# Predictive model for Arias intensity in northeastern region of India

Himanshu Rana[1] and G.L.S. Babu[1]

[1] Department of Civil Engineering, Indian Institute of Science, Bengaluru, Karnataka, India-560012.

**ABSTRACT**

This study presents a development of attenuation relationship for Arias intensity for Northeastern region of India. The attenuation relationship for Arias intensity is developed in terms of magnitude of seismic event (M), source distance of the earthquake (R) using two step regression analysis. The strong motion dataset is divided into two categories i.e. training dataset and validation dataset. Training dataset is used to calculate regression coefficients and validation dataset is used to examine the efficiency of predictive equations. The developed attenuation relationship for Arias intensity is compared with previously developed attenuation relationships. The comparison shows that the developed equation is efficient for Northeastern region of India.

**Keywords:** Northeastern India, Arias intensity, Two step regression analysis

## 1 INTRODUCTION

Northeastern region of India possesses very high seismicity. Due to these seismic events, this region experiences considerable damage to structures, roads, environment and life every year. Some researchers worked on deterministic and probabilistic seismic hazard based on different ground motion prediction equation for PGA, PGV etc. for this region (Das et al. 2006; Nath et al. 2012; Sitharam and Sil 2014).

Arias intensity is a measure of the intensity of ground shaking due to seismic waves (Tselentis, Danciu and Gkika, 2014). Some researchers developed empirical correlations for Arias intensity using regional earthquake data(Amiri et al., 2009; Chousianitis et al., 2014; Kayen and Mitchell, 1997; R.C.Wilson, 1985). Wilson and Keefer (1985) proposed attenuation relationship in terms of magnitude of seismic event and source distance

$$\log I_a = M - 2 * \log R - 4.1 \quad (1)$$

where $I_a$ is Arias intensity in m/s, $M$ is magnitude of seismic event and $R$ is source distance. The regression analysis is conducted using 30 accelerograms from different seismic events from different regions. Jibson (1987) published an equation using the dataset from Taba earthquake in Iran

$$\log I_a = 0.98 * M - 1.35 * \log R - 4.9 \quad (2)$$

Chousianitis et al. (2014) developed predictive model for Arias intensity ($I_a$) using dataset from Greece using two step regression analysis.

$$\log I_a = -4.822 + 0.919 * M - 1.284 * \log\sqrt{R^2 + 7.323} - 0.006\sqrt{R^2 + 7.312} + 0.397 S_1 + 0.530 S_2 \quad (3)$$

where $S_1$ and $S_2$ are the dummy variable for rock and soil sites. This equation is generated using dataset of 98 seismic events in Greece.

In the present study, predictive relationship for Arias intensity are developed. To develop predictive relationship for Arias intensity, strong motion dataset from 1986 to 2014 is used. The dataset is divided randomly into two categories (i) training dataset (ii) validation dataset. Two step regression analysis is conducted to training dataset to develop adequate predictive equation for Arias intensity for each site class and validation dataset is used to investigate the goodness of fit of this equation. Magnitude is used as energy parameter and source distance is used as energy attenuation parameter. It also considers the effect of site class.

## 2 STRONG MOTION DATASET

Strong motion dataset consists of 113 acceleration time records in two horizontal directions of 25 seismic events from 1986 to 2014. This data is obtained from the website http://strongmotioncenter.org and http://pesmos.in/2011/. This is the most comprehensive and homogenised data centre for Himalayan region. It handles a network of 200 strong motion accelerographs in the Himalayan region. Accelerograms are processed for base line correction and rotated the horizontal component to get the accelerogram in N-S and E-W direction. The dataset has good distribution of magnitude values for different source distances.

