Transaction A: Civil Engineering
Vol. 17, No. 1, pp. 62–73

Research Note

Seismic Risk Analysis of Iranian Construction Projects

M. Shokri-Ghasabeh[1], A. Bakhshiani[2], M. Mofid[3,*] and K. Hansen[4]

Abstract. *In this paper, the project earthquake occurrence risk coefficient is determined for each construction project that is located in one of Iran's twenty seismic regions. This coefficient is allocated, regardless of the current situation of the project, being in the plan or execution phase or even completed. This coefficient indicates the possibility of an earthquake occurrence during a project's life time. To find this coefficient, the Gutenberg-Richter linear relationship has been applied, in conjunction with the Poisson distribution. The Gutenberg-Richter linear equation expresses the relationship between the magnitude of an earthquake and the number of occurrences, during a fixed time, of that magnitude. To find the linear relationship for a series of earthquakes with different magnitudes occurring in the same seismic region, the Ordinary Least Square (OLS) has been used. Two linear regression assumption violations, which are variance heteroscedasticity and autocorrelation, have been tested on the available data. In the case of finding one or both of these two violations, The Generalized Least Square (GLS) has been applied to produce a better regression line. Moreover, the second order type of the Gutenberg-Richter relationship has also been determined to validate the linear one. In conclusion, by application of the Poisson distribution and by having the design earthquake's magnitude and project life time, the third parameter, which is the design earthquake occurrence risk, can be determined for a given construction project in a specific location in Iran.*

Keywords: *Earthquake risk analysis; Construction management; Project management; Risk management; OLS, GLS;*

INTRODUCTION

Nowadays, a variety of major construction projects all over the world are being planned for different purposes. Due to the complex nature of such projects, any consequent managerial decision, whether in the research, execution or maintenance phase, can dramatically change a project's scope, leading to either improvements or considerable losses. One of the main factors that influence the design decision of a project is consideration of the risk of natural disasters, especially in the case of earthquakes. The estimation of earthquake occurrence risk in some areas that are prone to having earthquakes is particularly important for a vital and important infrastructure, as it prevents or reduces the loss of financial capital. Generally speaking, both managers and engineers have become either too ignorant or too conservative about executing a project and designing it to withstand earthquakes. Poorly designed projects will cause life-threatening situations and financial damage, but on the other hand, conservatively planned projects will cause unnecessary cost and the waste of both materials and workforce.

As most major Iranian construction projects are exposed to the likely risks concerned with earthquake occurrence, any damage to important developmental and industrial projects, such as the construction of dams, energy power stations, refineries and so forth, creates not only economical, but also social, hygienic and political impacts. Strategic impacts are also included when these projects are located in the Iranian border regions. By taking Iran's vulnerability to earthquakes into consideration and acknowledging that earthquakes are known as the most frequent and

1. *School of NBE, University of South Australia (UniSA), Adelaide, SA, 5000, Australia.*
2. *Faculty of Engineering, Tehran University, Tehran, P.O. Box 11163-9313, Iran.*
3. *Department of Civil Engineering, Sharif University of Technology, Tehran, P.O. Box 11155-9313, Iran.*
4. *School of Mechanical Engineering, University of Adelaide, SA, 5005, Australia.*
*. *Corresponding author. E-mail: mofid@sharif.edu*

Received 1 January 2008; received in revised form 9 June 2009; accepted 15 December 2009

destructive natural disaster in Iran, it is evident that an assessment of the risk of earthquake occurrence is imperative. Indeed, a specific construction project undertaken in Iran can be more efficiently designed and managed if the earthquake occurrence risk, as a very crucial natural type of project risk, is addressed.

Iran and Earthquakes

Iran is situated in an area that is geologically and tectonically prone to earthquakes, called the Alpe-Himalayan belt or the Alpide. This area starts from the East of Asia, passes through Indonesia and Myanmar, North of India, Pakistan, Afghanistan, Iran and Turkey extends up to the Mediterranean Sea and, finally, to the Southwest of Europe. This region is circled in Figure 1. In spite of having a few shallow depth earthquakes, this area is known for medium depth earthquakes, which are a result of volcanic activity [1]. Specifically, as mentioned by Fattahi et al. [2], Iran is one of the most seismically active regions in this area, with numerous destructive earthquakes recorded both historically and instrumentally. As presented in Figure 1, Iran is coloured completely red, which indicates that this part of the world is very vulnerable to seismic activity (i.e. earthquakes).

Research on the Earthquake Occurrence Risk in Iran

Various research methods have been applied to evaluate the earthquake risk in Iran, and one of these methods is statistical-based research. Statistical-based research on the data of previous earthquakes collected from a given region can be contributed to two specialized branches of civil engineering: earthquake and construction management. Some similar research topics were previously undertaken by foreign and local researchers under the earthquake engineering branch of science (i.e. [3-9]). In regard to management, so far there has been little research on earthquakes as a project natural risk. In this paper, earthquakes are researched under the construction management concept to show the importance of the application of earthquake risk

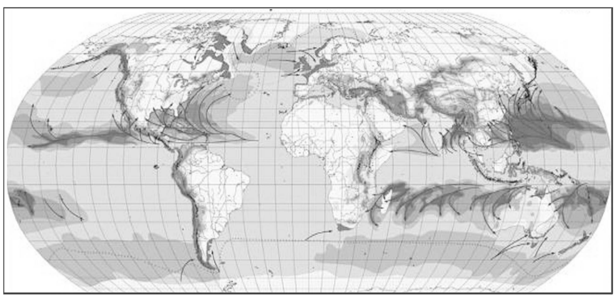

Figure 1. Iran in the world map of natural hazards.

in construction management. The result establishes a meticulous criterion to consider in a project's structural and execution plan. By means of the usage of some statistical tests, such as second order regression, GLS and the Poisson statistical model, as suggested by the authors, researchers can increase the accuracy of earthquake occurrence risk calculations.

