Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 56 (2): 17–28 (2019)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

Mathematical Modeling of Damaging Earthquakes in Pakistan

Asghar Ali[1*], Rashida Hussain[2], Javeed Shahid[2], and Faria Hamid[2]

[1]Department of Mathematics, Mirpur University of Science and Technology (MUST), Mirpur-10250 (AJK), Pakistan
[2]Department of Disaster Management University of Baluchistan Quetta, Pakistan

Abstract: Earthquakes are totally unpremeditated events and cannot be fully anticipated in advance. However, by analyzing historical knowledge and formulating mathematical models will offer researchers with an additional legitimate estimate of future losses. Earthquakes are the most deadly disaster in the world. The purpose of present study is to estimate the earthquake casualties and resulted damage costs for possible future earthquakes in Pakistan. On the basis of examination new mathematical models have been developed for estimation of earthquake casualties and related damage costs through multiple linear regression using matrices with correlation coefficient $\alpha = 0.01$. This study considers twenty-two most damaging earthquake that hit different regions of Pakistan from 1909 to 2017. The resulting models were multiple linear regression models explaining earthquake casualties and a total cost of damage through five independent variables magnitude, intensity, and depth of focus, location of epicenter and interval of earthquake.

Keywords: Mathematical Model, Earthquake, Depth of focus, Epicenter.

1. INTRODUCTION

The Islamic Republic of Pakistan has faced much in its 70-year life, in terms of man- made as well as natural disasters. Pakistan faces a severe threat from natural disasters such as flood, Tsunami, Storm, and Earthquakes. However, the worst disaster among them for Pakistan may be the Earthquake. Pakistan is one of the most seismically active countries in world, being crossed by several major faults. Historically Pakistan is indeed situated near extremely active fault line that could put risk to more than one hundred seventy million individual living in the country.

Pakistan geologically overlaps with the Indian and the Eurasian tectonic plates wherever its Sindh and Punjab provinces lie on the north-western corner of the Indian plate while Baluchistan and most of the Khyber-Pakhtunkhwa lie within the Eurasian plate which mainly comprises the Iranian plateau, some parts of the Middle East and the Central Asia [2]. The northern areas and Azad Kashmir lie mainly in Central Asia along the edge of the Indian plate and hence are prone to violent earthquakes where the two tectonic plates collide [2].

We limit this study only to the destructive earthquakes with magnitude 5 and higher. Data on casualties and cost of destructions due to these earthquakes were collected from State Disaster Management authority Muzaffarabad [10]. The information about the characteristics of the earthquakes was gathered from Pakistan Disaster Management authority Islamabad [11]. We considered the zone where the epicenter was located. Pakistan Meteorological Department (PMD) presently has a network of eleven seismic stations. By using recorded data, the seismicity and zoning maps of Pakistan have been developed by PMD [12]. Following different zones defined by PMD were considered in present study:

- Zone 1 - Kohistan-Kashmir
- Zone 2 - Quetta-Sibi
- Zone 3 - Southern Baluchistan
- Zone 4 - Hindu Kosh
- Zone 5 - Makran Coast

Received: January 2019; Accepted: June 2019
* Corresponding Author: : Asghar Ali <drali@must.edu.pk>

- Zone 6 - Runn of Kuchch
- Zone 7 - Sind-Punjab
- Zone 8 - Upper Punjab-NWFP
- Zone 9 - Western Baluchistan
- Zone 10 - Indian Kashmir
- Zone 11 - North Western Afghanistan-Tajikistan Border Region

The main objective of the study was to formulate such mathematical models which could estimate earthquake casualties and destruction cost. These models will be of great worth in formulating new programs for mitigation of earthquake hazards. According to the target of the study we try to find answers to the following questions:

- Question 1. To what extent are the major earthquake that strikes Pakistan be characterized in terms of intensity, magnitude, depth of focus, location of the epicenter and span?
- Question 2. How much destruction can earthquakes produce in terms of death, injuries, families affected and cost of destruction?
- Question 3. Which mathematical models can be formulated through regression analysis using matrices that approximately describe the earthquake casualties and cost of destruction due to destructive earthquakes?
- Question 4. How much significant are the developed mathematical models to assess the possible destructions due to an earthquake event?

1.1. Literature Review

Earthquake destruction depends upon various factors such as magnitude, intensity, interval, depth of focus etc. Mostly, earthquake destruction depends upon area where it occurs, if it hit a populated area then there will be more destruction than one that hits an un-populated area. The death tolls in earthquake depend on three main factors [1]: structural collapses, nonstructural cause and follow on disaster. According to a research study in 2016 [2], Pakistan and adjoining regions are divided into 14 seismogenic zones. Seismicity of each zone is studied considering also the major cities in the respective zone and type of infrastructure which is mainly responsible for earthquake disaster rather than earth- quake itself. There are four assumptions for multiple linear regressions [4], which are needed to be satisfied for truth worthy results. These assumptions are linearity, reliability of measurement, homoscedasticity, and normality. Linear regression is one of the fundamental models in statistics used to determine the relationship between dependent and independent variables [5]. An extension of this model, namely multiple linear regressions, is used to represent the relationship between a dependent variable and several independent variables. This study is based on multiple linear regression models using matrix notation and analyzing the model using a script approach with MATLAB.