## 3 REGRESSION MODEL FOR ARIAS INTENSITY

Arturo Arias (1970) proposed that an intensity scale should give a trace of potential damage created by a seismic event irrespective of the population of life or structures in the region. Arias intensity is defined as:

$$I_a = \Pi/2g\int_0^{T_d}(a(t))^2dt \tag{4}$$

where $g$ is acceleration due to gravity in m/s$^2$ , $T_d$ is the time for which the magnitude of the signal is more than critical acceleration. $a(t)$ is acceleration time history of earthquake waves. Arias intensity is considered as more reliable parameter to describe earthquake shaking to trigger slope instability as compared to peak ground acceleration (PGA) due to consideration of whole acceleration time history.

A predictive model is developed to determine Arias intensity in terms of energy attenuation parameter, energy parameter and site class for Northeastern region of India. This region has high amplification characteristics for seismic energy due to its topography (Sil and Sitharam 2017). This characteristic will have an impact on values of ground motion.

## 4 METHODOLOGY

Regression analysis is a method to develop a relationship between two or more random variables. Regression model for Arias intensity $I_a$ is as follows

$$\log I_a = a*M + b*\log R + c*R + d + \varepsilon \tag{5}$$

where $M$ is moment magnitude of the earthquake. $R$ is the source distance. $a$, $b$, $c$ and $d$ are regression coefficients. $\varepsilon$ is a random error which follows normal distribution having zero mean and variance $\sigma^2$.

When the regression analysis is performed for both energy attenuation parameter (source distance) and energy parameter (magnitude) simultaneously, coefficient for magnitude may affect the coefficient for distance in the regression analysis. This may lead to very high error coefficient in regression results. To avoid this, Joyner and Boore (1993) proposed two step regression process. In this approach, the coefficient for different parameters are calculated in different steps. So, the error in measuring one parameter does not affect the determination of regression coefficient for other parameter.

Two step regression method is used to derive the predictive relationship for Northeastern region of India. The random error is divided into two parts viz. (i) random error for different values of magnitude of seismic event $\varepsilon_e$ (ii) random error for different value of displacement for one value of magnitude $\varepsilon_r$. For the first step of regression analysis, analysis is conducted for different values of Source distance of earthquake for the same value of magnitude, which is shown as

$$\log I_a = c * \log R + d*R + g_i + \varepsilon_e \tag{6}$$

The random error coefficient $\varepsilon_e$ is normally distributed having zero mean and variance $\sigma_e^2$. $g_i$ is the amplitude factor for i$^{th}$ earthquake. This equation can be represented in a matrix form to conduct mathematical analysis regressively.

$$Y_1 = X_1B_1 + \varepsilon_1 \tag{7}$$

where $Y_1$ is vector of observations of dependent variable of $I_a$, $X_1$ is the matrix of independent parameter for each observation, $B_1$ is the matrix of unknown regressive coefficient, and $\varepsilon_1$ is the matrix of normally distributed random error with zero mean and variance $\sigma_r^2$. It is a multiple linear relationship between $Y_1$ and $X_1$. Equation (7) can be solved with least square method.

For the second step of regression analysis, regression analysis is conducted for different values of magnitude, which can be shown as

$$\hat{g}_i = a + b*M + \varepsilon_r + \varepsilon \tag{8}$$

where $\varepsilon_r$ is the random error and $\varepsilon = \hat{g}_i - g_i$ is the error affecting $g_i$. To determine all the regression coefficient, it can be represented in a matrix form

$$Y_2 = X_2B_2 + \varepsilon_2 \tag{9}$$

where $Y_2$ is the vector of $\hat{g}_i$, $X_2$ is the matrix of independent variable and $B_2$ is the matrix of regression coefficients. $\varepsilon_2$ is the vector of error $(\hat{g}_i - g_i) + \varepsilon_r$. $\varepsilon$ and $\varepsilon_r$ are uncorrelated and independent. Variance of $\varepsilon_2$ can be determined by using ordinary least square approach.