EARTHQUAKE OCCURRENCE POSSIBILITY DETERMINATION BY MEANS OF STATISTICAL MODELS

Magnitude of an Earthquake and Its Number of Occurrences: The Relationship

The best possible equation that can be used to describe the relationship between the number of earthquake occurrences and their magnitude was suggested by Gutenberg and Richter [10], which is a linear relation as follows:

$$\log(N) = A - bM, \qquad (1)$$

where N is the number of earthquakes over a defined time interval of magnitude greater than, or equal to M, and A and b are constant coefficients, which differ from one area to another. By considering the number of occurrences of earthquakes with different magnitudes over a fixed time interval, i.e. 100 years ago until now, A and b can be determined by means of a method such as an OLS (Ordinary Least Square) line. As an example, Banisadr [3] determined the following general regression line for all earthquakes in Iran:

$$\log(N) = 5.79 - 067M. \qquad (2)$$

The correlation coefficient, which is a positive number, shows the accuracy and validation of the regression line. The closer the correlation coefficient is to unity, the more accurate the regression line is. In statistical calculations, whether linear or non-linear, researchers always tend to move the correlation coefficient closer to 1, in order to obtain a more precise regression calculation. In addition, the relationship between earthquake magnitude and the logarithm of its reoccurrence over a defined time interval can be interpreted in another way by means of a second order relationship, as follows:

$$\log(N) = A + bM + CM^2, \qquad (3)$$

where, A, b and C can be either positive or negative constants. This relationship usually produces a better correlation coefficient than the linear model, although this second order relationship is in fact a curve, which tends toward a linear relation by having a very small C coefficient.

Seismic Regions of Iran

Iran has been classified differently by various Iranian and non-Iranian researchers in regard to its tectonic regions. For instance, Berberian [4], Stoklin [7] and Takin [8] divided Iran into 4, 9 and 4 seismic areas, respectively. On the other hand, Nowroozi [5] categorized Iran into 23 seismic areas in a thorough research work. The authors of this research applied Tavakoli's [9] seismic classification, which categorizes Iran into 20 seismic regions. The use of this classification is justified, as it accurately takes into consideration the tectonic structure of each area and ensures that each region includes one or more major active earthquake faults.

The Application of the Poisson Distribution

In the next stage of the research, a mathematical distribution model can be applied to estimate the probability of an earthquake's occurrence. Different models have been used for this purpose by different researchers. However, the Poisson model is considered as the most popular because of its simplicity [1]. In general, the Poisson model is expressed as follows [11]:

$$P_T(n,t) = \frac{(Nt)^n . e^{-Nt}}{n!}, \tag{4}$$

where, $P_T(n,t)$ is the percentage of occurrences of n earthquakes with magnitudes greater than or equal to M over the time interval, t. As mentioned by Bargi [1], this model can be simplified for the occurrence of only one earthquake, with a magnitude greater than or equal to M over the time interval of t, as follows:

$$P_T(t) = 1 - e^{-N.t}. \tag{5}$$

Therefore, by applying the Poisson distribution and by knowing the project's life period (t) and the possibility of an earthquake's occurrence ($P_T(t)$), the parameter N and, consequently, the magnitude of an earthquake that might possibly strike the project during its life period, can be determined. The main objective is to find the most intense earthquake that could possibly strike a project during its lifetime in order to design the project based on that possible earthquake.

Earthquake Magnitude Scales

There are four main earthquake magnitude scales, which are known as body-wave magnitude (mb), surface-wave magnitude (Ms), local magnitude scale (Ml) and moment magnitude scale (Mw) [12]. Currently, the Mw parameter is the most reliable and important scale for earthquake magnitude. However, there are few earthquakes in Iran that have been measured by Mw and this scale has not been used to determine their magnitudes. Moreover, there is not enough research in Iran on how to convert other earthquake magnitude scales to Mw. The second most reliable scale is the surface-wave magnitude (Ms). On the other hand, the body-wave magnitude (Mb) has been used to show the magnitude of most of the earthquakes recorded in Iran, especially since 1963. Since Mb is not a suitable scale for seismic study and estimation of earthquake occurrence [13], it is necessary to find a scientific-based relationship between these two magnitude scales in order to interpret all earthquake statistics with one scale which is Ms. Mirzaei [13] introduced a general relationship between Ms and Mb as follows:

$$Ms = d.Mb + c, \tag{6}$$

where d and c are constant coefficients. Mirzaei [13] also suggested some converter equations concerning four general areas in Iran (Table 1). To match Mirzaei's classification, some of Tavakoli's [9] 20 seismic regions are allocated to Mirzaei's 4 general divisions by the authors. The results can be found in the left column of Table 1. Therefore, by means of these equations, the body-wave magnitude (Mb) can be converted to surface-wave magnitude (Ms).

