Seismic Hazard Assessment using Machine Learning For Balochistan Region

By

Muhammad Asghar

(Registration No: 00000363978)

Department of Geotechnical Engineering

NUST Institute of Civil Engineering (NICE)

School of Civil & Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)

Seismic Hazard Assessment using Machine Learning For Balochistan Region

By

M.Asghar

(Registration No:00000363978)

A thesis submitted to the National University of Sciences and Technology,

Islamabad, in partial fulfillment of the requirements for the degree of

Master Of Science in

Geotechnical Engineering

Thesis Supervisor: Dr. Abbas Haider

NUST Institute of Civil Engineering

School of Civil & Environmental Engineering

National University of Sciences and Technology

Islamabad, Pakistan

THESIS ACCEPTANCE CERTIFICATE

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS Thesis written by Mr. Muhammad Asghar (Registration No. NUST-2021-MS-Geotech-0000363978), of NUST Institute of Civil Engineering has been vetted by undersigned, found complete in all respects as per NUST Statutes/ Regulations/ MS Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as point out by GEC members and foreign/ local evaluators of the scholar have also been incorporated in the said thesis.

Signature: ____________________

Name of Supervisor: Dr. Abbas Haider

Date: 28/03/2024

Signature (HOD): ____________________
HoD Geotechnical Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Date: ____________________

Signature (Associate Dean): ____________________

Date: 01/4/2024

Signature (Principal & Dean) ____________________
PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

Date: 02 APR 2024

Form TH-4

National University of Sciences and Technology

MASTER'S THESIS WORK

We hereby recommended that the dissertation prepared under our Supervision by: (Student Name Muhammad Asghar & Regn No. 00000363978).
Title: Seismic Hazard Assessment using Machine Learning Alogrithms for Balochistan Region be accepted in partial fulfillment of the requirements for the award of degree with (B+ grade).

Examination Committee Members

1. Dr. Badee Alshameri — Signature

2. Dr. Zain Maqsood — Signature

Supervisor Name: Dr. Abbas Haider — Signature

Head of Department
HoD Geotechnical Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

(Associate Dean)
Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

COUNTERSIGNED

Date: 02 APR 2024

Principal & Dean SCEE
PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

5x Pages

Certificate of Approval

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled "Seismic Hazard Assessment using Machine Learning Alogrithms for Balochistan Region" was conducted by Mr. Muhammad Asghar under the supervision of Dr. Abbas Haider. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the NUST Institute of Civil Engineering in partial fulfillment of the requirements for the degree of Master of Science in Field of Geotechnical Engineering. Department of NUST Institute of Civil, Engineering University of National University of Sciences and Technology, Islamabad.

Student Name: Muhammad Asghar Signature:

Examination Committee:

a) GEC Member 1: Dr. Badee Alshameri Signature:

(Designation & Office Address)

HoD Geotechnical Engineering

Room #011, NICE, SCEE, NUST

b) GEC Member 2: Dr. Zain Maqsood Signature:

(Designation & Office Address)

Assistant Professor Geotechnical Engineering

NICE, SCEE, NUST

Supervisor: Dr. Abbas Haider Signature:

HOD: Dr. Badee Alshameri Signature:

HoD Geotechnical Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Associate Dean: Dr. Syed M. Jamil Signature:

Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

Dean/Principal: Dr. M. Irfan Signature:

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

Author’s Declaration

Author’s Declaration

I Muhammad Asghar hereby state that my MS thesis titled “Seismic Hazard Assessment using Machine Learning Alogrithms for Balochistan Region” is my own work and has not been submitted previously by me for taking any degree from the National University of Sciences and Technology or anywhere else in the country/ world.

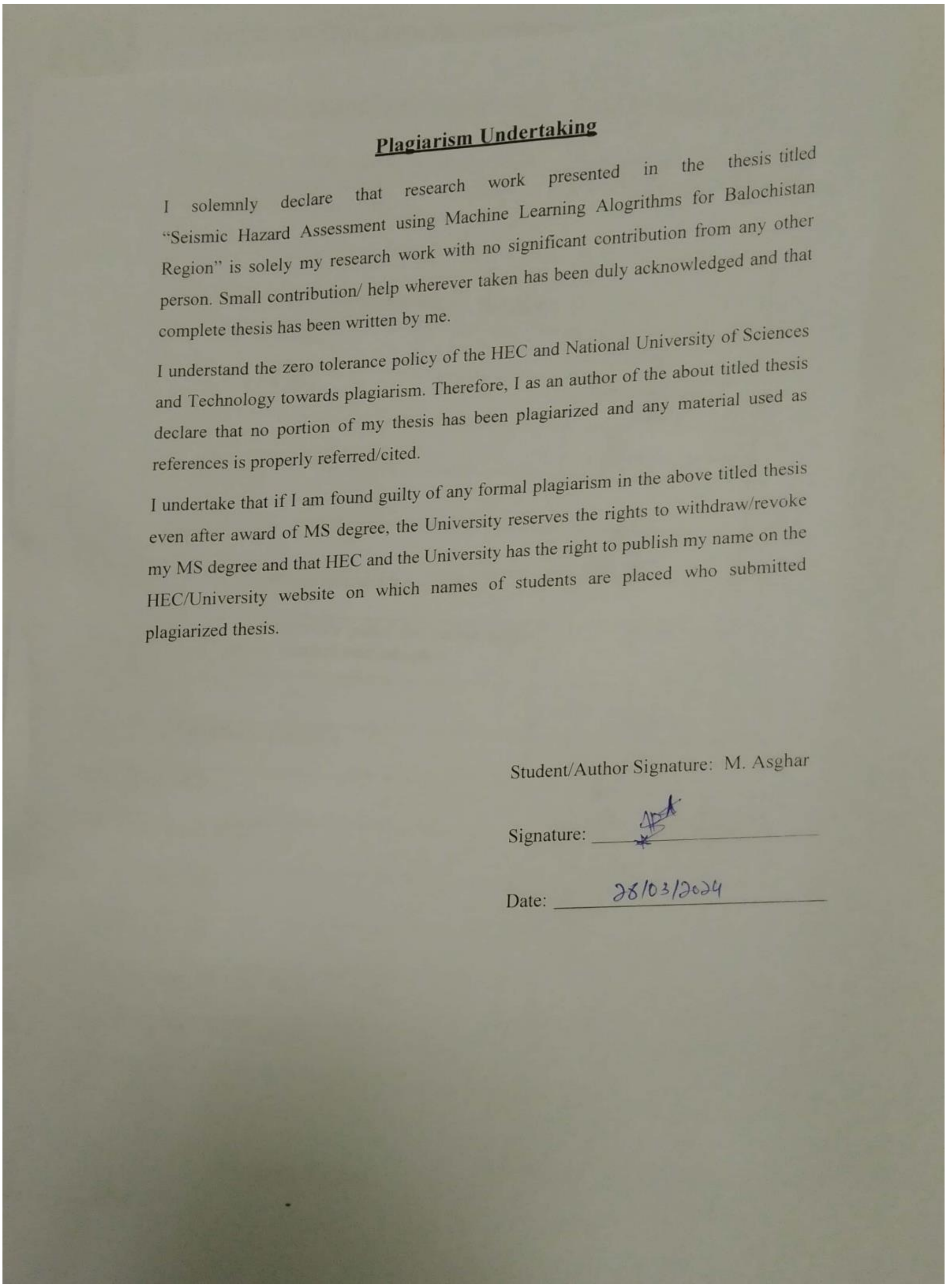
At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Muhammad Asghar

Signature: ______________________

Date: 28/03/2024

Plagiarism Undertaking



Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled "Seismic Hazard Assessment using Machine Learning Alogrithms for Balochistan Region" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and National University of Sciences and Technology towards plagiarism. Therefore, I as an author of the about titled thesis declare that no portion of my thesis has been plagiarized and any material used as references is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and the University has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student/Author Signature: M. Asghar

Signature: ______________

Date: 28/03/2024

Acknowledgements

I want to express my sincere gratitude towards my research supervisor Dr. Abbas Haider who continuously and convincingly conveyed a spirit of hardworking and steadfastness to contrive and complete this project. Without his painstaking efforts, support and guidance, completion of this project would not have been possible.

I would also like to extend my thanks to Dr. Badee Al-Shameri and Dr. Zain Maqsood. Their keen interest in my research work and their priceless effortstowards completion of my thesis work cannot be disremembered.

I am highly indebted to Dr. Badee Alshameri for all the inspiration and guidance Igot from him during my studies. Finally, I am extremely grateful to my parents for their love, support, and hard work, thanking them for their endless patience and encouragement when it was needed the most.

Abstract

The Balochistan Region is in a seismically active belt that has been exposed to many destructive earthquakes in the past in the boundary of Pakistan. In this study, Peak Ground Acceleration (PGA) in the Balochistan Region has been predicted using Gene Expression Programming (GEP) after comparing its efficiency with Artificial Neural Network (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS) within range of 28.070º to 31.000º N latitude to 66.710º and 69.330º E. This is a simple approach to assess seismic hazard assessment using machine learning algorithm. A homogeneous catalogue was obtained by collecting data from Pakistan Meteorological Department (PMD) which comprises 2021 earthquake events in the past divided into 600 data sets. An Expression Tree (ET) of six inputs (longitude, latitude, depth, magnitude, seismic energy, and logarithmic seismic moment) and one output (PGA) was run on GEP Expo tool 5.0. The most suitable model with a correlation function R^2 of 91% was obtained after several iterations of GEP. All results were statistically evaluated and validated from data of United States Geological Survey (USGS) and International Seismic Centre (ISC). The present study explicitly provides the applicability of GEP to predict Peak Ground Acceleration.