The analysis of local network of earthquake in Pakistan is referred from the PMD (Pakistan Metrological Department) [6] covers a period of 1905 to 2007 and its comparison is done with the Global Catalog of National Earthquake Information Center (NEIC). Pakistan and adjoining region lying between longitude 60^o E to 78^o E and latitude 20^o N to 45^o N is selected for the study. From the NEIC catalog, a sub-catalog is obtained for Pakistan and surrounding region, it contains 8635 events from 1963 to 2004. Finally, to get a rather homogeneous catalog for the region, the two catalogs were merged by considering different data properties and different data analysis techniques prepared by different data collection agencies. The results of this study can be employed for the earthquake prediction research.

Rizwan et al. [7], analyzed available previous seismic data in terms of maximum annual recorded intensity of earthquakes on Richter scale at various locations of Pakistan with a view to know future earthquakes intensity and return periods required for making decision regarding the design of lifeline systems and other structures.

In 2014, Urrutia et al. [8] formulated mathematical models that could estimate the earthquake casualties and destruction costs through regression analysis using matrices. Author considered 30 earthquakes that hit Philippines from inclusive year 1968-2012. The proposed mathematical models with predictors intensity x_1, magnitude x_2, depth x_3, epicenter x_4 and duration x_5 and dependent variables deaths y_1, injured y_2, cost of damage y_3, families affected y_4, as follows:

$$\ln \hat{y}_1 = -13.405 - 0.001x_1 + 2.214x_2 - 0.013x_3 + 0.374x_4 - 0.001x_5$$

$$ln\,\hat{y}_2 = -13.426 + 0.360x_1 + 1.899x_2 + 0.015x_3 + 0.444x_4 - 0.001x_5$$
$$ln\,\hat{y}_3 = -16.905 + 0.083x_1 + 2.577x_2 - 0.043x_3 + 0.630x_4 + 0.002x_5$$
$$ln\,\hat{y}_4 = -17.075 + 0.234x_1 + 2.520x_2 - 0.033x_3 + 0.383x_4 + 0.001x_5$$

Allison, Nicole, and Waters in 2016 [9] presented multiple linear regression models that explained total destruction resulting from an earthquake through four independent variables: whether or not a tsunami occurred (tsunami − dummy), whether or not the earthquake occurred in a developed nation (developed − dummy), intensity (intensity) and number of injuries (total − injuries)

$$Y = -19971 + 5335.98x_1 + 4203.51x_2 + 2333.30x_3 + 0.21428x_4$$

Where coefficient is in millions of dollars in 2015.

Multiple regression for one of the models that predicts death is the combination of magnitude, ln (magnitude) and epicenter with highest R2=0.563 and that is given below:

$$\hat{y}1 = 59578.966 + 11182.497x_2 - 71139.37\ln(x_2) + 167.795x_4$$

The model for predicting injuries with highest $R_2 = 0.685$ is the combination of magnitude and ln (epicenter) with ln(injured) are:

$$\ln(\hat{y}2) = 16.697 + 2.664x_2 + 1.241\ln(x_4)$$

Model that predicts cost of destruction is the combination of ln (intensity) and epicenter with highest $R_2 = 0.611$ is:

$$\ln(\hat{y}3) = -18.787 + 9.477\ln(x_2) + 0.719x_4$$

Timothy [13] represents the development of mathematical model for the estimation of required maintenance for a homogenous facilities portfolio using multiple linear regressions. The study shows how a facilities manager can take historical facility attribute data from a maintenance work-order system and develop a prediction equation by using multiple regression analysis for predicting required maintenance. The derived prediction equations results were compared with those of three popular models discussed in the research. The prediction equations results strongly correlated with all three of the models. Present study has reviewed these models and applied to collected data of earthquake that occurred in Pakistan from 1909 to 2017.

2. MATERIALS AND METHODS

2.1 Data Collection

The first step in this study was to persuade variables that may be included to develop the model. Different variables as shown in block diagram Fig.1 were used to formulate mathematical models that will evaluate the casualties and total cost of destruction. All 22 earthquakes with magnitude 5 or more occurred in Pakistan between 1909 and 2017 were examined. Data about most destructive earthquake that used in this research are presented in Table 1 & 2.