INVESTIGATION ON THE VIOLATION OF THE LINEAR REGRESSION ASSUMPTIONS

Variance Heteroscedasticity

According to Judge et al. [14], all regression models are based on a series of assumptions and linear regression methods, including Ordinary Least Square (OLS), which also require the validity of some initial

Table 1. Conversion relationships proposed by Mirzaei [13] (the event regions are matched by the authors).

Event Regions	$M_S - M_b$ Relationship	Magnitude Range
Zagros (9, 10, 11, 12, 13)	$M_S = 1.79M_b - 4.32$	$4 \leq M_b \leq 6.2$
Azarbaijan-Alborz-Kopedagh (14, 15, 16, 17, 18, 19, 20)	$M_S = 2.01M_b - 5.28$	$4 \leq M_b \leq 6.2$
Central and eastern parts (4, 6, 7, 8)	$M_S = 2.00M_b - 5.28$	$4.1 \leq M_b \leq 6.2$
Mokran region (1, 2, 3, 5)	$M_S = 1.58M_b - 3.11$	$4 \leq M_b \leq 5.9$

assumptions. One of these assumptions is the existence of variance homoscedasticity, which indicates that each data set has a variance equal to the other data sets. Variance homoscedasticity is statistically defined as follows [15]:

$$E(U_i^2) = \sigma_i^2. \quad (7)$$

If a statistical test proves that data variance homoscedasticity exists, it implies that the OLS method is applicable and the results will be precise. However, if the test shows the existence of data variance heteroscedasticity (difference of variances), then OLS will not be valid and another method, called the Generalized Least Square (GLS) method, should be applied. In GLS, the main relationship is transformed into other relationships that cover the homoscedasticity of variance. Therefore, after this transformation, the researchers are able to analyze the new relationship by OLS.

Variance Heteroscedasticity Identification

To identify variance heteroscedasticity, some methods, which have been studied and designed by different researchers and are based on professional statistical methods, are more often applied. The most reliable and applicable test is the Goldfeld-Quandt test, which includes four stages. The interpretation factor for this test is λ, which presumably follows the F distribution and must be compared to the critical F parameter yielded from this distribution. In fact, heteroscedasticity exists if λ is greater than F and, in this case, the GLS method should be applied. Conversely, if λ is less than F, the Ordinary Least Square model is applicable, on account of variance homoscedasticity. As described above, λ follows the F statistical distribution and, if the number of records is high, it can also follow a normal distribution.

Problem Solution

A common problem is that the variances are usually unknown. To solve this problem, the following three-stage solution is considered [15]:

- An assumption is made to establish a relationship between variance and the mathematical expectation (E) of squares of residuals of the main relationship, such as:

$$E(u_i^2) = \sigma^2 X_i, \quad (8)$$

$$E(u_i^2) = \sigma^2 / X_i^2, \quad (9)$$

$$E(u_i^2) = \sigma^2 / X_i, \quad (10)$$

$$E(u_i^2) = \sigma^2 / X_i^2. \quad (11)$$

- The main relationship is then transformed into a relationship with homoscedasticity, which can apply OLS.
- Linear regression is calculated using OLS for the data set.

After adopting all three different relationships between variance and mathematical expectation, the resulting regression lines should be tested in regard to their accuracy. To validate the accuracy of the new regression lines, correlation coefficients should be calculated. The closer the correlation coefficients are to 1, the more reliable the adopted relationship between variance and mathematical expectation will be.

Auto-Correlation

Absence of autocorrelation is the other important assumption of linear regression models. This means that linear regression models are based on the assumption that the residuals (u_i) are dependent on one other. On the other hand, autocorrelation as an assumption violation means that the residuals of an observation are completely independent from the residuals of other observations.

According to Gujarati [15], the relation between residuals is as follows:

$$U_t = \rho U_{t-1} + \varepsilon_t, \quad (12)$$

where, ρ is a coefficient that linearly relates two consecutive residuals.

Autocorrelation Identification

There are a few different statistical tests that can be applied to discover the existence of data autocorrelation. The most famous identification test is called the Durbin-Watson, d, test. It is known as the d test, since the determinant parameter is introduced as d. Once the d parameter is found, and by having three factors, including the number of data sets, the confidence level percentage and explanatory variables, two other constants of d_U and d_L are derivable from the Durbin-Watson table. According to Table 2, by comparing d against d_U and d_L, the existence of either negative or positive autocorrelation can be proven.

Table 2. Autocorrelation identification according to the Durbin-Watson table.

Condition	Result
$0 < d < d_L$	Positive autocorrelation
$d_L \leq d \leq d_U$	No positive autocorrelation
$4 - d_L < d < 4$	No autocorrelation
$4 - d_U \leq d \leq 4 - d_L$	No negative autocorrelation
$d_U < d < 4 - d_U$	Negative autocorrelation

Problem Solution

GLS is the definite solution of autocorrelation in which the main relationship is transformed into another relationship that does not have autocorrelation and, therefore, can be analyzed by OLS. Generally, ρ is unknown and, instead, must be replaced by a new coefficient. Different researchers have proposed different methods for determining ρ. However, the Durbin-Watson d test is the most popular method and is formulated as follows [15]:

$$\hat{\rho} \approx 1 - \frac{d}{2}, \tag{13}$$

where d is the parameter determined from the Durbin-Watson test. The symbol $\hat{\rho}$ is not equivalent to ρ, but is an estimation of this parameter. Also, in cases where the amount of data is low, another equation for ρ, which is given by Theil-Nagar, is applicable. With use of the parameter ρ, the line equation, which has no autocorrelation, can be written. As a result, OLS regression can be applied, since the new relationship has the regression assumption. To check the efficiency of the new line, its correlation coefficient should be compared with the initial line. The closer the correlation coefficient is to 1, the more reliable the new line will be.