Keywords Balochistan Region – Seismic Hazard Assessment - Peak Ground Acceleration (PGA) - Gene Expression Programming (GEP) - Artificial Neural Network (ANN) - Adaptive Neuro Fuzzy Inference System (ANFIS) - Pakistan Meteorological Department (PMD)- United States Geological Survey (USGS) – International Seismic Center (ISC) - GEP Expo tool 5.0.

Contents

List of Figures

List of Tables

Seismic Hazard Assessment using Machine Learning Algorithms : A novel approach for Balochistan Region.

CHAPTER 1: INTRODUCTION

The major environmental problem which has affected worldwide is natural hazard. Statistical data shows that 40% of economic and social disasters are caused by natural hazards [1]. Natural hazard can be described as physical activity which has the most influential outcomes and leads to significant damage and loss of human lives. Natural hazards can be categorized as floods, seismic hazards, landslides, tsunamis, volcanoes, and erosion. These hazards are classified based on their origin and outcomes they yield [2]. Pakistan is a seismically active region and subjected to severe earthquakes in the past few decades. The geological studies reveal that there is an overlap of Indian and Eurasian tectonic plates over the past 30 to 40 million years [3]. The seismo-tectonic environment and unplanned urbanization along with population growth of the country had previously resulted in seismic hazards [4]. Moreover, the fatalities caused due to seismic hazards have a significant rate and need assessment and mitigation measures (Najam, 2021).

To overcome this issue several probabilistic seismic hazard assessments (PSHA) have been done in the past. Pakistan Building Code had been updated from the past few decades resulting in several updated models. These models were programmed using conventional techniques including recording of earthquakes using several instruments and by past geodetic surveys. The updated seismic assessment till now used Peak Ground Acceleration (PGA) for return period of 50, 100, 200, 500 and 1000 years. Latest Building Code Of Pakistan (BCP) provides a framework of seismic zone mapping where Pakistan is divided into 5 zones (zone 1, 2A, 2B, 3 and 4) based on Peak Ground Acceleration (PGA) [5] [6].

This study focuses on using Gene Expression Programming (a machine learning technique) as a new and efficient alternative way Seismic Hazard Assessment. With the advancement in the computational methods and their day by day increasing efficiency has led the use of machine learning preferable [7]. The replacement of seismic hazard assessment using deterministic approach with machine learning for the rapid analysis of daily accumulated and growing volumes of seismic data gives more precise results in short span of time [8]. More accurate and precise results can be achieved for seismic assessment by using several machine learning techniques. [9]. This study focuses on seismic hazard assessment using gene expression programming because this technique of machine learning has not been used in this domain in Pakistan before. Although some studies had been done in the past related to assessment of geophysical data but not specifically used for seismic hazard assessment. [10] [11]

All the research in the past on seismic hazard assessment for Pakistan were executed using spatially smoothed gridded technique along with conventional area sources approach which require time and are far apart from realistic seismic hazards. [12]. Although Probabilistic Seismic Hazard Assessment (PSHA) is a simple technique and provides useful outcomes for engineering design, but it is based on conventional methods which are time consuming and slow. Modification in PSHA or alternative techniques should be used for better risk mitigation. [13] . Gene Expression programming when used for seismic hazard assessment may prove to be more efficient and less time consuming.

Gene Expression Programming is a technique of machine learning used in this research. A mathematical model is a pre-requisite to evaluate the performance of proposed tasks in scientific research. With advancement in technology, the traditional approaches for assessment of different phenomena have been replaced by computational methods. [14] Gene Expression Programming (GEP) is considered as an effective tool in this regard. Optimized parameters of GEP give accurate results. GEP is a more advanced form of Genetic Programming (GP), a subset of machine learning that produces models which depend on genetic evaluation. GEP is a new technique and the evolution of computer programs that strongly considers the character linear chromosomes comprising of genes arranged as head and a tail. The chromosomes play role of genomes that are subjected to rearrangement by means of mutation, transposition, root transposition, gene transposition, gene recombination, one- and two-point recombination. The chromosomes encode expression trees, which are the targets to be studied [15]. The production of these separate entities (genome and expression trees) with distinct functions permits the algorithm to yield high efficiency, which greatly surpasses existing adaptive techniques. Based on neural network and regression techniques, GEP is an effective and more precise optimization technique. A problem-independent solution based on the Darwinian reproduction principle emerged from the computer program.

GEP is an efficient method due to linear fixed width of genetic programming. Furthermore, GEP is the simplest mechanism that further enables the evolution of complicated and non-linear programmes due to multi-genic behavior because of the genetic process on the chromosome level. Five sets make up the entire GEP: Function, Terminal, Fitness Measure, Parameters, and Criteria sets. In GEP, each specimen is set as a genome, which is a fixed-length linear string. Additionally, genetic operators are utilized to modify chromosomes during the reproduction stage.

The process continues by first collecting data randomly, then the best fitness of population is chosen based on error criteria and outliers are identified. Further, the best combination is selected because of mutation, crossover and direct cross-over. This process is also termed as *"learning"*. After running several cycles model having high precision is created by reaching maximum iterations [16]. Genetic programming is actually a modelling technique working on Darwinian evolutionary theory which is used to produce a best fitted model for the concerned study by a predefined structure.

The chronology of the processes undergoing GEP modelling proceed in the following manner :-

1. Based on recorded data (population) number of chromosomes are produced randomly.
2. The chromosomes formed in the first step then generate mathematical equations.

3. Each chromosome is then used to check suitability with targeted function. This is an iterative process.
4. To create modified individuals from other chromosomes genetic operators are applied which are the GEP algorithm.
5. Now more chromosomes are created by several iterations for several generations and model is formed with most efficient producing results.

1.1 Problem Statement

Probabilistic Seismic Hazard Assessment (PSHA) for seismically active regions of Pakistan using smooth gridded approach and other conventional techniques is a time consuming process which is less precise and should be replaced with Machine learning Techniques like GEP which may have more precise results like European and Gulf countries. [10] [12] [7]

1.2 Aim and Objectives

- To propose empirical relation between depth, time, longitude, latitude, Magnitude and Peak Ground Acceleration (PGA) for upcoming seismic wave in Balochistan Region of Pakistan while using GEP for seismic hazard assessment.
- To compare the efficiency of seismic hazard assessment using GEP with other machine learning algorithms.
- To provide aid by finding Peak Ground Acceleration (PGA) of upcoming earthquake to overcome disasters or worst case scenarios.

1.3 Scope

The goal of this study is to propose a correlation for seismicity which would help in controlling post-earthquake disasters. As discussed earlier, the conventional techniques for seismic hazard assessment like PSHA and geodetic surveys are time consuming and require more effort. Seismic activities had led to many losses both financial and human in the past. An effective correlation when developed using GEP may help in overcoming disaster that usually occurs after a high scale seismic activity. GEP has been used previously in other countries for seismic hazard assessment but not in Pakistan. It has wide applications in civil and geotechnical works. This research mainly focuses on the use of GEP for prediction of magnitude using various parameters of earthquake/ seismic wave (longitude, latitude, depth, Peak Ground Acceleration (PGA), seismic moment, seismic energy).

1.4 Literature Review

The Indian continental plate has been interacting with the Erosion plate boundary for millions of years, shrinking the continental lithosphere by more than 2000 kilometers to form massive rock ranges in central Asia. The subcontinental region experiences high seismic activity because of continental collision. As a consequence of such collisions, numerous enormous mountain constructions have been formed (for example, Kiether and Sulaiman ranges, Hindukush Mountains, Karakorum Mountains, and the Pamir ranges) [17]. These devastating earthquake sources in the past century support the necessity for hazard analysis in order to devise policies for reducing seismic risk in the area [18]. The formulation of suitable mitigation solutions at the urban level should incorporate seismic hazard assessment into consideration and encompass all pertinent geological, social, economic, physical, and structural components [19]. Recent years have seen several studies on earthquake risk assessment. The most current and pertinent work was done through a project from 2009 to 2018 called the Worldwide Earthquake Model (GEM) on a global scale. The main goal of the model was to combine and develop multiple national and regional models to offer a homogeneous global earthquake hazard and risk model. A Worldwide Assessment Report (GAR) was compiled by the United Nations Office for

Disaster Risk Reduction (UNISDR) to minimize the risks of many disasters, including earthquakes. This is yet another admirable global effort.