2.1.1. Relation between dependent and independent variables

The relationship between the dependent and independent variables is obtained by using Pearson's

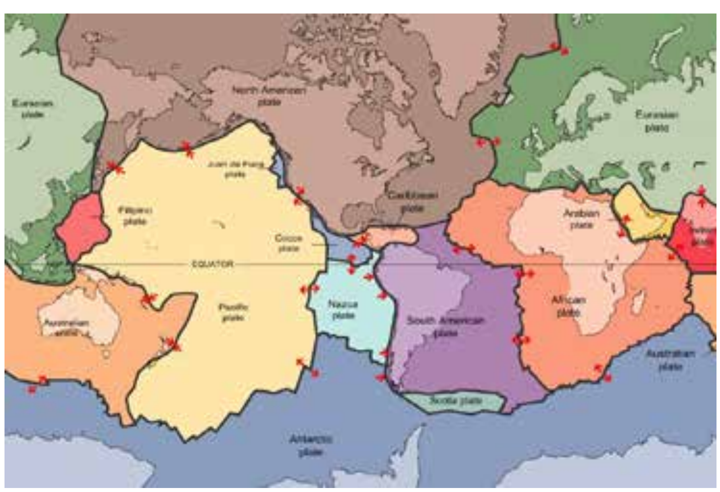

Fig. 1. Tectonic Plates

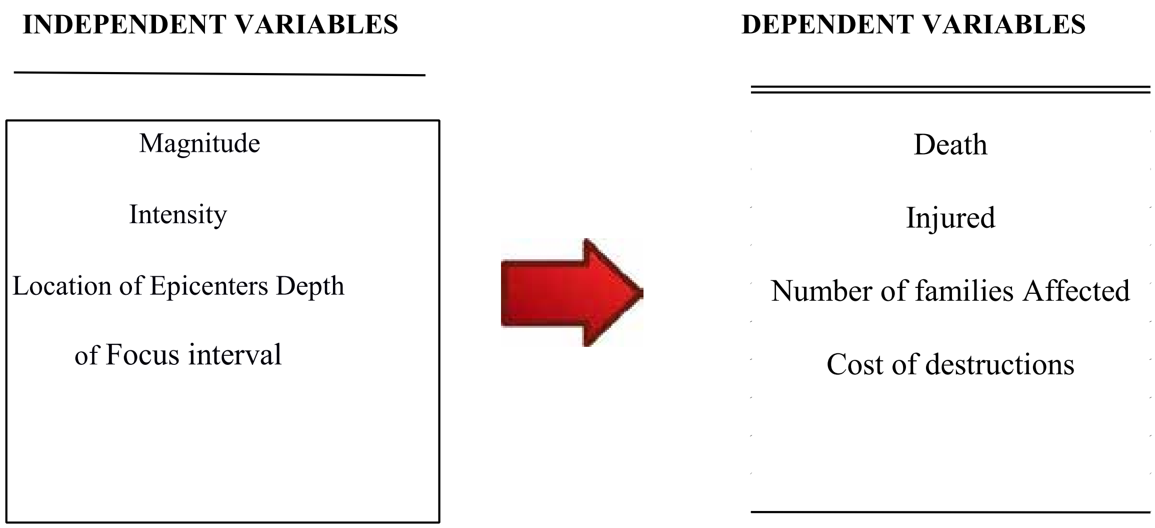

Fig. 2. Block Diagram

Table 1. Data for independent variables

No	Year of earthquake occurrence	Magnitude (x1i)	Intensity (x2i)	Interval (x3i)	Epicenter (x4i)	Depth (x5i)
1	1909	7.1	9	25	3	60
2	1935	7.7	10	180	2	17
3	1945	8	10	60	5	25
4	1966	6.7	7	35	3	10
5	1966	6.1	6.9	25	1	250
6	1971	6.8	7	20	7	230
7	1974	6.6	7	30	4	180
8	1974	7.4	6	22	4	22
9	1990	6.2	6	34	3	7250
10	1995	7.5	5.2	31	5	33
11	1997	6.9	7	30	3	33
12	2001	7.7	10	120	6	22
13	2002	5.8	5	37	10	33
14	2002	5.5	75	28	10	31
15	2004	5.5	5	27	1	10
16	2005	7.6	8	45	8	26
17	2008	7	6	25	2	15
18	2011	5	7	25	3	54
19	2011	7.4	7	20	6	300
20	2013	7.7	7	35	6	83
21	2013	5.6	7	15	3	15
22	2016	6.30	5	60	11	204.40

Table 2. Data for Casualties

No	Year of earthquake occurrence	Deaths (y1)	Injuries (y2)	Affected Families (y3)	Total cost of estruction (y4) (in rupees)
1	1909	100	72	2500	250,000
2	1935	60,000	4000	6000	600000000
3	1945	2000	3000	3000	200000000
4	1966	12	150	4860	486000000
5	1966	12	15	1300	130000000
6	1971	100	12	1000	1000,000,00
7	1974	4	10	1	100,000
8	1974	5300	1700	4400	440000000
9	1990	11	40	10	1000,000
10	1995	4	6	12	1200,000
11	1997	100	57	500	50000000
12	2001	2030	166800	400,000	796320000000
13	2002	17	30	1500	150000,000
14	2002	23	60	15	1500,000
15	2004	24	40	1	100,000
16	2005	87,351	75266	314474	387100000000
17	2008	216	370	100	10000,000
18	2011	4	5	200	20000000
19	2011	4	5	200	20000000
20	2013	34	80	10,000	1000,000,000
21	2013	800	700	1000	1000,000,000
22	2016	4	100	3	5600,000