CASE STUDY: RESEARCH ON IRAN'S EARTHQUAKES

Data Collection: Iran's Earthquake Statistics

Statistical data concerning earthquakes in Iran can be collected from a variety of different resources. These include books on Iran's earthquakes, comprised of historical and updated editions such as Berberian [16]. In addition, seismography stations that record earthquakes sometimes publish their data on specific seismology websites. In this investigation, data from 11539 earthquakes was collected by the authors, which was recorded from 1904 to the end of November 2007 (over 103 years). Many of the recorded earthquakes were measured by either Mb or Ms magnitude scales and some of them were measured by both. Earthquakes that were measured neither by Mb nor Ms, were eliminated from the list. As a result, 4302 earthquakes were chosen to be analyzed. These earthquakes were consequently allocated to all 20 seismic regions of Iran, according to their occurrence location. The results are presented in Table 3. To facilitate the process, the converter relationships suggested by Mirzaei [15], derived from Table 1, were utilized to unify all the magnitude scales to conform to the Ms scale. The next step was to select earthquakes with a higher accuracy. The earthquakes that were recorded between 1904 and 1964 are less accurate than the ones recorded after 1964, due to the primitive nature of the equipment that existed before 1964. Therefore, to increase the accuracy of earthquake data used in this research, the earthquakes with a magnitude (original Ms or the converted Mb) greater than or equal to 5 were chosen from among those which occurred between 1904 and 1964. On the other hand, for earthquakes that occurred after 1964, those with a magnitude (original Ms or the converted Mb) greater than or equal to 4 were chosen. The total number of earthquakes meeting all criteria was 2007, which were allocated to one of the 20 seismic regions of Iran mentioned earlier.

In the next stage, earthquake magnitudes from 4 to 7 Richter were sorted incrementally by an interval of 0.5 Richter (Table 4) and 0.3 Richter (Table 5), and earthquakes were allocated to these magnitude intervals. The advantage of using 0.3 increments over 0.5 increments is that the amount of data is more ideal for the application of statistical tests. In the next stage, both ordinary linear regression and second order regression were applied to the data for all twenty seismic regions of Iran, including the data based on both 0.5 and 0.3 Richter increments. Due to the statistical tailor-made nature of the data based on 0.3 Richter increments, the results from this data set will be analyzed through statistical tests in a further

Table 3. The number of selected earthquakes for this study in 20 seismic areas of Iran.

Region	Number of Earthquakes	Region	Number of Earthquakes	Region	Number of Earthquakes
1	96	9	94	17	79
2	36	10	161	18	126
3	64	11	252	19	168
4	142	12	1861	20	133
5	58	13	217	All Iran's seismic regions	4302
6	203	14	123		
7	130	15	227		
8	78	16	54		

Table 4. The number of earthquakes allocated to each seismic area of Iran based on 0.5 Richter increments.

Number of Earthquakes with Magnitude Greater than or Equal to a Given Magnitude.							
Regions	$\geq$ **4**	$\geq$ **4.5**	$\geq$ **5**	$\geq$ **5.5**	$\geq$ **6**	$\geq$ **6.5**	$\geq$ **7**
1	68	32	22	9	2	1	1
2	22	11	5	2	2	1	0
3	33	15	6	5	3	2	1
4	69	44	21	14	10	6	4
5	42	26	16	13	5	2	1
6	86	44	28	21	10	3	1
7	61	40	25	16	8	4	1
8	39	27	21	14	9	3	1
9	46	34	29	18	8	3	0
10	83	41	18	7	2	1	0
11	113	58	37	21	8	8	2
12	799	388	168	89	22	6	0
13	93	41	22	15	6	1	0
14	63	47	21	12	6	5	3
15	111	75	47	25	18	13	8
16	26	22	15	13	12	7	4
17	33	26	18	11	6	1	1
18	73	58	36	25	16	8	4
19	100	75	50	33	15	11	7
20	47	30	15	5	1	0	0
All regions	2007	1134	620	368	169	86	39

Table 5. The number of earthquakes allocated to each seismic area of Iran based on 0.3 Richter increments.

Number of Earthquakes with Magnitude Greater than or Equal to a Given Magnitude											
Regions	$\geq$ **4**	$\geq$ **4.3**	$\geq$ **4.6**	$\geq$ **4.9**	$\geq$ **5.2**	$\geq$ **5.5**	$\geq$ **5.8**	$\geq$ **6.1**	$\geq$ **6.4**	$\geq$ **6.7**	$\geq$ **7**
1	68	45	31	25	15	9	4	2	1	1	1
2	22	17	11	8	4	2	2	1	1	1	0
3	33	22	14	6	5	5	3	2	2	1	1
4	69	52	35	22	17	14	11	10	6	6	4
5	42	34	25	16	15	13	7	4	2	1	1
6	86	60	39	32	24	21	11	7	4	2	1
7	61	53	34	27	19	16	11	6	4	4	1
8	39	34	27	21	18	14	10	9	5	1	1
9	46	35	34	30	23	18	10	8	5	2	0
10	83	51	40	20	13	7	3	2	1	0	0
11	113	77	57	41	30	21	12	8	8	3	2
12	799	499	363	212	131	89	40	19	8	4	0
13	93	53	41	27	19	15	6	6	6	1	0
14	63	52	37	28	15	12	8	5	5	3	3
15	111	86	65	53	37	25	21	18	13	11	8
16	26	22	20	17	14	13	12	11	7	6	4
17	33	27	23	18	15	11	6	4	2	1	1
18	73	62	48	38	30	25	17	11	8	7	4
19	100	90	69	58	42	33	19	14	12	10	7
20	47	39	28	17	9	5	2	1	0	0	0
All regions	2007	1410	1041	716	495	368	215	148	100	65	39

section. The results for the data based on 0.5 Richter increments are shown for comparison with the previous research, but will not be analyzed further in this study. In addition, the second order regression operation is introduced for the first time by the authors in this paper.