1.4.1 Seismic Risk Analysis

Since seismic risk analysis has gained widespread attention, thousands of lives have been saved thanks to mitigating measures (Cornell, 1968). A few research groups associated with several universities in Pakistan made a serious effort to anticipate earthquakes a couple of years ago. They used geodetic methods and techniques (such as GPS and SAR monitoring) and seismological instruments to successfully identify and limit priority locations where prevention and seismic risk reduction actions should be targeted . The Seismic Crisis that occurred in Muzaffarabad in 2005 with the M 6.1 powerful earthquake and some large scale aftershocks is an extension of this highly effective integration of seismological and geodetic information. Instead of using GPS data to estimate the area's typical two-dimensional ground velocity and strain fields as is more commonly done, the researches were done to reconstruct the velocity and strain patterns along specially selected transects that were properly oriented in conformity with information about the main regional tectonic settings [20] [21]. Balochistan is in a seismically active area [22] . As per PMD, Pakistan has seen 58 destructive earthquakes throughout its history (PMD, 2007). Between 1905 and 2008, the nation saw five extremely destructive large earthquakes that caused significant property and human life destruction. The first of these five significant quakes, with a magnitude of 8.0, struck the northwest Himalayas in 1905. 10,000 structures were demolished, and more than 20,000 people were murdered. The second one, with a magnitude of 7.7, struck Quetta in 1935, destroying the entire city and killing 35,000 people. The third devastating earthquake, which had a magnitude of 8.0, happened close to the shore of Makran. As a result, 4,000 people also died. With something like a magnitude of 7.6, the fourth and most destructive earthquake occurred in 2005. More than 3.5 million people were affected and 86,000 people died when the earthquake struck the country's northeast [23] (PMD, 2007). The analysis finds that there are no disaster prevention projects. The Balochistan region is susceptible to several earthquake-related risks. Urban regions are currently subject to a variety of issues, including climate change and natural disasters, due to a large scale of human relocation, environmental factors, a bad economy, a complex demographic nexus, and infrastructure and diverse functional systems. Policymakers, planners, and managers may make wise decisions and take appropriate action to lessen the effects of natural catastrophes and new threats [24].

Finding a suitable approach that can extensively integrate numerous forms of data is such a complex and difficult undertaking because the assessment of an earthquake hazard of any specific location contains diverse casual variables arising from distinct dimensions of susceptibility [25]. There are numerous ways and approaches for assessing hazards, including Analysis of Nonlinear Dynamics, Failure Mechanism Identification Turkish Method [26]. However, all of these models and techniques are challenging and demand for a high level of knowledge. [27].

1.4.2 Computational Approach (Machine Learning)

The accuracy of the machine learning (ML) based method was superior to the other traditional methods, enabling its application for seismic hazard assessment. [28]. One of the biggest challenges to developing the model is the inability to integrate certain other influencing variables in the future, such as topography, which requires more data and may require large funding. [29]. One of the new epistemic frontiers in the application of artificial intelligence (AI) technology to catastrophe risk management is ensuring ethical, inclusive, and unbiased machine learning techniques. [30]. Due to its capacity to analyze quantities and sources of data that could not otherwise be readily elaborated, the adopted machine learning model demonstrated effectiveness in decreasing CPU time and model-development costs, like all big data technologies. [31]. Indeed, land use managers and planners might employ machine learning as a tool in their daily work. Municipalities with a high priority for intervention are identified by the suggested study, allowing stakeholders to use this tool to prioritize any preventive steps. [32]. The process also makes it possible to determine the most crucial factors to take into account in a multi-risk analysis that combines seismic and hydraulic factors. In other words, using ML-based techniques enables evaluating the variables that are best suited to classify the observations according to overall risk. In fact, the investigation revealed factors related to various risk kinds, which are better at communicating with one another and carrying the most data. The methodology, however, also enables the identification of factors that do not interact with those of a similar kind and, as a result, cannot be applied [33]. Nonetheless, it is important to prevent misuse, therefore important factors like applicability, prejudice, and ethics should be properly considered. A few of the ethical issues regarding a potential abuse of AI technologies include the weakening of cognitive abilities of users, the possibility of spiteful use, the emergence of issues with arrangement of data and systems, the dependence of results on user bias, and the potential preference of the "wrong" problems in relevance to generated outcomes. [34]. Users may be upset over a lack of accountability and transparency, while criticalities frequently draw on an intrinsic mismatch between the algorithm's designers and the communities where the research is done. Moreover, undeveloped machine learning techniques may be applied in circumstances requiring high levels of safety before they are ready. According to Gevaert et al., disaster-risk-management experts are continually looking for knowledge on how to effectively convey the outcomes and uncertainties of machine learning algorithms in order to lower unrealistic expectations. Also, it's important to identify sensitive populations and check them for bias removal. Additionally, the ML model needs to be enhanced by utilizing modern datasets that employ cutting-edge methods to represent various locations [35]

1.4.3 Overview Of Gene Expression Programming (GEP)

Genetic expression programming (GEP) is a type of evolutionary algorithm used to find solutions to problems in various fields, including computer science, engineering, and biology. [36]. In GEP, a population of candidate solutions, represented as computer programs or expressions, are evolved through successive generations by applying genetic operators such as mutation, crossover, and selection. The fitness of each solution is evaluated based on how well it solves the problem being addressed. [37]. GEP is particularly well-suited for problems that require the creation of complex solutions or programs, as it can combine smaller solutions to form larger and more complex ones. It can also be used to optimize solutions that involve non-linear relationships between variables. [37]. One advantage of GEP is that it can generate solutions that are not limited by pre-defined structures or functions, allowing it to explore a wide range of possible solutions. However, this also makes GEP more computationally expensive and requires careful parameter tuning.Overall, GEP is a powerful tool for solving complex problems, and its applications are wide-ranging and varied.

1.4.4 Overview Of Artificial Neural Network (ANN)

An artificial neural network (ANN) is a computational model inspired by the structure and function of biological neural networks found in the human brain. ANNs are composed of interconnected processing nodes, also known as artificial neurons, that are organized into layers. In an ANN, information flows through the network in the form of signals that are processed by each neuron and transmitted to the next layer of neurons. The neurons in each layer receive input signals, perform some processing on those signals, and then transmit their output to the next layer of neurons. This process continues until the output of the final layer is produced. The connections between neurons in an ANN can be either weighted or unweighted. Weighted connections allow the network to learn from input data by adjusting the strength of the connections between neurons. This is done through a process called training, where the network is fed input data and the weights of the connections are adjusted to minimize the error between the output produced by the network and the desired output. ANNs have found widespread applications in various fields, including computer vision, natural language processing, speech recognition, and many others. They are particularly useful in situations where traditional programming approaches are difficult or impossible, and where large amounts of data are available for training the network.

1.4.5 Overview Of Adaptive Neuro Fuzzy Inference System (ANFIS)

An Adaptive Neuro Fuzzy Inference System (ANFIS) is a hybrid computational model that combines the capabilities of artificial neural networks (ANNs) and fuzzy logic systems (FLSs) to enable efficient and accurate inference and decision-making. The ANFIS model consists of five layers of neurons, with each layer performing a specific function. The first layer, known as the input layer, receives input data, while the second layer, known as the fuzzification layer, converts the input data into fuzzy sets using membership functions. The third layer, known as the rule layer, determines the rules that govern the input-output relationships based on the fuzzy sets generated in the second layer. The fourth layer, known as the defuzzification layer, converts the fuzzy output into crisp output, and the final layer, known as the output layer, produces the final output. The ANFIS model can learn from input data through a process called training, which involves adjusting the parameters of the model using a training algorithm such as backpropagation. The model can be used for a wide range of applications, including prediction, classification, and control. ANFIS has several advantages over other traditional modeling approaches, including the ability to handle complex nonlinear relationships, handle uncertain and imprecise data, and adapt to changing conditions. It is widely used in various fields such as finance, medicine, engineering, and environmental studies.

Chapter 2: Research and Methodology

This study has been conducted using machine learning tool i.e. Gene Expression Programming with the help of GEP Expo tool 5.0. The methodology proceeds with collection of data from government departments. The most authentic data related to seismic variations can be gathered from Pakistan Meteorological Department PMD.

The Pakistan Meteorological Department (PMD) is a self-governing organization entrusted with issuing public weather warnings and prediction for protection, safety, and general information.

In addition to meteorology, it also monitors and investigates meteorological phenomena, astronomical events, hydrology, astrophysics research, climatic changes, and studies on a eronautical engineering and renewable energy sources in different regions of the nation, centrally located in Islamabad.

PMD provides authentic data gathered through various meteorological devices and seismic sensors. The data about seismic records of the site under study were fetched from the past years record of seismic history to evaluate longitude and latitude with corresponding occurrence of magnitude.

PMD has offices and research facilities in all provinces and territories of the country so that data can be collected across the whole country.

2.1.Site Geology

The site under study in this research is Balochistan. It is a vulnerable region when it comes to seismic concerns. Balochistan has a wide range of landforms ranging from high lands skirting the mountains to plains and deserts.

The tectonics of Balochistan are extensively characterized by a well-developed and explicitly example of major fault systems in a regime of convergence where one fault type terminates against another. The Chaman transform fault zone medially traversing the entire province intersects with the central Zhob and the Makran convergence zones. These fault zones are of direct relevance to hydrological control to direct reservoir.

In this perspective the following zones play a major part in seismic activities and are listed as follows :-

- Chaman Transform Zone Main Fault (Kharan, Panjgur and Turbat)
- Chaman, Omach Nal, Bhalla Dor,
- Internal Convergence Zone Main Fault (Ziarat)
- Barkhan, Mekhtar-Kohlu, Hamai, Zhob.
- Chaghai Makran Convergence Zone Main Fault (Karat, Mastang)
- Mashki Chah, Dalbundin, Ahmadwal, Usman, Siahan, Ladgasht, Panjgur, Hoshab, Aghol, Ras Malan, Nai RUD, West Makran, Ormara.

2.2.Methodology

The keen study of geographical site and the faults present in the region led to gathering the data of several magnitudes that may facilitate in the hazard assessment of seismic waves in the said region. This data is then organized in ascending order of year wise seismic events occurring so that best suitable output of Peak Ground Acceleration can be predicted via Gene Expression Programming (GEP).