Table 3. Pearson's coefficient of correlation

Intensity, Parameter	Magnitude	Depth of focus	Location of epicenters	Time interval	Coefficients of correlation
Death(y1)	0.378	0.339	-0.226	0.060	0.491*
Injured(y2)	0.469*	0.325	-0.205	0.160	0.452*
Affected Families(y3)	0.441*	0.341	-0.214	0.206	0.338
Total destruction Cost(y4)	0.453*	0.314	-0.197	0.168	0.431*

Table 4. Pearson's coefficient of correlation

Parameters	ln(death)	ln(injuries)	ln(families)	ln(total destruction)
Intensity	0.607**	0.580**	0.653**	0.565**
Magnitude	0.508*	0.500*	0.489*	0.443*
Depth	-0.483*	-0.498*	-0.316	-0.160
Epicenter	-0.086	0.073	0.082	0.246
Time interval	0.536*	0.576**	0.355	0.394

coefficient of correlation, as shown in the Table 3. Data shows that the magnitude of considered earthquakes is not significantly correlated with any of the dependent variables. Unlike the magnitude, the intensity of considered earthquakes is significantly correlated with injuries, number of affected families and total cost of destruction at 0.05 levels. The coefficient of correlation for depth and location of epicenter are not significantly correlated with any of the dependent variables but time interval is significantly correlated with death, injuries and total cost of destruction. This also shows that all the dependent variables correlate with one or two independent variables. Since all assumption for the multiple linear regressions does not meet, therefore to formulate the mathematical models that could estimate the causalities and destruction costs the dependent variables were transformed.

Table 4 shows that the transformed variable death is significantly correlated with magnitude, intensity and time interval. The transformed variable injuries are significantly correlated with intensity, and magnitude. The transformed variable affected families are significantly correlated with intensity and magnitude. The transformed variable total cost of destruction is significantly correlated with intensity and magnitude. After transformation all the assumption for the multiple linear regression analysis were satisfied. The assumption for normality is tested by using (P-P) plot, the homoscedasticity was checked by plotting the predicted values and residuals on a scatter plot, and the multi-collinearity was tested by the variance inflation factor (VIF) value.

2.1.2. Multiple linear regression by using matrices

Multiple linear regression analysis is used to investigate the relationship between dependent and independent variables when the number of variables exceed by two. The knowledge of matrix theory can facilitate the mathematical manipulation considerably. Relationship between k independent variable $x_1, x_2..., x_k$ and n observations $y_1, y_2, ..., y_n$, each of which can be expressed by equation.

$$y_i = \beta 0 + \beta 1x1i + \beta 2x2i + ... + \beta kxki + \epsilon .$$

Where β determines the contribution of independent variable X's and ϵ is random error. This model essentially represents n equations describing how the response values are generated in the scientific process. Using matrix notation, we can write the following equation:

$$y = X\beta + \epsilon ,$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Then the least squares method for estimation of β, involves finding b. The criterion is to minimize a sum of squares of residuals, which can be written as:

$$SSE = \epsilon\epsilon'$$
$$SSE = (y - b)'(y - X)$$

is minimized. This minimization process is involves for solving b

$$\begin{aligned} SSE &= (y - Xb)'(y - X) \\ &= (y' - b'X')(y - Xb) \\ &= y'y - y'Xb - b'X'y + b'X'Xb \end{aligned}$$

by using identity

$$b'X'y = y'Xb$$
$$\begin{aligned} SSE &= y'y - b'y'X - b'X'y + b'X'Xb \\ &= y'y - 2`X'y + b'X'Xb \end{aligned}$$

$$\frac{\partial}{\partial b}(SSE) = 0$$

$$\frac{\partial}{\partial b}(SSE) = 0 - 2X'y + 2b'X'y$$

$$bX'X - X'y = 0$$
$$X'Xb = X'y$$

On the assumption that the inverse matrix exists, the equations have a unique solution, which is the vector of ordinary least-squares estimates. The result reduces to the solution of b in:

$$(X'X)b = X'y$$

Notice the nature of the X matrix. Apart from the initial element, the ith row represents the x-values that give rise to the response y_1. Writing:

$$A = X X' = \begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} & \cdots & \sum_{i=1}^{n} x_{ki} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{1i}{}^{2} & \sum_{i=1}^{n} x_{1i} \; x_{2i} & \cdots & \sum_{i=1}^{n} x_{1i} \; x_{ki} \\ . & . & . & \cdots & \\ . & . & . & \cdots & \\ . & . & . & \cdots & \\ \sum_{i=1}^{n} x_{ki} & \sum_{i=1}^{n} x_{ki} x_{1i} & \sum_{i=1}^{n} x_{ki} \; x_{2i} & \cdots & \sum_{i=1}^{n} x_{ki}{}^{2} \end{bmatrix}$$

$$g = X'y = \begin{bmatrix} g_0 = \sum_{i=1}^{n} y_i \\ g_1 = \sum_{i=1}^{n} x_{1i} \; y_i \\ . \\ . \\ . \\ g_k = \sum_{i=1}^{n} x_{ki} \; y_i \end{bmatrix}$$

allows the normal equation to be put in the matrix form

$$Ab = g$$

If the matrix A is non-singular, we write the solution for the regression coefficient as

$$b = A^{-1} g = (X`X)^{-1} x'y$$

Thus, we can obtain the prediction equation or regression equation by solving a set of k + 1 equation in a like numbers of unknowns. This involves the inversion of the $k + 1$ by $k + 1$ matrix $X`X$ [14]

2.1.3. Mathematical model for estimation of death

By using the relation

$$b = X'y(X'X)^{-1}$$

and following matrix was obtained:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} 99.05 \\ 749.88 \\ 694.909 \\ 5531.56 \\ 473.83 \\ 549.69 \end{bmatrix} \begin{bmatrix} 22 & 153.1 & 147.6 & 1903.3 & 109 & 929 \\ 153.1 & 119.65 & 1044.29 & 12733.6 & 733.9 & 7234.7 \\ 147.6 & 1044.29 & 1008.2 & 12651.62 & 736.1 & 6527.7 \\ 1903.4 & 12733.6 & 12651.62 & 362920.36 & 10073.4 & 62945 \\ 109 & 733.7 & 736.1 & 10073.4 & 733 & 4582 \\ 929 & 7234.7 & 6527.7 & 62945 & 4582 & 69283 \end{bmatrix}^{-1}$$

$b_0 = -4.773$; $b_1 = 0.616$; $b_2 = 0.810$; $b_3 = -0.012$; $b_4 = 0.013$; $b_5 = 0.013$

Hence, the number of deaths caused by an earthquake event can be computed by using the regression equation:

$$ln\hat{y}_1 = -4.773 + 0.616x_1 + 0.810x_2 - 0.012x_3 + 0.013x_4 + 0.013x_5$$

Table 5. Model Summary

Model	R	R Square	Adjusted R square	Standard error of the estimate
1	0.766	0.587	0.458	2.26762

Table 6. ANOVA

Regression	Sum of Squares	df	mean squares	F	sig.
Regression	117.056	5	23.411	4.553	0.009
Residual	82.273	16	5.142		

Table 7. Model Summary

Model	R	R Square	Adjusted R square	Standard error of the estimate
1	0.790	0.624	0.506	2.04509

Table 8. ANOVA

Regression	Sum of Squares	df	mean squares	F	sig.
Regression	110.968	5	22.194	5.306	0.005
Residual	66.918	16	4.182		
Total	177.887	21			

Table 9. Model Summary

Model	R	R Square	Adjusted R square	Standard error of the estimate
1	0.757	0.572	0.439	2.6748

Table 10. ANOVA

Regression	Sum of Squares	df	mean squares	F	sig.
Regression	153.285	5	30.657	4.285	0.012
Residual	114.475	16	7.155		
Total	267.760	21			

Table 11. Model Summary

Model	R	R Square	Adjusted R square	Standard error of the estimate
1	0.699	0.489	0.329	3.62135

Table 12. ANOVA

Regression	Sum of Squares	df	mean squares	F	sig.
Regression	200.420	5	40.084	3.057	0.040
Residual	209.827	16	13.114		
Total	410.247	21			

2.1.4. SPSS output for multiple regression

SPSS output yields that the coefficient of determination R2 = 0.587 which implied that 58.70% of the variation in the number of death during an earthquake is explained by the regression equation and is significant with p-value of 0.009.

2.1.5. Mathematical model for estimation of injuries

By using the relation

$$b = X'y(XX)^{-1}$$

the following Matrix was obtained:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} 99.05 \\ 749.88 \\ 694.909 \\ 5531.56 \\ 473.83 \\ 549.69 \end{bmatrix} \begin{bmatrix} 22 & 153.1 & 147.6 & 1903.3 & 109 & 929 \\ 153.1 & 119.65 & 1044.29 & 12733.6 & 733.9 & 7234.7 \\ 147.6 & 1044.29 & 1008.2 & 12651.62 & 736.1 & 6527.7 \\ 1903.4 & 12733.6 & 12651.62 & 362920.36 & 10073.4 & 62945 \\ 109 & 733.7 & 736.1 & 10073.4 & 733 & 4582 \\ 929 & 7234.7 & 6527.7 & 62945 & 4582 & 69283 \end{bmatrix}^{-1}$$