Results of OLS and Second Order Regression Operations

Tables 6 and 7 present the Gutenberg and Richter relationship for each area of Iran by means of the Ordinary Least Square method, using the data based on both 0.5 and 0.3 Richter increments. Tables 8 and 9, on the other hand, present the second order regression curves resulting from earthquake data. The curves are obtained by means of the second order regression, with data based on both 0.5 and 0.3 Richter increments.

Violation of Linear Regression Assumptions

Once the results of the OLS operation had been obtained for the data based on 0.3 increments, they were statistically tested against variance heteroscedasticity and autocorrelation.

Table 6. Resulting lines from the application of OLS for different seismic areas of Iran based on 0.5 Richter increments.

Regions	OLS Regression	Correlation Coefficient
1	$\log(N) = 4.599 - 0.682M$	0.959
2	$\log(N) = 3.413 - 0.533M$	0.962
3	$\log(N) = 3.303 - 0.472M$	0.973
4	$\log(N) = 3.454 - 0.412M$	0.991
5	$\log(N) = 3.895 - 0.543M$	0.967
6	$\log(N) = 4.490 - 0.613M$	0.951
7	$\log(N) = 4.155 - 0.561M$	0.952
8	$\log(N) = 3.758 - 0.504M$	0.917
9	$\log(N) = 3.621 - 0.458M$	0.914
10	$\log(N) = 5.171 - 0.797M$	0.992
11	$\log(N) = 4.261 - 0.546M$	0.967
12	$\log(N) = 6.356 - 0.836M$	0.977
13	$\log(N) = 4.900 - 0.715M$	0.939
14	$\log(N) = 3.655 - 0.461M$	0.982
15	$\log(N) = 3.574 - 0.383M$	0.993
16	$\log(N) = 2.469 - 0.252M$	0.937
17	$\log(N) = 3.947 - 0.562M$	0.914
18	$\log(N) = 3.628 - 0.418M$	0.975
19	$\log(N) = 3.672 - 0.404M$	0.986
20	$\log(N) = 5.127 - 0.824M$	0.944
All regions	$\log(N) = 5.615 - 0.567M$	0.995

Table 7. Resulting lines from the application of OLS for different seismic areas of Iran based on 0.3 Richter increments.

Regions	OLS Regression	Correlation Coefficient
1	$\log(N) = 4.692 - 0.697M$	0.969
2	$\log(N) = 3.599 - 0.566M$	0.956
3	$\log(N) = 3.439 - 0.505M$	0.965
4	$\log(N) = 3.368 - 0.397M$	0.981
5	$\log(N) = 4.068 - 0.578M$	0.957
6	$\log(N) = 4.486 - 0.612M$	0.964
7	$\log(N) = 4.047 - 0.538M$	0.951
8	$\log(N) = 3.872 - 0.523M$	0.874
9	$\log(N) = 3.643 - 0.462M$	0.915
10	$\log(N) = 5.248 - 0.811M$	0.990
11	$\log(N) = 4.381 - 0.569M$	0.982
12	$\log(N) = 6.435 - 0.849M$	0.979
13	$\log(N) = 4.457 - 0.618M$	0.930
14	$\log(N) = 3.752 - 0.483M$	0.983
15	$\log(N) = 3.559 - 0.381M$	0.994
16	$\log(N) = 2.437 - 0.247M$	0.948
17	$\log(N) = 3.895 - 0.552M$	0.952
18	$\log(N) = 3.603 - 0.417M$	0.985
19	$\log(N) = 3.700 - 0.408M$	0.986
20	$\log(N) = 5.130 - 0.820M$	0.966
All regions	$\log(N) = 5.614 - 0.566M$	0.996

Variance Heteroscedasticity

To find the existence of data variance heteroscedasticity, the parameter, λ, was determined for each of the 20 seismic regions. This parameter was subsequently compared with F from the F distribution (confidence level of 5%). Variance heteroscedasticity was discovered for the data obtained from regions 1, 2, 7, 8, 11, 13, 16, 17 and 18. To solve the problem defined before, four relationships between the variance and mathematical expectation of the squares of residuals (Equations 8 to 11) were examined for each of these regions, which were:

$$E(u_i^2) = \sigma^2 X_i,$$

$$E(u_i^2) = \sigma^2 X_i^2,$$

$$E(u_i^2) = \sigma^2 / X_i,$$

Table 8. Resulting second order regression curves for different seismic areas of Iran based on 0.5 Richter increments.