$$\text{Var}(\varepsilon_2) = \text{Var}((\hat{g}_i - g_i) + \text{Var}(\varepsilon_r) \tag{10}$$

where $Var(\varepsilon_r)$ is a diagonal matrix having constant value. The value of $B_2$ is determined by using the following equation

$$B_2 = (X_2^TX_2)^{-1}X_2^TY_2 \tag{11}$$

## 5 DATA PROCESSING

For this study, only acceleration time histories are used which has two horizontal components. Arias intensity is given as follows:

$$I_a = \sqrt{I_{EW}^2 + I_{NS}^2} \tag{12}$$

where $I_a$ is the resultant Arias intensity, $I_{EW}$ is Arias intensity in east-west direction and $I_{NS}$ is Arias intensity in north-south direction. The source distance ($R$) is considered as energy attenuation parameter and moment magnitude ($M$) is considered as a parameter

Table 1 Results from two step regression analysis for Arias Intensity

| Site class | a | b | C | d | RMSE |
|---|---|---|---|---|---|
| Soil | 0.2074(±0.05) | -0.3454(±0.13) | -0.00273(±0.00063) | 0.874(±0.25) | 0.317 |
| Rock | 0.1533(±0.0924) | -0.179(±0.0678) | -0.001052(±0.00076) | -0.7515(±0.067) | 0.281 |

of source energy. To consider the site effect in this study, the local site is classified in two different classes according to shear wave velocity in upper 30 m of the site (i) rock sites(700 < $V_{s30}$ < 1400 m/s) and (ii) soil sites (200 < $V_{s30}$ < 700m/s). Separate regression analysis is conducted for different site condition.

The effectiveness of the regression equation depends on goodness of fit. The goodness of fit increases with increase of data set. The strong motion dataset for rock type-site and soil type is divided into two categories randomly (i) training dataset (ii) validation dataset. The training dataset is a larger one and used to calculate the regression parameters for each categories. The validation dataset is used to validate the results of regression analysis.

The training dataset consist of 70 Arias intensity calculated from 18 seismic events having moment magnitude range between 3.7 and 7.5 and the range of source distance varies between 17 km to 389 km. The validation data set consists of 53 Arias intensity determined from 13 seismic events. The magnitude and source distance range of validation dataset are 3.7 to 7.5 and 30 km to 400 km respectively. In these datasets, many events are not present for moment magnitudes from 5 to 6. The validation and training dataset have good data coverage over the range of source energy parameter and energy attenuation parameter as shown in the Fig. 1.

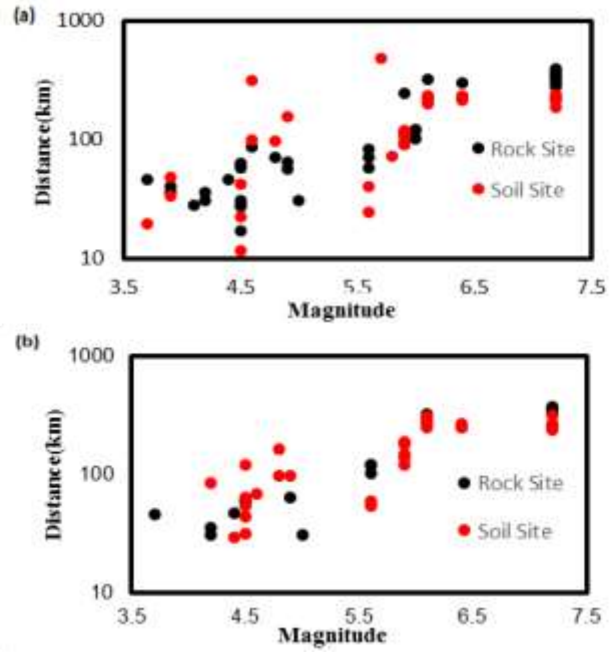


Fig. 1. Distribution of magnitude and source distance for (a) training and (b) validation dataset

## 6 RESULTS

A two step regression analysis is performed for both rock and soil sites. The results of the regression analysis are presented in Table 1.