Then the regression coefficients were obtained by using MATLAB

$$b_0 = -3.754;\ b_1 = 0.586;\ b_2 = 0.621;\ b_3 = -0.012;\ b_4 = 0.172;\ b_5 = 0.016$$

Hence the number of injured persons resulting from an earthquake event can be computed by using the regression equation

$$ln\hat{y}_2 = -3.754 + 0.586x_1 + 0.621x_2 - 0.012x_3 + 0.172x_4 + 0.016x_5$$

2.1.6. SPSS output for multiple regression

SPSS output yields that the coefficient of determination R_2 = 0.624 which implied that 62.40% of the variation in the number of death during an earthquake is explained by the regression equation and is significant with p-values of 0.005.

2.1.7. Mathematical model for estimation of affected families

By using the relation, the following matrix was obtained

$$b = X'y(X'X)^{-1}$$

the following matrix was obtained:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} 134.08 \\ 1011.739 \\ 933.45 \\ 9296.62 \\ 682.88 \\ 6667.86 \end{bmatrix} \begin{bmatrix} 22 & 153.1 & 147.6 & 1903.3 & 109 & 929 \\ 153.1 & 119.65 & 1044.29 & 12733.6 & 733.9 & 7234.7 \\ 147.6 & 1044.29 & 1008.2 & 12651.62 & 736.1 & 6527.7 \\ 1903.4 & 12733.6 & 12651.62 & 362920.36 & 10073.4 & 62945 \\ 109 & 733.7 & 736.1 & 10073.4 & 733 & 4582 \\ 929 & 7234.7 & 6527.7 & 62945 & 4582 & 69283 \end{bmatrix}^{-1}$$

Then regression coefficient was obtained by using MATLAB

$b_0 = -8.401; b_1 = 1.595; b_2 = 0.510; b_3 = -0.010; \quad b_4 = 0.318; b_5 = -0.018$

Hence the number of affected families resulting by an earthquake event can be computed by using the regression equation.

$$ln\hat{y}_3 = -8.401 + 1.595x_1 + 0.510x_2 - 0.010x_3 + 0.318x_4 - 0.018x_5$$

2.1.8. Mathematical model for estimating total cost of damage

By using the relation

$$b = X'y(X'X)^{-1}$$

The following matrix was obtained

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} 384.52 \\ 2760.202 \\ 2617.824 \\ 31826.776 \\ 1974.33 \\ 17619.92 \end{bmatrix} \begin{bmatrix} 22 & 153.1 & 147.6 & 1903.3 & 109 & 929 \\ 153.1 & 119.65 & 1044.29 & 12733.6 & 733.9 & 7234.7 \\ 147.6 & 1044.29 & 1008.2 & 12651.62 & 736.1 & 6527.7 \\ 1903.4 & 12733.6 & 12651\,62 & 362920.36 & 10073.4 & 62945 \\ 109 & 733\,7 & 736\,1 & 10073.4 & 733 & 4582 \\ 929 & 7234.7 & 6527.7 & 62945 & 4582 & 69283 \end{bmatrix}^{-1}$$

Then the regression coefficients were obtained by using MATLAB:

$b_0 = 0.824$; $b_1 = 1.717$; $b_2 = 0.356$; $b_3 = -0.005$; $b_4 = 0.585$; $b_5 = -0.004$

Hence the total cost of damage resulted by an earthquake event can be computed by using the regression equation

$$ln\hat{y}_4 = 0.824 + 1.717x_1 + 0.356x_2 - 0.005x_3 + 0.585x_4 - 0.004x_5$$

2.1.9. SPSS output for multiple regression

SPSS output yields that the coefficient of determination $R_2 = 0.48$ which implied that 48.90% of the variation in the cost of destruction during an earthquake is explained by the regression equation and is significant with p-values of 0.040.

3. Characteristics of Destructive Earthquakes in Pakistan

Table 13 shows the Mean, Median and standard Deviation of the different independent Variables.

3.1 Intensity

Table 14 shows frequency distribution of intensity for twenty two earthquakes occurred in the Pakistan that has been considered in the present study. As overall 36.4% earthquakes are of intensity 7. Earthquakes of intensity 7 are classified as destructive. Moreover 22.7 % of the considered earthquakes are of intensity 5 while 18.2 % of earthquakes are of intensity 6. And 13.6 % of considered earthquakes are of intensity 10 which are classified as most destructive earthquakes. 4.5 % of the considered earthquakes are of intensity 8 and 9 respectively which are considered as very destructive.

3.2 Magnitude

Table 13 shows that the mean magnitudes of the earthquakes are 6.71 with standard deviation of 0.92. According to Richter scale, seismic activities with magnitude 6 to 7 are strong earthquakes and accompanied by local destructions near the epicenters. First class seismological station can observe these earthquakes wherever occur within the earth.