Regions	Second Order Regression	Correlation Coefficient
1	$\log(N) = 5.178 - 0.900M + 0.020M^2$	0.960
2	$\log(N) = 6.039 - 1.560M + 0.098M^2$	0.979
3	$\log(N) = 4.738 - 1.012M + 0.049M^2$	0.981
4	$\log(N) = 4.364 - 0.754M + 0.031M^2$	0.995
5	$\log(N) = 1.041 + 0.530M - 0.098M^2$	0.990
6	$\log(N) = 0.370 + 0.936M - 0.141M^2$	0.988
7	$\log(N) = 0.265 + 0.902M - 0.133M^2$	0.992
8	$\log(N) = -1.060 + 1.308M - 0.165M^2$	0.991
9	$\log(N) = -1.373 + 1.496M - 0.186M^2$	0.995
10	$\log(N) = 3.791 - 0.257M - 0.051M^2$	0.995
11	$\log(N) = 2.963 - 0.058M + 0.040M^2$	0.972
12	$\log(N) = 2.088 + 0.834M - 0.159M^2$	0.996
13	$\log(N) = -0.130 + 1.252M - 0.190M^2$	0.973
14	$\log(N) = 4.588 - 0.810M - 0.031M^2$	0.980
15	$\log(N) = 4.087 - 0.576M + 0.018M^2$	0.995
16	$\log(N) = 0.886 + 0.343M - 0.054M^2$	0.969
17	$\log(N) = 0.224 + 0.839M - 0.127M^2$	0.949
18	$\log(N) = 1.474 + 0.392M - 0.074M^2$	0.998
19	$\log(N) = 3.012 - 0.160M - 0.023M^2$	0.989
20	$\log(N) = -3.161 + 2.559M - 0.338M^2$	0.999
All regions	$\log(N) = 4.422 - 0.120M - 0.041M^2$	0.999

$$E(u_i^2) = \sigma^2/X_i^2.$$

The accuracy of the new lines, resulting from GLS, was determined by their correlation coefficient. Sometimes, one of these assumptions was more accurate than the rest. However, it was also probable that no assumption was more accurate than the original OLS regression line. In this case, the original line would be the 'best' regression line.

Autocorrelation

To investigate the existence of data autocorrelation, the data was examined by the Durbin-Watson d test. The calculated d factor was subsequently compared to d_U and d_L, which were obtained from a Durbin-Watson's table with a confidence level of 5%. In accordance with Table 2, data from regions 5, 6, 9, 12 and 17 have autocorrelation, and need to be investigated. To solve this issue, the Theil-Nagar ρ parameter was applied and new lines were produced by means of the GLS operation. Again, the new lines were compared with the original regression lines, in regard to their correlation coefficient.

The 'Best' Regression Lines of Iran's Seismic Regions

To complete the research, for each seismic region of Iran, the 'best' regression line was chosen from among all possible alternatives, in regard to their correlation coefficient. As explained before, the closer the correlation coefficient is to 1, the more precise the regression line is. In this comparison, all regression lines resulting from the OLS and GLS operations, which are concerned with solving either variance hetroscedasticity or autocorrelation, were taken into account. All these lines were resulted from regression operations on data based on 0.3 Richter magnitude increments. Finally, the 'best' regression lines for all Iran's seismic regions are presented in Table 10.

Application of the Poisson Distribution

By using the 'best' regression line in conjunction with the Poisson distribution, the probability of an earthquake occurrence of a certain Ms magnitude over a fixed time period can be determined.

Table 9. Resulting second order regression curves for different seismic areas of Iran based on 0.3 Richter increments.

Regions	Second Order Regression	Correlation Coefficient
1	$\log(N) = 4.758 - 0.723M + 0.002M^2$	0.969
2	$\log(N) = 6.239 - 1.579M + 0.095M^2$	0.972
3	$\log(N) = 5.467 - 1.265M + 0.069M^2$	0.978
4	$\log(N) = 4.620 - 0.866M + 0.043M^2$	0.989
5	$\log(N) = 0.903 + 0.609M - 0.108M^2$	0.980
6	$\log(N) = 0.754 + 0.786M - 0.127M^2$	0.993
7	$\log(N) = 1.168 + 0.541M - 0.098M^2$	0.974
8	$\log(N) = -1.637 + 1.541M - 0.188M^2$	0.953
9	$\log(N) = -1.127 + 1.369M - 0.171M^2$	0.988
10	$\log(N) = 2.831 + 0.140M - 0.091M^2$	0.996
11	$\log(N) = 2.674 + 0.071M - 0.058M^2$	0.989
12	$\log(N) = 2.140 + 0.800M - 0.154M^2$	0.998
13	$\log(N) = 2.083 + 0.293M - 0.085M^2$	0.940
14	$\log(N) = 4.072 - 0.586M + 0.007M^2$	0.984
15	$\log(N) = 4.048 - 0.564M + 0.017M^2$	0.996
16	$\log(N) = 0.930 + 0.318M - 0.051M^2$	0.977
17	$\log(N) = 0.305 + 0.794M - 0.122M^2$	0.984
18	$\log(N) = 2.138 + 0.132M - 0.050M^2$	0.995
19	$\log(N) = 3.215 - 0.226M - 0.017M^2$	0.987
20	$\log(N) = -1.103 + 1.695M - 0.249M^2$	0.998
All regions	$\log(N) = 4.514 - 0.154M - 0.037M^2$	0.999

A Problem: Finding an Earthquake Occurrence Probability

In the following analysis, we assume a project is to be undertaken in a given city in Western Azerbaijan, within seismic region 18 in Iran and is designed based on a 6.7 Richter earthquake. The designers and managers intend to find out the possibility of an earthquake with a magnitude of 6.7 Ms Richter targeting the project within its life expectancy of 30 years.