Compiled data as mentioned earlier through government organizations is collected with different inputs (longitude, latitude, seismic energy, seismic moment, earthquake magnitude) and an output (PGA). The existing trend of PGA analyzed through conventional techniques is then run through GEP tool to predict the upcoming peak value in order to take necessary measures.

The data has been divided into two sets i.e. training and testing datasets. Then GEP executes its operation and iteration of algorithms proceeds in order to yield best outcome having suitable statistical errors.

Best GEP combinations are selected for plotting and analyzing the hazard assessment of seismic activity governed due to tectonic activities.

2.3.GEP Execution on The Current Study

The working phenomena of GEP has already been explained in Chapter 2. The execution proceeds by initiating the user interface and then running the program. The program operates on desired mathematical operations which will be discussed in next chapter. The running continues till the best GEP with statistical parameters has been achieved. Now output can be expressed mathematically by the interpreting VBA code and the implementing it in real life problems. Therefore, GEP facilitates as mentioned many times by saving time and yielding required results very effectively.

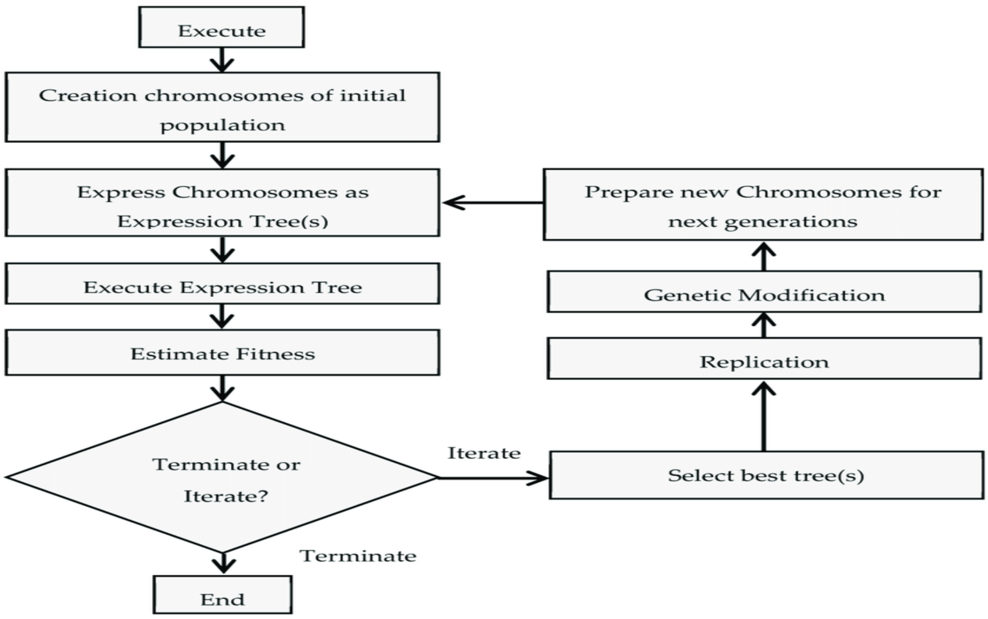


Figure 1 Schematic Diagram Of The Execution Of GEP [37]

2.4.ANN Execution on The Current Study

The data base collected from PMD is again assessed for ANN technique using MATLAB. The input and output parameters are assigned to the MATLAB tool for ANN. After several iterations the best suitable output prediction is selected based on regression value close to 1.

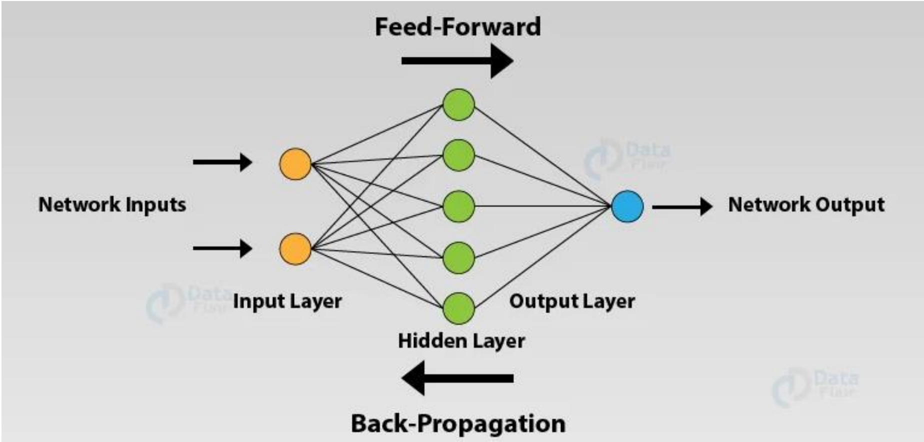


Figure 2-Schematic Diagram Of ANN [30]

2.5.ANFIS Execution on The Current Study

The procedure to analyze data for prediction and data accuracy is same as explained in previous section but the difference is only about back propagation and fuzzy logics used in this technique. The fuzzification makes this ML algorithm most useful then ANN due to more robustness than other ML algorithms.

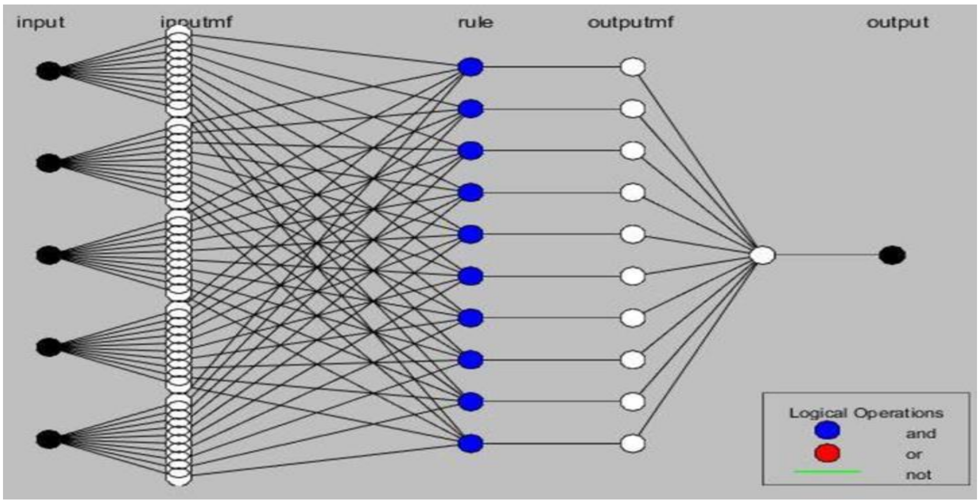


Figure 3-Schematic Diagram Of ANFIS] [37]

2.6.Performance Evaluation Of Machine Learning Algorithms

Based on the deeply explained methodology discussed in previous sections the results for this study have been generated based on R^2 value. As can be seen in the Fig-6. GEP has the closest value to 1 which depicts that the GEP is preferable than ANN and

ANFIS. The results and prediction model in the next chapter are therefore generated and assessed based on the execution of GEP.

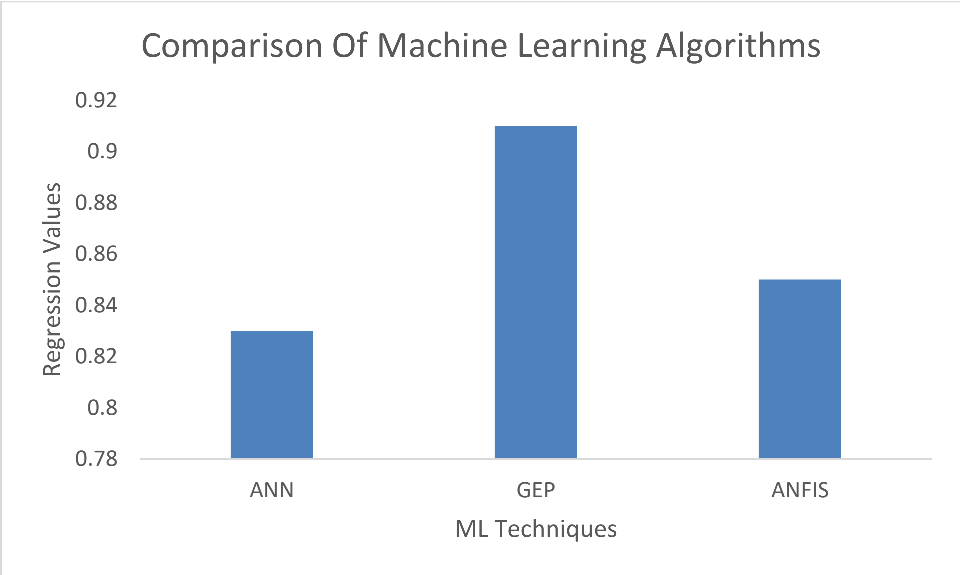


Figure 4-Performance Evaluation Of ML Algorithms

Chapter 3: Results and Discussions

In this chapter the findings of the current study will be discussed in detail along with graphs as validation or proof of the innovations under study.