3.3 Depth of Focus

Table 13 shows that the average depth of considered earthquakes is 86.51 kilometer with standard deviation of 97.16. Shallow earthquakes are those earthquakes whose depth of occurrence ranges from 0 to 70 kilometer. The earthquakes with shallower depths bring more destructions than those which have greater depths.

3.4 Location of Epicenter

Table 15 shows the frequency distributions with respect to the location of epicenter of the considered earthquakes. Table 7 shows that the most of earthquake has epicenter located at Zone

Table 13. Means and standard Deviations of the Independent variables

Variable	Mean	Standard Deviation
Intensity(x1)	6.96	1.61
Magnitude(x2)	6.71	0.92
Depth of Focus (x3)	86.51	97.16
Location of epicenter (x4)	4.95	3.03
interval (x5)	42.22	37.83

Table 14. Distribution of the intensity of the earthquake

Intensity	Mean	Percentage
5	5	22.7 %
6	4	18.2 %
7	8	36.4 %
8	1	4.5 %
9	1	4.5 %
10	3	13.6 %
	n=25	100%

Table 15. Epicenter location of earthquake

Zone	Names	Frequency	%
Zone 2	Kohistan-Kashmir	2	9.1%
Zone 3	Quetta-Sibi	6	9.1%
Zone 4	Southern Baluchistan	2	27.3%
Zone 5	Hindukush	2	9.1%
Zone 6	Makran Coast	2	9.1%
Zone 7	Runn of Kuchch	1	9.1%
Zone 8	Sind-Punjab	1	4.5%
Zone 9	Upper Punnjab-NWFP	1	4.5%
Zone10	Western Baluchistan	2	4.5%
Zone 11	Indian Kashmir	1	9.1%
	Northwestern Afghanistan-Tajikistan	n=25	4.5 %
	Kohistan-Kashmir		100 %

3. The 27.3% earthquake considered in study has epicenter at Southern Baluchistan. This implies that southern Baluchistan is very prone to be stricken by the earthquake or any seismic activity.

3.5 Interval of Earthquakes

The average interval of the earthquake considered in the present study is 42.22 seconds with standard deviation of 37.83 as shown in table 5. Earthquakes with longer time interval bring more destructions than those with shorter one.

3.6 Destructions caused by Earthquakes

Table 16 shows the Mean, Median and standard Deviation of the different Dependent Variables. The mean death owning to earthquake is 7188.3183

Table 16. Mean Medians and Standard Deviation of the Dependent variables

Variable	Mean	Median	Standard Deviation
Death(y1)	7188.3183	29	21963.51653
Injured(y2)	114578.0909	66	38183.36629
Affected Families(y3)	34139.8182	1000	105436.2386

as associated to median 29, which is more lifelike estimate of death. The most number of deaths aroused during the Muzaffarabad earthquake in October 8, 2005 which left 87,371 persons dead. It was an earthquake that resulted in enormous destruction of properties and huge loss of lives.

The mean number of people injured is 114578.0909 while the median is 66. The largest number of injured happened on 26 January 2001 Gujarat earthquake that hit India and some area of southern Pakistan which wounded 166,800 persons. The affected families have a mean of 34139.8182 and median of 1000. The largest families were affected in Muzaffarabad earthquake which was about 3147774. The cost property destruction has a mean of 55518265909 and median of 75000000. The largest cost of destruction occurs in earthquake of October 8, 2005. The earthquake left destruction cost of 3.5 billion $. Also the largest cost of destruction occurred in earthquake of Jan 26, 2001 whose epicenter was about 9 km south-southwest of the village of Chobari in Bhachau Taluka of Kutch District of Gujarat, India. The interpolate earthquake reached 7.7 on the moment magnitude scale and had a maximum felt intensity of X (Extreme) on the Mercalli intensity scale. The total destruction cost in this earthquake was about 7.7 billion$.

4. RESULTS & DISCUSSION

4.1 Characteristics of Earthquakes

36.4 percent of the total considered earthquake event that hit Pakistan from 1909 to 2017 were of intensity 7. These earthquakes classified as destructive. The average magnitudes of the earthquakes that are considered in research are:

- ✔ 6.71 With standard deviation of 0.92. According to Richter scale, seismic activities with magnitude 6 to 7 are strong earthquakes and accompanied by local destructions near the epicenters. First class seismological station can observe them wherever they occur within the earth.
- ✔ The average depths of considered earthquakes from 1909 to 2017 are 86.51 kilometer with standard deviation of 97.16. Shallow earthquakes are those quakes with depth ranges from 0 to 70 kilometer. The earthquakes with shallower depths bring more destructions than those which have greater depths.
- ✔ The 27.3 percent earthquake considered in study has epicenter at Zone 3 and that is in southern Baluchistan. This implies that southern Baluchistan is quit prone to be stricken by the earthquake or any seismic activity.
- ✔ The average interval of the earthquake considered in this study is 42.22 seconds with standard deviation of 37.83. Earthquakes with longer time interval bring more destructions than those with shorter time interval.

