakes, small and large, are recorded, whereas old records show only large magnitude events. Furthermore, recorded events almost always refer to places of habitation, neglecting the possible stronger earthquakes in unpopulated areas. The incompleteness of earthquake catalog creates biases in the database both in time and space. Accordingly, the resulting recurrence relationships may not represent the true long term rates. For this reason, it is necessary to determine the time period over which the data in a given magnitude interval is completely reported. The annual rates of earthquake occurrences with respect to their magnitude range are then computed by considering only the corresponding time intervals.

To adjust the number of earthquakes artificially in the earthquake catalogs for the incompleteness that is present in observations, the method proposed by Stepp (1973) has been widely used. In this procedure earthquakes are grouped in magnitude intervals and each magnitude interval is modeled as a point process in time. The basic concept of statistical estimation that the variance of the sample mean is inversely proportional to the number of observations in the sample is utilized. Thus, the variance can be made as small as desired by making the number of observations large enough, provided that data reporting is complete in time and the process is stationary, i.e. the mean, variance and other moments of each observation are the same. The earthquake sequence can be modeled by a Poisson distribution. If \( k_1, k_2, \ldots, k_n \) are the number of earthquakes per unit time interval, the unbiased estimate of the mean rate per unit time interval of this sample is:

\[
\lambda = \frac{1}{n} \sum_{i=1}^{n} k_i
\]

and its variance is
\[ \sigma^2 = \frac{\lambda}{n} \]  
(3)

where, \( n \) is the number of unit time intervals. Taking the unit time interval as one year gives,

\[ \sigma_\lambda = \sqrt{\frac{\lambda}{T}} = \frac{\sqrt{\lambda}}{\sqrt{T}} \]  
(4)

Here, \( \sigma_\lambda \) is the standard deviation of the estimate of the mean and \( T \) is the sample length in years. Thus, assuming stationarity, \( \sigma_\lambda \) is expected to behave as \( 1/\sqrt{T} \) in a given sample in which the mean rate of occurrence in a magnitude interval is constant. If the mean rate of occurrence is constant, we expect stability to occur only in the subinterval that is long enough to give a good estimate of the mean, but short enough that it does not include intervals in which reports are incomplete. Once this time period is determined, then the mean annual rates of earthquake occurrences for that magnitude range is computed by considering only that time period.

5 Unified Turkish Earthquake Catalog

As a part of this study, a seismic database is compiled for the earthquakes that occurred in Turkey by applying the necessary modifications within the framework of the concepts stated in the previous sections. It is expected that this seismic database can be used in seismic hazard studies. The seismic hazard created at a site is due to not only earthquakes happening right at that site, but also due to earthquakes in a certain vicinity, which is generally accepted to be approximately as 250 kms in radius, depending on the sensitivity of the analysis. Accordingly, a 2° North-South and East-West extension of coordinates of Turkey are selected to constitute the probable event domain in space, resulting in an area bounded by 34° – 44° North latitudes and 24° – 47° East longitudes. The time period considered is from 1900 to 2004.

Five large data sets are used in the compilation of the seismic database. These are provided by the Earthquake Research Department (ERD) of General Directorate of Disaster Affairs of Turkey (GDDA-ERD, 2004), Kandilli Observatory and Earthquake Research Institute of the Bogazici University (KOERI, 2004), International Seismological Centre (ISC, 2004a and ISC, 2004b) and United States Geological Survey (USGS, 2004a). Also, a small data set is reported by ERD, in which earthquakes are given in more than one magnitude scale (Inan et al., 1996). Another small data set of large earthquakes is made available by USGS, in which large earthquakes prior to 1973, the starting year of the main USGS catalog, are accessible (USGS, 2004b). All of these component catalogs have unique properties and the details of these can be found in Deniz (2006).