Training Dataset			Linking Function
R^2	RMSE	R	
0.81	0.64	0.89	Addition
0.78	0.68	0.88	Addition
0.72	0.77	0.85	Addition
0.75	0.73	0.85	Addition
0.77	0.70	0.87	Addition
0.81	0.64	0.90	Addition
0.83	0.59	0.91	Addition
0.85	0.56	0.92	Multiplication
0.87	0.73	0.87	Addition
0.89	0.75	0.87	Addition
0.91	0.72	0.96	Addition
0.83	0.60	0.91	Addition

0.85	0.72	0.94	Subtraction
0.83	0.71	0.94	Addition
0.87	0.71	0.94	Addition
0.88	0.68	0.88	Addition
0.86	0.59	0.87	Addition
0.85	0.66	0.92	Addition
0.84	0.88	0.94	Addition
0.84	0.85	0.85	Addition

Table 1 GEP Parameters

The algorithm for the model of *PGA* is re-presented as expression tree as shown in Fig. The ET was decoded to derive the empirical relationships. The ET for *PGA* comprises of four basic algebraic operators i.e., +, −, x and ÷.

No. of inputs Used variables	No. of Chromosomes	Head size, Number of genes	Program size, No of literals	Duration (min)
6,4	30	8,3	36,9	120
	30	8,3	39,9	100
6,5			42,10	80
	50	8,3	36,7	70

			36,9	60
6,6	50	10,5	79,23	60
			77,22	50
			79,27	40
	100	10,3	33,11	120
			49,18	100
			49,15	80
	150	8,5	64,16	70
			76,18	60
			78,14	50
			78,19	40
5,4	200	3,5	32,6	120
			38,11	100
		5,5	42,10	80
			40,8	70
			35,6	60

[a]The operations employed included +, -, *, /, sqrt, x^3

[b]The operations employed included +, -, *, /, sqrt, x^3, x^2

[c]The operations employed included +, -, *, /, sqrt, x^3, x^2, 3Rt

[d]The operations employed included +, -, *, /, sqrt, x^3, exp, sin, cos, atan, ln

[e]The operations employed included +, -, *, /, sqrt, x^3, x^2, 3Rt

[f]The operations employed included +, -, *, /, sqrt, x^3, x^2, pow

[g]The operations employed included +, -, *, /, sqrt, x^3, exp, sin, cos

[h]The operations employed included +, -, *, /, sqrt, x^3, x^2,3Rt,4Rt, exp, ln

[i]The operations employed included +, -, *, /, sqrt,

[j] The operations employed included +, -, *, /, sqrt, x^3, exp, sin, cos, atan,ln

[k]The operations employed included +, -, *, /, sqrt, x^3, x^2

[l]The operations employed included +, -, *, /, sqrt, x^3, x^2,3Rt,4Rt, exp, ln

[m]The weight of the “+, -, *” operations was four times that of others.

[n]The weight of the “+, -, *” operations was seven times that of others.

[o]The weight of the “*” operations was four times that of others.

[p]The weight of the “+, -, *” operations was three times that of others.

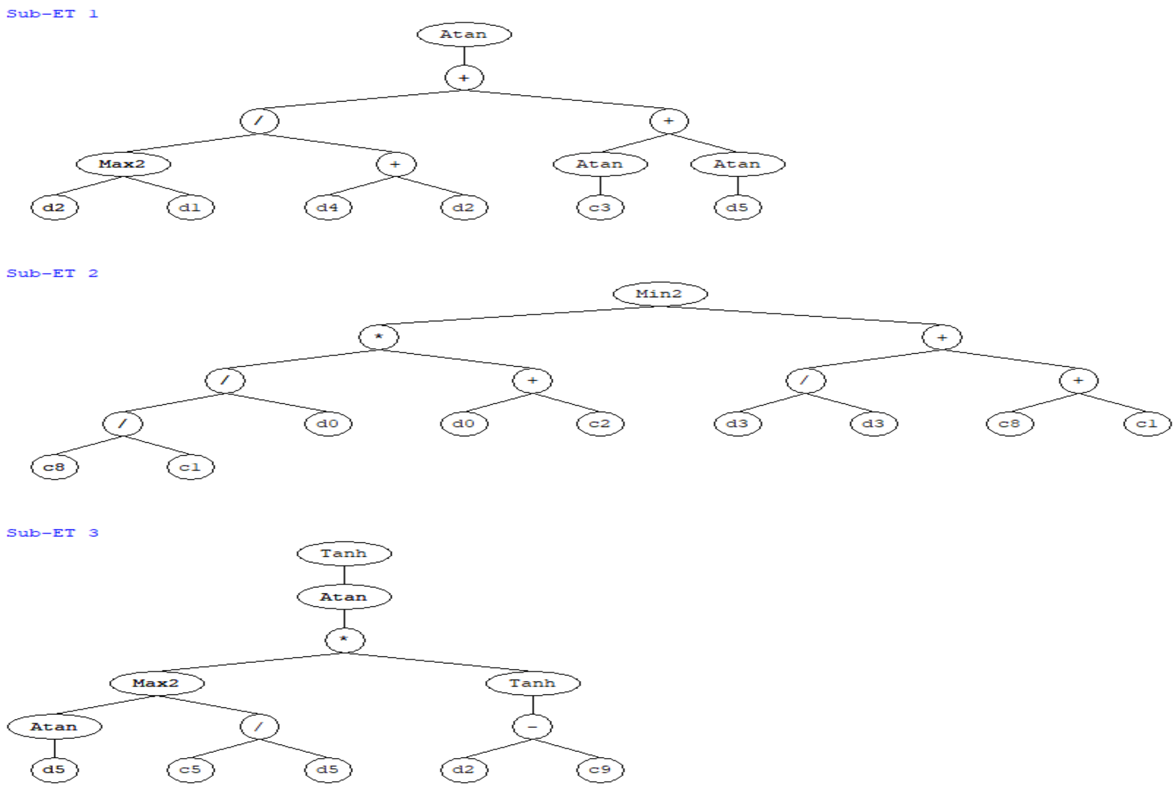


Figure 5 Expression Tree

A visual comparison of the anticipated model and the researched data for *PGA* is shown in Fig-7. The findings presented in this figure reveal a significant correlation which has been shown in Pearson Matrix in Table-1. The quantity of datasets has a significant impact on the suggested model's reliability which is also shown in descriptive statistics below in Table-2. To get better results, diverse history of seismic data from 1901 to 2022 was kept under study.

	Latitude	Longitude	Depth_km	Energy_log	Seismic Moment_l

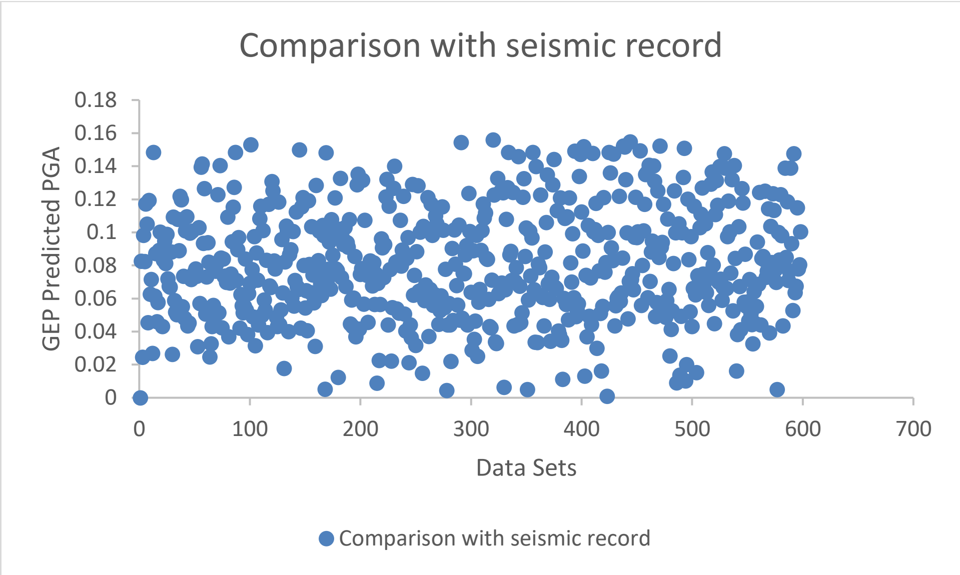


Figure 6 Comparison with seismic record

					og
Latitude	1				
Longitude	0.021041	1			
Depth_km	0.168129	-0.16857	1		
Energy_log	-0.00687	-0.17939	0.143611	1	
Seismic moment_log	-0.00687	-0.17939	0.143611	0.156423	1
Mw	-0.00687	-0.17939	0.143611	0.136222	0.145322

It should be noted that numerous trials and algorithms were conducted to determine the database's veracity. This produced 2022 datasets, which were used to create the corresponding empirical model. The training, validation, and testing sets of the database were randomly chosen for this investigation. The model was trained using the training data, and the validation data was used to confirm the model's generalizability. Throughout the testing process, many expressions were tested on the collected data. The statistical errors can be calculated by the Eqs (1)-(5).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(e_i - m_i)(e_i - m_i)}{n}} \tag{1}$$

$$\text{MAE} = \frac{\sum_{i=1}^{n}|e_i - m_i|}{n} \tag{2}$$

$$\text{RSE} = \frac{\sum_{i=1}^{n}(m_i - e_i)(m_i - e_i)}{n} \tag{3}$$

$$\text{RRMSE} = \frac{1}{|\bar{e}|}\sqrt{\frac{\sum_{i=1}^{n}((e_i - m_i)^2}{n}} \tag{4}$$

$$\text{R} = \frac{\sum_{i=1}^{n}(e_i - \bar{e}_i)(m_i - \overline{m_i})}{\sqrt{\sum_{i=1}^{n}(e_i - \bar{e}_i)^2 \sum_{i=1}^{n}(m - \overline{m_i})^2}} \tag{5}$$

In Table 2-Descriptive Statistics of Input Parameters, the descriptive statistics are displayed. It is advised to employ the provided formulas for this set of data to make accurate forecasts of the PGA model.