The Solution by Means of the 'Best' Regression Line

To find the solution, we need to find the value of N from the 'best' regression line of region 18. According to Table 10, the 'best' regression line is calculated as follows:

$$\log(N) = 3.603 - 0.417\ M.$$

Therefore:

$$\log(N) = 3.603 - 0.417 \times 6.7 \rightarrow \log(N) = 0.8091 \rightarrow N = 6.44.$$

As suggested by Bargi [1], N must be divided by 100 to be placed within the Poisson distribution. Thus:

$$N = 0.0644,$$

$$P_T(t) = 1 - e^{-Nt},$$

$$P_T(t) = 1 - e^{-(0.0644)(30)} \rightarrow P_T(t) = 0.855.$$

The result from the calculations shows that the possibility of an earthquake occurrence with a magnitude of 6.7 Richter in the 18th region, over 30 years, is 85.5%, which would be considered very crucial by decision makers. As a result, this fact should be

Table 10. The 'Best' regression line of Iran's seismic regions.

Regions	The 'Best' Line	Correlation Coefficient
1	$\log(N) = 4.692 - 0.697M$	0.969
2	$\log(N) = 3.684 - 0.582M$	$1.011(\text{Het } \sigma \sim X^2)^*$
3	$\log(N) = 3.439 - 0.505M$	0.965
4	$\log(N) = 3.368 - 0.397M$	0.981
5	$\log(N) = 4.068 - 0.578M$	0.957
6	$\log(N) = 4.486 - 0.612M$	0.964
7	$\log(N) = 4.120 - 0.551M$	$0.998(\text{Het } \sigma \sim (1/X^2))$
8	$\log(N) = 4.012 - 0.548M$	$0.959(\text{Het } \sigma \sim (1/X^2))$
9	$\log(N) = 3.643 - 0.462M$	0.915
10	$\log(N) = 5.248 - 0.811M$	0.990
11	$\log(N) = 4.424 - 0.577M$	$1.009(\text{Het } \sigma \sim (1/X^2))$
12	$\log(N) = 6.435 - 0.849M$	0.979
13	$\log(N) = 4.509 - 0.627M$	$0.958(\text{Het } \sigma \sim (1/X^2))$
14	$\log(N) = 3.752 - 0.483M$	0.983
15	$\log(N) = 3.559 - 0.381M$	0.994
16	$\log(N) = 2.475 - 0.253M$	$1.001(\text{Het } \sigma \sim (1/X^2))$
17	$\log(N) = 3.986 - 0.568M$	$1.008(\text{Het } \sigma \sim (1/X^2))$
18	$\log(N) = 3.603 - 0.417M$	0.985
19	$\log(N) = 3700 - 0.408M$	0.986
20	$\log(N) = 5.130 - 0.820M$	0.966
All regions	$\log(N) = 5.614 - 0.566M$	0.996

* Het means the line is resulted from the application of GLS to solve heteroscedasticity.

taken into consideration when this project is being undertaken.

The Solution by Means of the Second Order Regression Curve

In the case of using the second order regression curve, which is not as accurate as the 'best' line, in regard to its correlation coefficient, the result is as follows:

$$\log(N) = 2.138 + 0.132M - 0.050M^2 \log(N)$$

$$= 2.138 + 0.132 \times 6.7 - 0.050 \times 6.7^2,$$

$$\log(N) = 07779 \rightarrow N = 6.00 \rightarrow N = 0.06,$$

$$P_T(t) = 1 - e^{-Nt},$$

$$P_T(t) = 1 - e^{-(0.06)(30)} \rightarrow P_T(t) = 0.835.$$

Interpretation of Results

A comparison between the results from the application of both the regression line and second order regression curve are presented in Figure 2. Generally, the result from application of the second order regression validates the previous solution from application of the 'best' regression line. It is shown that either of these methods is applicable, since the difference between the results is negligible. On the other hand, the linear regression line has been tested with statistical methods against heteroscedasticity and auto-correlation, while the second order curve is typical and untested. As a result, the authors suggest that the 'best' line is more reliable than the second order curve because it has been optimized in this research.

Finally, the result, which is a coefficient, can be considered as the earthquake occurrence risk for a