4.2 Proposed Mathematical Models

$$\ln\hat{y}1 = -4.773 + 0.616x1 + 0.810x2 - 0.012x3 + 0.013x4 + 0.013x5$$

$$\ln\hat{y}2 = -3.754 + 0.586x1 + 0.621x2 - 0.012x3 + 0.172x4 + 0.016x5$$

$$\ln\hat{y}3 = -8.401 + 1.595x1 + 0.510x2 - 0.010x3 + 0.318x4 - 0.018x5$$

$$\ln\hat{y}4 = 0.824 + 1.717x1 + 0.356x2 - 0.005x3 + 0.585x4 - 0.004x5$$

The first model estimate number of deaths and is significant with p-value of 0.009 which has a strong positive linear relationship with correlation coefficient of R=0.766. The second model predicts number of injured persons and is significant with p-values of 0.005 which has a strong positive linear relationship with correlation coefficient of R=0.790. The third model predicts number of affected families and is significant with p-values of 0.012 which has a strong positive linear relationship with correlation coefficient of R=0.757. The fourth model predicts total cost of destructions and is significant with p-values of 0.040 which has a strong positive linear relationship with correlation coefficient of R=0.699.

5. CONCLUSION

Unlike hurricanes and a few different natural hazards, earthquakes can happen at any time of the year and occur without any warning. The current research has successfully proposed such mathematical models which can estimate the earthquake casualties and destruction costs using multiple linear regression analysis. The multiple linear regression analysis method requires data to be linear whereas our available data was nonlinear so we first used transformation to make it linear and then applied the regression analysis. Thus the resulted mathematical models estimate earthquake casualties and destruction costs more efficiently as compared to the existing model and can be concluded as:

- ✔ The most number of deaths provoked during the Muzaffarabad earthquake in October 8, 2005 which left 87,371 persons dead. It was an earthquake that resulted in enormous destruction of properties and huge loss of lives, largest no of families about 3147777 were affected.

- ✔ The largest cost of destruction occurs in earthquake of October 8, 2005. The earthquake left destruction cost of 3.5 billion $. Also the largest cost of destruction occurred in earthquake of Jan 26, 2001 and total destruction cost in this earthquake was about 7.7 billion $.

6. REFERENCES

1. Crampin, S., R. Evans., B.K. Atkinson, Earthquake prediction: a new physical basis, *Ceophys, Journal of Royal Astronomy Society,* 76:147-156(1984).
2. Sarwar ,F., S. Iqbal., M. Qaisar., A. Rehman., F. Akhtar., S.M. Raza, Earthquake Statistics and Earthquake Research Studies in Pakistan, *Open journal of earthquake Research,* 5(2):97-105 (2016).
3. Mejia, J., K. Rojas., N. Valeza, & A.R. Villagracia, Earthquake prediction through Kannan-Mathematical-Model Analysis and Dobrovolsky-based clustering Technique, *Conference: DLSU Research Congress,* 3:1-6 (2015).
4. Osborne, J.W, & E. Waters, The Effects of Minimum Values on Data Transformations, *American Education Research Association, New Orleans*, LA 8: 1-5 (2002).
5. Brown, S.H, Multiple Linear Regression Analysis: A Matrix Approach with MATLAB, *Alabama Journal of Mathematics,* 1-3 (2009).
6. Sarwar, F., S. Iqbal, & L.R. Kamal, An analysis of Pakistan's local network catalog of earthquake for the period of 1905-2007, *Science International,* 23(1): 13-17 (2011).
7. Rizwan, S.A., M.A. Khan., S.A. Debaj, Application of probability theory in the earthquake risk assessment and its consequent possible reduction, *12WCEE* 2000: (1948).
8. Urrutia, J.D., L.A. Bautista, & E.D. Baccy, Mathematical models for estimating earthquake casualties and destruction cost through regression analysis using matrices, *Journal of Physics: Conference Series,* 495:1-15 (2014).
9. Nicole, A, & E. Water, Statistical modeling of earth- quake destruction, Honors Program Thesis. *University of Northern Iowa*: 39: (2016).
10. State Disaster Management Authority Muzaffarabad, http://sdma.pk/.
11. National Disaster Management Authority Islamabad (NDMA), *www.ndma.gov.pk/*
12. Pakistan metrology department and NORSAR. Seismic Hazard Analysis and Zonation for Pakistan. *Azad Jammu and Kashmir:* July (2007).
13. Nipp,T.M, University of Tennessee, Knoxville," Development of a Mathematical Model for the Estimation of Required Maintenance for a Homogenous Facilities Portfolio Using Multiple Linear Regression", May 2017.
14. Ye, M.W, Probability and Statistics for Engineers and scientists, Ninth Edition (2014).