The Arias intensity for different values of magnitude is shown in the Fig. 2 for rock sites. The rate of decrease of Arias intensity with source distance is higher for lower magnitude. A comparison of Arias intensity for rock sites and soil sites is also shown in the Fig. 3. The values of Arias intensity are higher for soil sites as the attenuation of seismic waves are higher in rock. In soil sites, the seismic waves are reflected and refracted and travel farther from source as compared to rock sites.

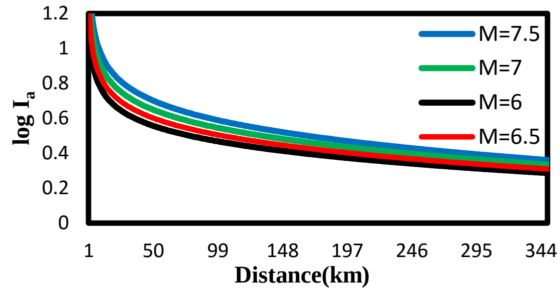


Fig. 2. Comparison of Arias intensity for different values of magnitude for rock site calculated using developed equation

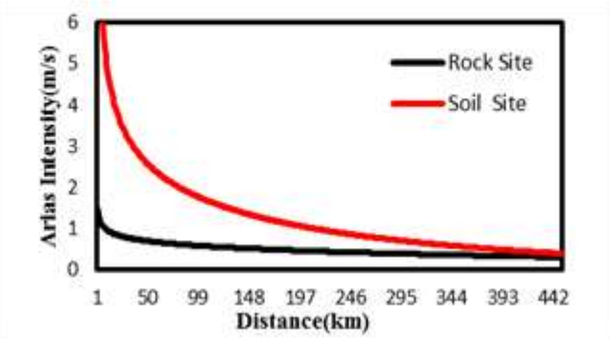


Fig. 3. Comparison of attenuation of Arias intensity for different classes of site

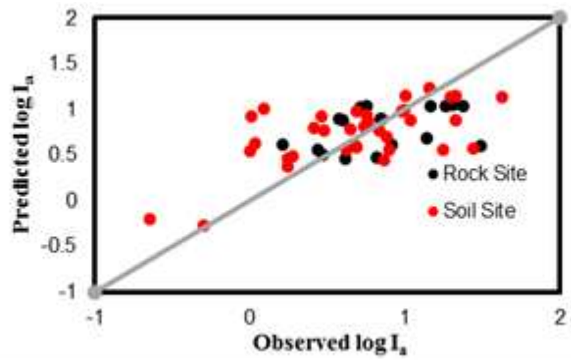


Fig. 1. Comparison of Arias intensity predicted by developed equation and observed Arias intensity.

For validation dataset, a comparison between observed and computed Arias intensity is shown in the Fig. 4. The line of agreement, shown in the Fig. 4, is acceptable for prediction of Arias intensity. The developed equation slightly overestimate the Arias intensity at low values and in good agreement at higher values for soil sites. For rock sites, the developed equation shows good agreement at low values and slightly underestimate for higher values of Arias Intensity. The developed equation in this study is also compared with other attenuation relationships developed by other researchers to show the efficiency of the developed relationship at regional level.

For comparison with other studies, the predictive equation for rock site is used and compared with the values of Arias intensity calculated from developed equation and the equations given by other authors for magnitude 7. The comparison is shown in the Fig 5. The Arias intensity calculated by Eq. (1) and Eq. (2) shows higher values at small distances as compared to the values of Arias intensity calculated by the developed equation. These studies show high rate of decrease of Arias intensity with source distance. The Arias intensity values, calculated using Eq. (3) are lower than as compared to the values of Arias intensity calculated by the present study. In the Greece region, the attenuation of seismic waves is higher as compared to attenuation of seismic waves in Northeastern region of India.

Thus, it is noted that the equation proposed by this study provides very efficient estimations of Arias intensity for Northeastern region of India as compared to previously available predictive equations. The root mean square error (RMSE) of Arias intensity is also less for this equation. The previous studies underestimate the values of Arias intensity, which is attributed to the difference in seismic energy propagation efficiency due to topography and structural geology. So, the use of the developed equation is rationalized to predict the ground motion in this specific region.