It should be noted that numerous tests were conducted to evaluate the database's consistency and validity. The datasets that considerably (up to 20%) diverged from the overall trend were regarded as negligible while developing or assessing the performance of the models.

Latitude		*Longitude*		*Depth_km_*		*Energy_logE_*		*Seismic_moment*		*Mw*	
Mean		Mean		Mean		Mean		Mean		Mean	
Standard Error		Standard Error		Standard Error		Standard Error		Standard Error		Standard Error	
Median		Median		Median		Median		Median		Median	
Mode		Mode		Mode		Mode		Mode		Mode	

Standard Deviation		Standard Deviation		Standard Deviation		Standard Deviation		Standard Deviation		Standard Deviation	
Sample Variance		Sample Variance		Sample Variance		Sample Variance		Sample Variance		Sample Variance	
Kurtosis		Kurtosis		Kurtosis		Kurtosis		Kurtosis		Kurtosis	
Skewness		Skewness		Skewness		Skewness		Skewness		Skewness	
Ran		Ran		Ran		Ran		Ran		Ran	

ge		ge		ge		ge		ge		ge	
Minimum		Minimum		Minimum		Minimum		Minimum		Minimum	
Maximum		Maximum		Maximum		Maximum		Maximum		Maximum	
Sum		Sum		Sum		Sum		Sum		Sum	
Count		Count		Count		Count		Count		Count	

Table 2-Descriptive Statistics of Input Parameters

To comprehend the maximum error % in the relations, the entire database, gathered from past and predicted, has been shown with absolute error in the corresponding data point as shown in Fig-10. The predicted and real outputs of the model can be found to be reasonably close to one another, with the least mean error. Furthermore, the frequency of maximal mistakes is quite insignificant. 80% of anticipated PGA values have been found to have inaccuracy under 0.2g. The whole database, actual and predicted PGA outputs, and absolute error in the corresponding data point have all been displayed, as illustrated in Fig-10 , to help perceive the highest error percentage in the relations. It can be shown that the model's outputs and predictions are reasonably accurate, with an average error in *PGA* of 0.2g and a maximum error of 0.42g. Additionally, the maximum error occurs with a very low data set. 20% of predicted *PGA* findings have been found to have inaccuracy under 0.2g. For the external validation of the suggested GEP models, different checks which validate statistically are also provided in Table-2. to have inaccuracy under 0.2g.

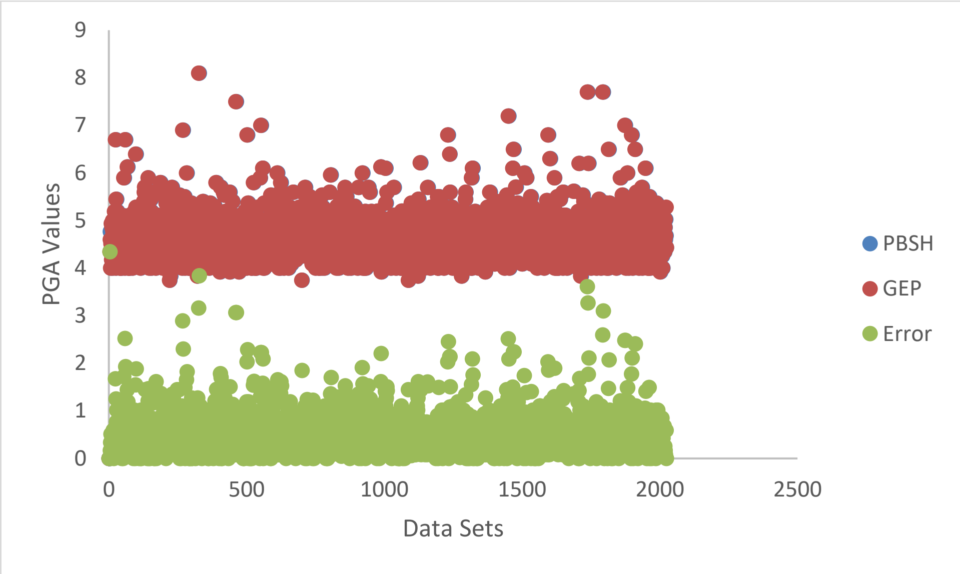


Figure 7-Absolute Errors

3.1.Sensitivity And Parametric Analysis

The relative contribution of different variables to the prediction of PGA is investigated by conducting sensitivity analysis (SA) using Eqs. (6)-(7).

$$N_i = f_{max}(x_i) - f_{min}(x_i) \quad (6)$$

$$\mathrm{SA} = \frac{N_i}{\sum_n^{j=1} N_j} \quad (7)$$

Where the remaining input parameters are held fixed at their mean values, and $f_{max}(x_i)$ and $f_{min}(x_i)$ represent the highest and lowest values of the anticipated output based on ith input domain, respectively. The outcomes of the sensitivity analysis are shown in Fig-11 for predicted PGA, respectively. From a geotechnical engineering perspective, it is clear from the Fig-11 that the input elements' contributions to the predicted PGA are depending on many factors.

As the topography is terraneous to plain areas so the PGA increases from low to high values from mountainous regions and then remains constant. Due to irregularities and heterogenous soil strata considering several geophysical features including ground water one can consider the irregularities in longitude. [38]

Considering the depth factor, which is more crucial, as the seismic wave moves away from focal point or origin and the depth increases moving towards the surface PGA decreases due to different soil strata and other morphological features present underground. [39]

Seismic Energy and Seismic Moment show similar trends as the intensity of seismic wave decreases and PGA is reduced moving towards the surface so is the case with these factors. Irregularity in magnitude is again subjected to changing geographical position along with unique topography of site under study. [40]

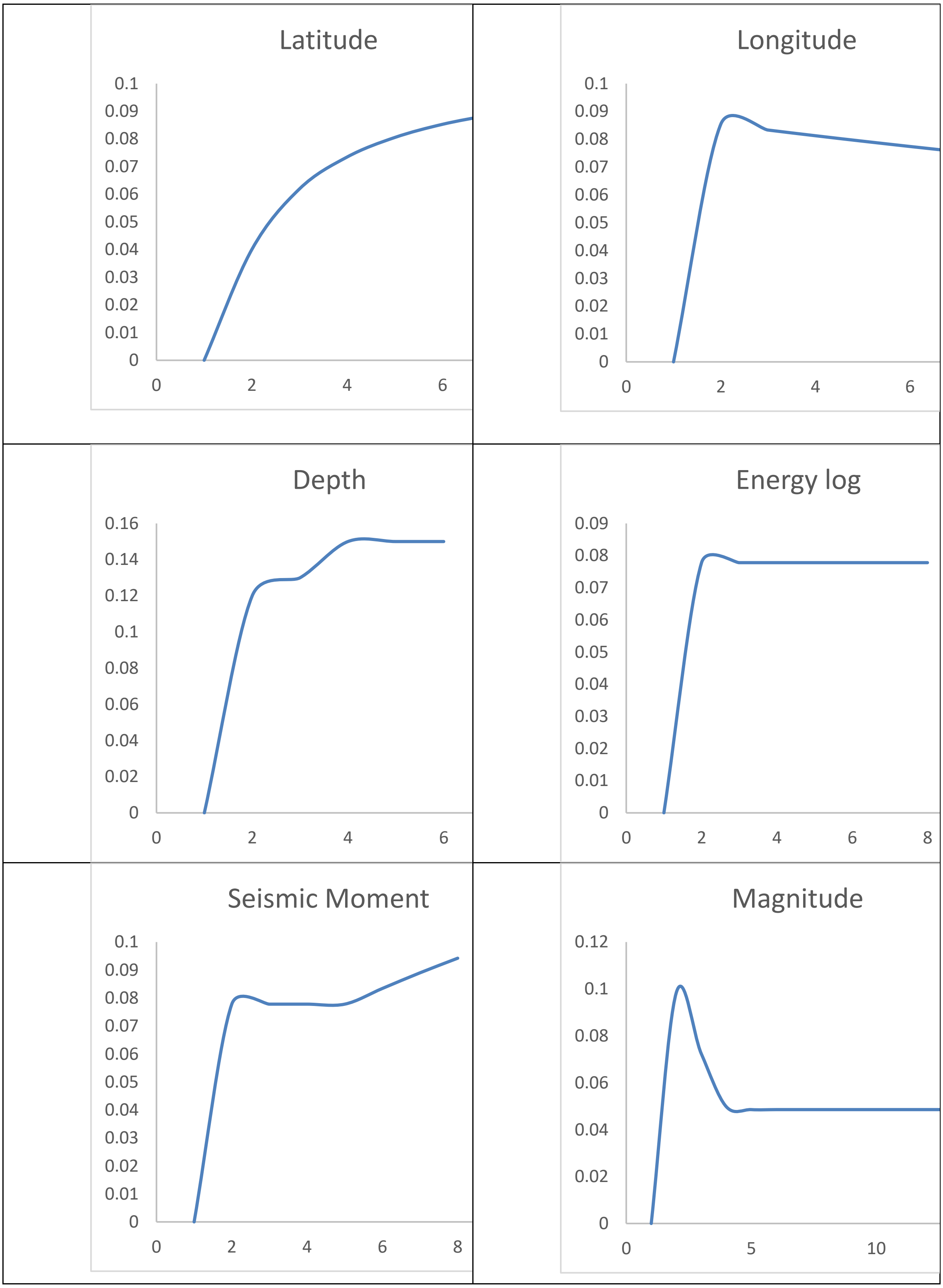


Figure 8 Trend Of Input Parameters

The usefulness of the most crucial input elements in analyzing of the trends of PGA is also assessed in this research using the analysis of parameters. This is done by keeping all input variables at mean values, and the variation in characteristics is seen as single input variable is increased from its minimum to maximum value. Using the GEP model as a basis, the results of a parametric analysis are displayed in Fig 11.