The data sets listed above, formed the preliminary input catalogs giving information on date, time, geographic location, depth and magnitude of recorded events in the combined (pooled) catalog. The data in the pooled catalog are
screened to avoid duplicate events. For this purpose, time and space windows are specified to identify the same earthquake. To take care of especially old records, a time window of 3 minutes and a space window of 0.2° latitudes and longitudes are selected. In other words, any two events reported in two different catalogs, that fall within the bounds of these time and space windows are assumed to belong to the same event. For relatively recent data, required precautions are taken to avoid any mistake due to this assumption. 28530 events are accumulated from the component catalogs. Total number of different events complying with the pre-defined bounds is found out to be 16015 at the end. Approximately 44% of the records are determined to belong to some other events with the above specified same-event-neighborhood. The small differences in geographic location, time and depth of an event as reported by different catalogs are eliminated by assigning the average values of the corresponding parameters to that event. Magnitude records of all institutions in all scales are determined and it is observed that even records of different institutions in the same scale may differ. This is attributed to the way of processing the available instrumental data or data collection capabilities of these institutions (Ambraseys and Melville, 1982). After performing a set of statistical analyses, such differences are determined to be systematic for only $M_s$ records of the KOERI (2004) catalog. Such records are adjusted and then the magnitude of each earthquake in a selected scale is considered to be the average of the available records due to different sources.

The earthquakes in the pooled catalog are reported in different scales. These are converted to the moment magnitude ($M_w$) scale by using Equations 1.a – 1.d. The resulting seismic database contains 4752 earthquakes with magnitudes $M_w \geq 4.5$ and is called as the Unified Turkish Earthquake Catalog. This catalog is available, in digital format in Deniz (2006).

It is recommended that secondary event identification and completing the available earthquake catalog should be carried out in seismic source scale rather than the whole geographical domain. In other words, screening of the secondary events and adjustment for incompleteness should be done by considering only the earthquake databases of individual seismic source zones, which are composed of the earthquakes associated with these seismic source zones.

Within the context of the previous discussions, the whole pooled catalog is analyzed with respect to main shocks and secondary events in seismic source scale (based on a comprehensive set of seismic source zones). Applying the space and time windows given in Table 1, 8.67 % of the pooled catalog records are identified as foreshocks, 55.68 % as main shocks and the remaining 35.65 % as aftershocks. The catalog composed only of main shocks is given in Deniz (2006). For PSHA purposes, instead of artificially completing the available catalogs, the seismicity parameters of the seismic source zones are adjusted, based on the completely reported time periods for different earthquake magnitude levels. For this reason, complete catalogs are not virtually available. However, their modified seismicity parameters can also be found in Deniz (2006).
6 Concluding Comments

In this study, the main deficiencies of earthquake catalogs are discussed and procedures for correcting these deficiencies are presented. A Unified Turkish Earthquake Catalog is compiled, where all earthquakes are listed only in the moment magnitude scale. The following points should be emphasized in compiling seismic databases based on existing earthquakes catalogs.

There are a number of different data sources providing past earthquake data. The most complete earthquake catalog should be constructed by using all of the available seismic databases. The presence of five different magnitude scales (Mc, Ml, M, Ms and Mw) in the available earthquake catalogs makes it necessary to perform regression analyses to develop conversion equations with respect to a selected scale. The moment magnitude (Mw) is selected as the target scale in this study. During the development of conversion equations among different magnitude scales, statistical methods considering the random errors both in the independent and dependent variables should be employed. Here, the orthogonal regression procedure is utilized. This method is observed to yield larger converted Mw values as compared to the standard least squares regression method, especially for large magnitude earthquakes.

In seismic hazard analysis, main shocks and secondary shocks (foreshocks and aftershocks) should be differentiated and alternative analyses should be carried out. Also, the earthquake catalogs should be corrected for incomplete reporting, especially for the early parts of the available data sets.

References

or at: http://library.metu.edu.tr/ (search by author’s name: Deniz, Aykut).