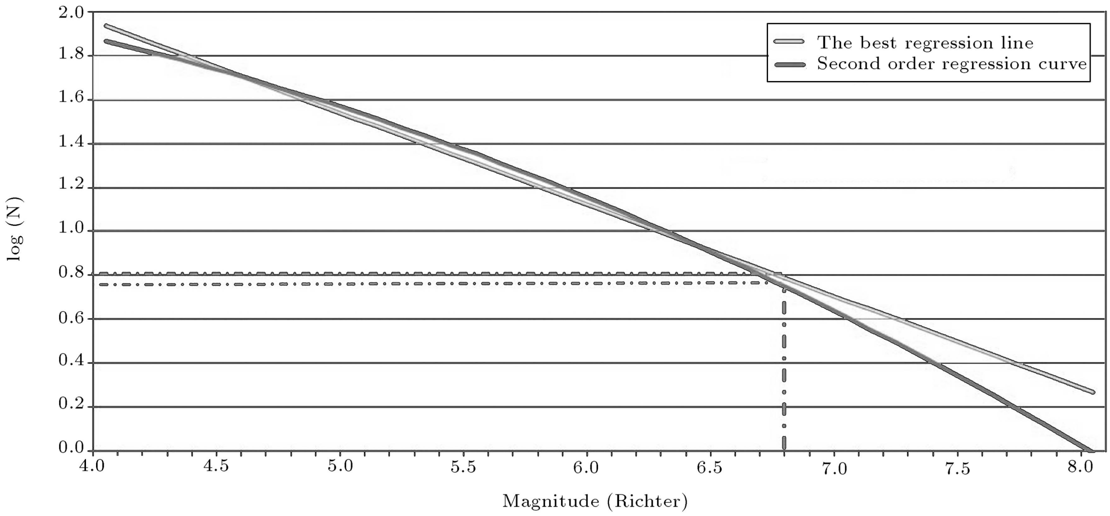

Figure 2. Comparison between the results of the 'best' regression line and the second order regression curve in the aforementioned example.

particular project. It should be taken into account that the result will vary for projects undertaken in another region of Iran, even when they are designed based on the same magnitude of earthquake due to the difference between 'best' lines.

SUMMARY

Statistical research on natural hazards influencing a project, which include earthquakes and floods as the most important in Iran, must be continually updated. In general, the absence of this type of research, which should be undertaken by a construction management researcher, will result in a lack of mutual understanding between engineers and managers. Therefore, construction managers are encouraged to scientifically analyze the natural risks pertaining to each project, in order to obtain better standards by which to evaluate their projects and support their decisions. In this research, the possibility of an earthquake occurrence during the life time of a particular project is determined. This is achieved by applying the Gutenberg-Richter linear relationship in conjunction with the Poisson distribution. To find the 'best' linear relationship between the magnitude of an earthquake and the number of its occurrence, the Ordinary Least Square (OLS) regression is used. The results are checked against two linear regression assumption violations: variance heteroscedasticity and autocorrelation. Generalized Least Square (GLS) is applied to solve the problem, once the possible violation has been diagnosed. As a result, the design earthquake occurrence risk can be determined for a given construction project in a specific location in Iran. This is achieved through application of the Poisson distribution, together with the 'best' regression line, and by knowing the design earthquake's magnitude and project life time.

FURTHER STUDIES

Below are some alternatives and recommendations for undertaking further research.

1. Using another relationship to relate the magnitude of an earthquake with the number of that earthquakes' occurrence during a fixed time, instead of the Gutenberg-Richter linear relationship applied in this research.
2. Providing other relationships, which include new parameters other than "magnitude" and "number of earthquake occurrences", i.e. duration of an earthquake, length of a fault and so forth.
3. Analyzing more unified data, which are measurable under the same scale, i.e. Mw.
4. Applying more statistical tests, in order to increase the scientific validity of the research.
5. Using other distributions, such as the Gamble distribution, instead of the Poisson distribution used in this research.

REFERENCES

1. Bargi, K., *Fundamentals of Earthquake Engineering*, Tehran University Publisher (2000).
2. Fattahi, M., Nazari, H., Bateman, M.D., Meyer, B., Sébrier, M., Talebian, M., Le Dortz, K., Foroutan, M,, Ahmadi Givi, F. and Ghorashi, M. "Refining the OSL age of the last earthquake on the Deshir fault, central Iran", *Quaternary Geochronology* (2009), doi: 10.1016/j.quageo.2009.04.005

3. Banisadr, M. "The seismicity of Iran (1900-1969)", *Tech. Res. & Standard Bureau.* (1971).
4. Berberian, M. "Preliminary map of epicenters and focal depth", *Geological Survey of Iran* (1973).
5. Nowroozi, A. "Seismotectonic provinces of Iran", *Bull. Seism. Soc.*, **66**, pp. 1249-1276 (1976).
6. Moinfar, A.A., Naderzadeh, A. and Maleki, E., *A New Seismic Hazard Map for Implementation in the National Physical Planning of Iran*, Kluwer Academic Publishers, pp. 289-296 (2000).
7. Stocklin, J. "Structural history and tectonics of Iran: A review", *Bull Am. Assoc Petrol. Geol.*, **52**, pp. 1229-1258 (1968).
8. Takin, M. "Iranian geology and continental drift in the middle east", *Nature*, **235**(5334), pp. 147-150 (1972).
9. Tavakoli, B. and Ghafoori Ashtiany, M. "Seismic hazard assessment of Iran", Online, assessed 1 Jun. 2003, Url:www.IIEES.com
10. Gutenberg, B. and Richter, C. "Seismicity of the earth", *Geol. Soc. Am. Bull.*, **56**, pp. 603-667 (1945).
11. Walpole, R., *An Introduction to Applied Statistics*, Science and Thechnology University Publisher (1992).
12. Wallace, T.C., *Global Modern Seismology*, Mashhad, Ferdowsi University Publisher (2002).
13. Mirzaei, N., *Basic Parameters For Iran's Earthquakes*, Danesh Negar Publisher (2002).
14. Judge, G.G., Griffiths, W.E., Carter Hill, R. Lütkepohl, H. and Lee, T.C., *Introduction to the Theory and Practice of Econometrics*, John Wiley & Sons Ltd (1988).
15. Gujarati, D.N., *Basic Econometrics*, McGraw-Hill International Edition (1995).
16. Berberian, M., *Natural Hazards and the First Earthquake Catalogue of Iran*, Ehya Ketab Publisher (1995).