The depth is a crucial factor influencing the PGA values among other parameters, according to well-done analysis. The *PGA* values would increase with an uprise in the depth value up to a certain limit and then it becomes constant, and vice versa. Fig-11 shows that the characteristics of PGA continue to show a significant effect with varying values of Mw and Depth. In Fig-12. data sets have been combined to avoid difficulty in understanding and then their contribution to PGA values is represented to see which parameter has the most dominant effect on the overall behavior of targeted output i.e., PGA.

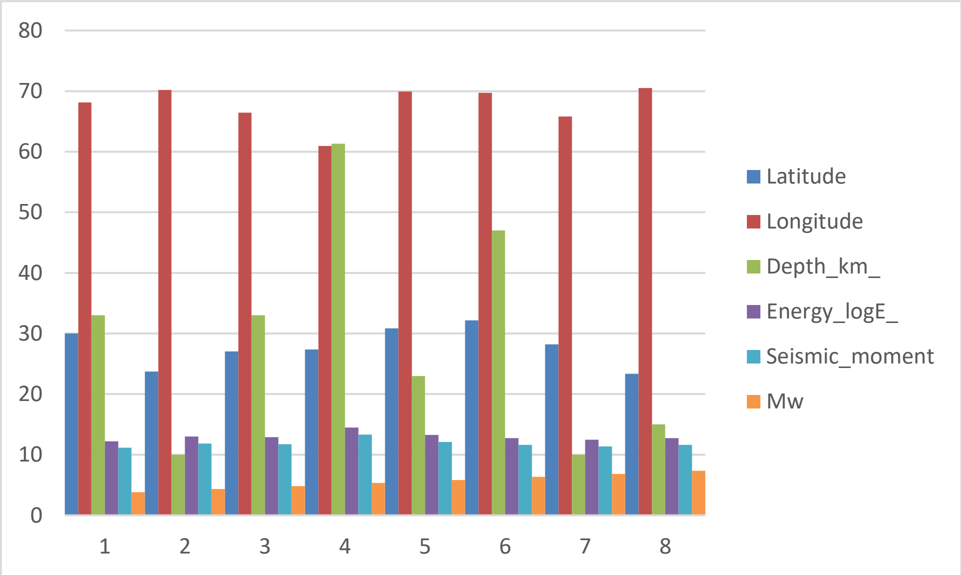


Figure 9-Variable Importance

3.2. Validation of the trend under study with United States Geological Survey (USGS) and International Seismic Center (ISC)

To validate the current study with International data obtained from official sites of USGS and ISC the following trends were observed which can be seen and compared with Fig 11. magnitude shows variation in correspondence with the PGA values. This may be due to carelessness in installment of seismic instruments or sensors.