## 7 CONCLUSION

Strong motion datasets from this region are used in the present study. Predictive equations for both rock and soil sites are developed to calculate Arias intensity. Arias intensity used for the development of regression model is resultant of two horizontal components of Arias intensity (N-S and E-W). The strong motion datasets used in this study cover the magnitude range from 3.7 to 7.2 recorded from 1986 to 2014. The attenuation relationship for Arias intensity for Northeastern region is not available in any previous studies.

The two step regression analysis is conducted for the determination of attenuation relationship of Arias intensity. To separate the determination of regression coefficient for energy released by the source from

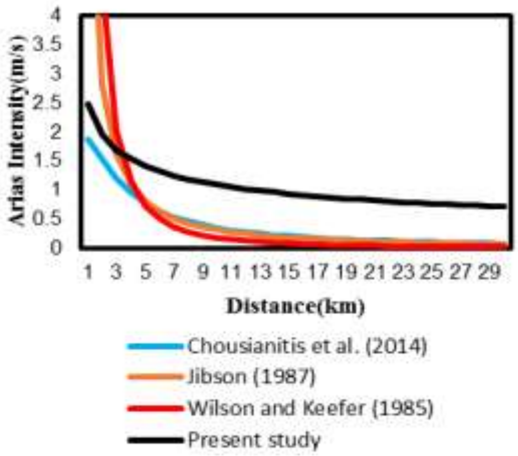


Fig. 2. Comparison of present study with previous study for rock site.

attenuation effect, this method is adopted for regression. The local geological feature leads to the high amplification of seismic energy in this region, which makes the previously developed global model inefficient in this region. The strong motion dataset is subdivided into two categories training dataset and validation dataset. To investigate the efficiency of regression results, validation dataset is used.

## REFERENCES

Amiri, G. G., Khorasani, M., Hessabi, R. M., and Amrei, S. A. R. (2009). “Ground-Motion Prediction Equations of Spectral Ordinates and Arias Intensity for Iran.” Journal of Earthquake Engineering, 14(1), 1–29.

Chousianitis, K., Del Gaudio, V., Kalogeras, I., and Ganas, A. (2014). “Predictive model of Arias intensity and Newmark displacement for regional scale evaluation of earthquake-induced landslide hazard in Greece.” Soil Dynamics and Earthquake Engineering, Elsevier, 65, 11–29.

Das, S., Gupta, I. D., and Gupta, V. K. (2006). “A Probabilistic Seismic Hazard Analysis of Northeast India.” Earthquake Spectra, 22(1), 1–27.

Hsieh, S. Y., and Lee, C. T. (2011). “Empirical estimation of the newmark displacement from the arias intensity and critical acceleration.” Engineering Geology, Elsevier B.V., 122(1–2), 34–42.

Jibson, R. W. (1993). “Predicting Earthquake-Induced Landslide Displacements Using Newmark’s Sliding Block Analysis.” *Transportation Research Record*, (8), 9–17.

Joyner, W. B., and Boore, D. M. (1981). “Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 imperial valley, California, earthquake.” Bulletin of the Seismological Society of America, 71(6), 2011–2038.

Nath, S. K., Thingbaijam, K. K. S., Maiti, S. K., and Nayak, A. (2012). “Ground-motion predictions in Shillong region, northeast India.” Journal of Seismology, 16(3), 475–488.

Wilson, R.C., Keefer, D. K. (1985). “Predicting areal limits of earthquake-induced landsliding In: ZionyJI. editor. Earthquake Hazards in the Los Angeles Region –An Earth-science Perspective" U.S. Geological Survey Professional Paper 1360, 317–45.

Sitharam, T. G., and Sil, A. (2014). “Comprehensive seismic hazard assessment of Tripura and Mizoram states.” Journal of Earth System Science, 123(4), 837–857.