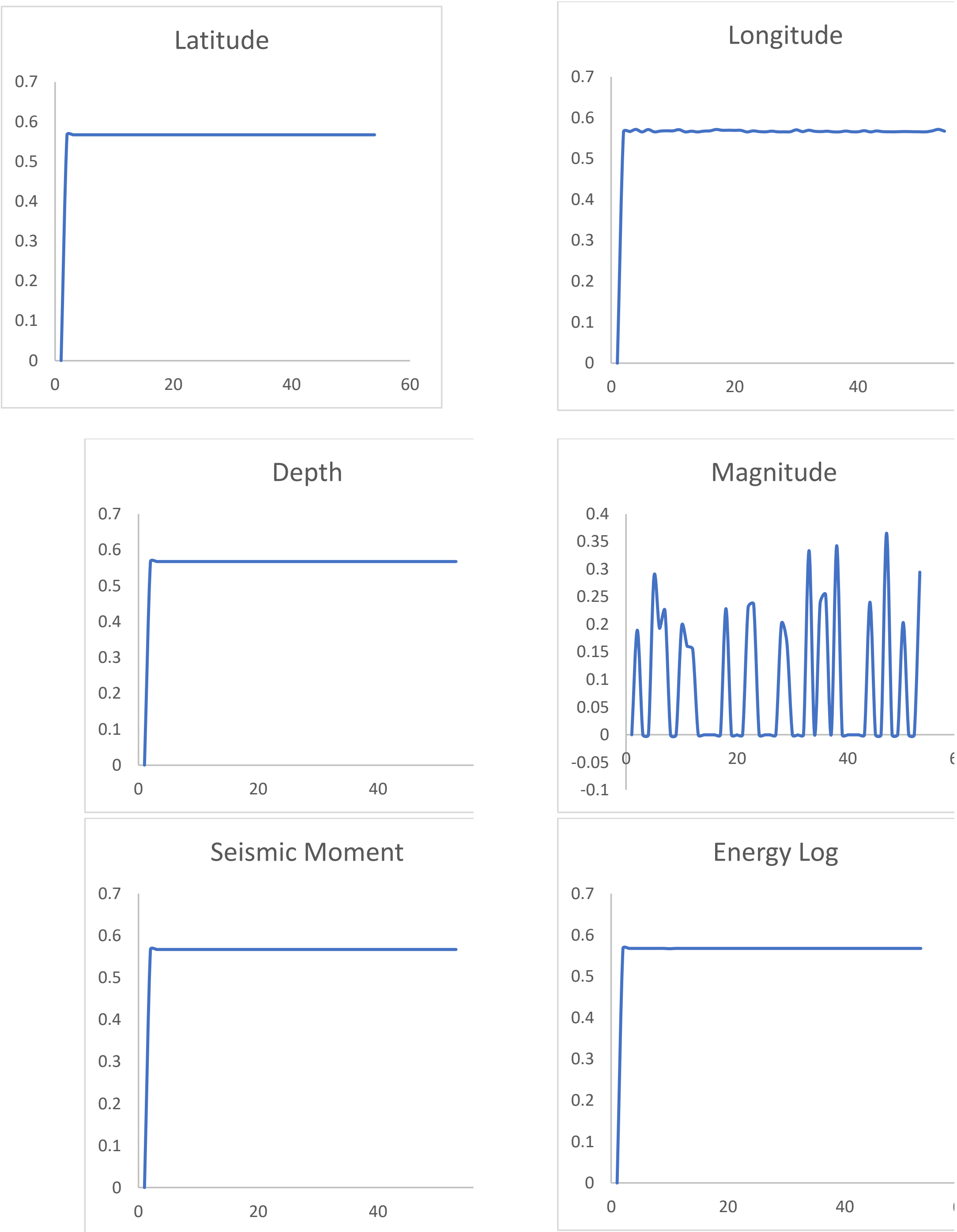


Figure 10 Trend Of Input parameters from USGS Data

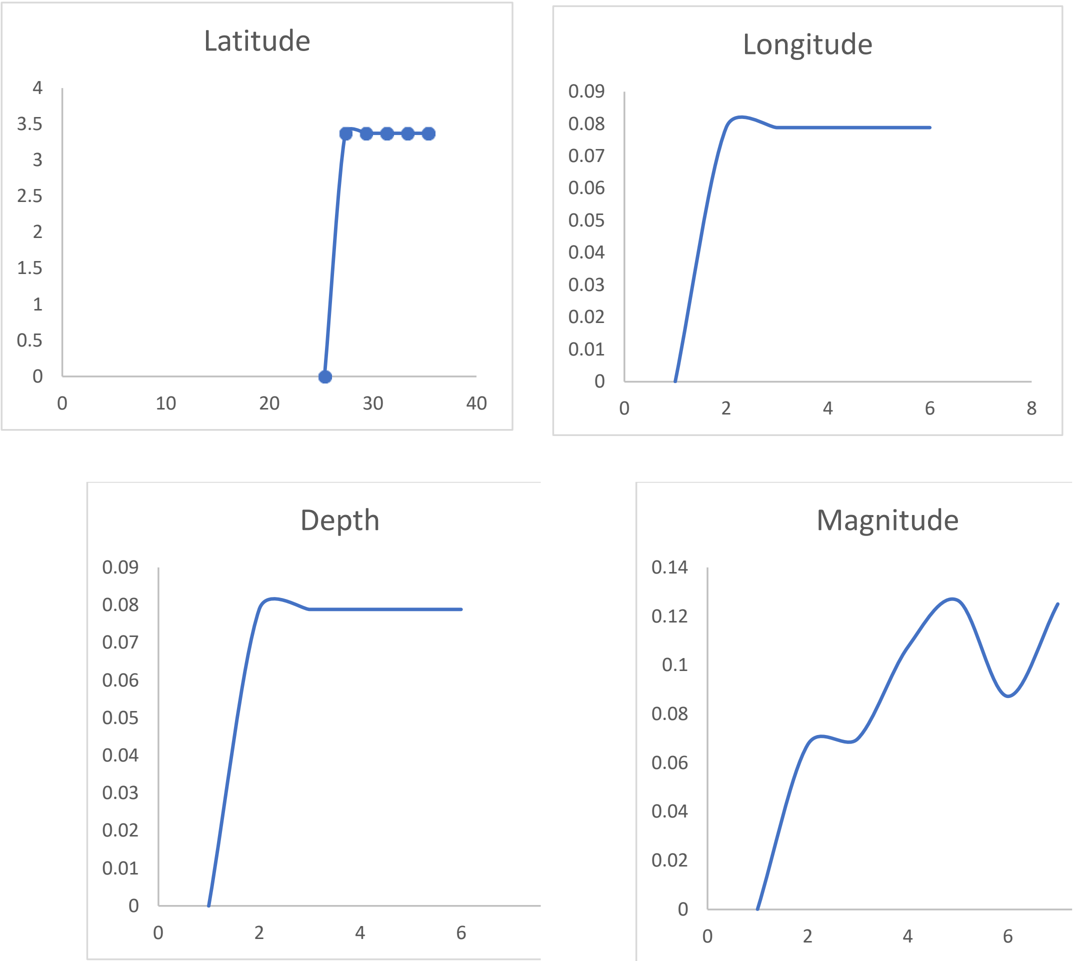


Figure 11 Trend Of Input Parameter from ISC Data

3.3.Ground Motion Prediction Equation

The prediction equations (8)-(10) considering PGA as an output and six input parameters mentioned below give the relationship and validity of interdependency of variables defined in Pearson Matrix earlier in this chapter.

$$a = tan(d)^2 \frac{(long^2)}{(SM)^2 + (d)^2} + tan6.83 + \tan\,(Mw)^2) \tag{8}$$

$$b = a + \left(\frac{-7.09}{4.05}\right) \div (lat^2 + 4.33) - 3.04 + E \tag{9}$$

$$y = b + (\tan(Mw^2))(\frac{6.90}{Mw^2})(\tan(d^2) + 1.05 \qquad (10)$$

Where, a and b are defined variables.

Long = longitude

SM = Seismic Moment

d = Depth

Mw = Magnitude

Lat = Latitude

E = Seismic Energy

y = PGA (Peak Ground Acceleration)

Chapter 4: Conclusions

This study predicts the applicability of Gene Expression Programming for Seismic Hazard Assessment while considering Peak Ground Acceleration (PGA) as a targeted output for various input parameters which have a significant impact on seismic wave as can be viewed in the sensitivity analysis section of this study.

In the Fig-15 comparison of the outputs of PGA by the application of several ML techniques can be seen which lie close to each other (i.e., 0.2g-0.25g) and validate the accuracy of results. The diverse data of 2021 points of Balochistan Region was divided into 600 data sets in order to provide clusters of data sets close to each other so that statistical outliers could be discriminated, and the best model could be achieved.

Fig-15 clearly explicit that the outcomes generated for different Machine Learning techniques show similar trend as that obtained by Probabilistic Seismic Hazard Assessment with less than 5% error in the values compared in the estimated errors.

Thus from the above discussion it can be concluded that the GEP technique is by far the most accurate among other techniques (i.e. ANNs, ANFIS) for predicting the PGA value based on recorded seismic history. The equation of 10^{th} order degree can be seen in Fig. 15 which can also be used as a prediction model but due to its complexity GEP value is more preferred.

The prediction model may help in designing buildings and structures like bridges, roads or dams which can resist the worst scenarios of upcoming high PGA values leading to huge loss of humanity and economy.

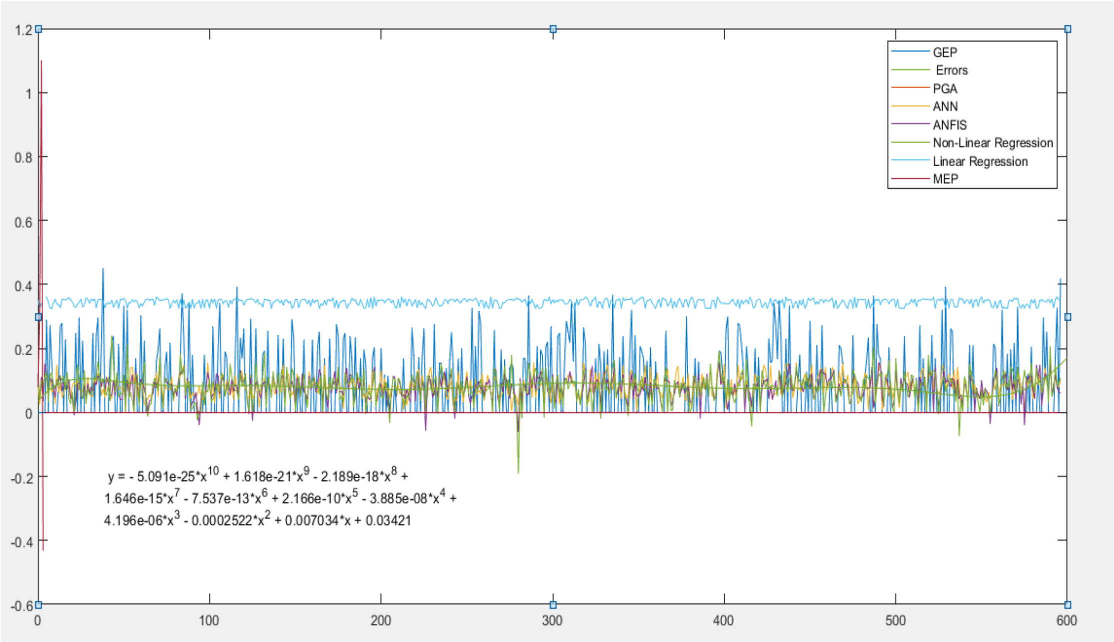


Figure 12-Comparison Of ML outcomes

References

[1] S. B. A. B. S. &. D. J. (. (. S. C. Choudhury, "Flash Flood Susceptibility Mapping Using GIS-Based AHP Method.," *Spatial Modelling of Flood Risk and Flood Hazards,* pp. pp. 119-142, 2022.

[2] M. A. H. P. H. R. &. A. Z. Bordbar, " Multi-hazard spatial modeling via ensembles of machine learning and meta-heuristic techniques," *Scientific Reports,* pp. 1-17, 2022.

[3] R. C. F. A. &. J. K. H. Quittmeyer, "The seismicity of Pakistan and its relation to surface faults.," *Geodynamics of Pakistan,* pp. 271-284, 1979.

[4] M. &. K. M. Y. Shafique, " Earthquake hazards and risk mitigation in Pakistan.," *Disaster Risk Reduction Approaches in Pakistan.Springer, Tokyo.,* pp. 101-117, 2015.

[5] A. Q. H. S. Z. U. R. Z. K. Z. &. A. Q. Bhatti, "Probabilistic seismic hazard analysis of Islamabad, Pakistan," *Journal of Asian Earth Sciences, 42(3),* pp. 468-478, 2011.

[6] F. A. Z. S. R. A. &. R. I. A. Najam, "An updated probabilistic seismic hazard assessment (PSHA) for Pakistan," *Bulletin of Earthquake Engineering, 19(4),* pp. 1625-1662, 2021.

[7] F. Soleimani, "Probabilistic seismic analysis of bridges through machine learning approaches," *Structures,* pp. Vol. 38, pp. 157-16, 2022, April.

[8] R. L. J. H. W.-l. L. J. L. W. L. P. G. G. C. M. T. G. Liu, "Scaling relation of earthquake seismic data," *Statistical Mechanics and its Applications,* pp. pp. 2092-2102, 2018.

[9] N. &. K.-B. O. Bektaş, "Conventional RVS Methods for Seismic Risk Assessment for Estimating the Current Situation of Existing Buildings: A State-of-the-Art Review," *Sustainability,* pp. 14(5), 2583, 2022.

[10] D. Z. V. B. O. &. K. A. Melkov, " Temporal and Spatial Geophysical Data Analysis in the Issues of Natural Hazards and Risk Assessment (in Example of North Ossetia, Russia," *Applied Sciences,* pp. 12(6), 2790., 2022.

[11] F. &. Y. T. Kaneko, "A New Function for Frequency Analysis in the Risk Assessment of Offshore Oil-Related Platforms," *Journal of Ship*

Research, pp. 1-11, 2022.

[12] S. A. P. K. H. I. G. M. M. R. A. R. &. E. W. Khan, " A New Framework for Earthquake Risk Assessment in Developing Countries," in *15th World Conference on Earthquake Engineering (WCEE)*, 2012.

[13] G. K. V. G. P. A. &. N. A. Panza, " Why are the standard probabilistic methods of estimating seismic hazard and risks too often wrong," *Earthquake hazard, risk and disasters,* pp. pp. 309-357, 2014.

[14] M. R. K. ,. B. K. ,. Aydin Shishegaran, "Computational predictions for estimating the maximum deflection of," *International Journal of Impact Engineering,* 2020.

[15] H. Saghi, M. Behdani, R. Saghi, A. R. Ghaffari and S. Hirdaris, "Application of Gene Expression Programming Model to Present a New Model for Bond Strength of Fiber Reinforced Polymer and Concrete," *Journal of Civil Engineering and Materials Application,* 2020.

[16] N. G.-M. M. T. Alireza Arabshahia, "Development of applicable design models for concrete columns confined," *Structures,* pp. 225-244, 2020.

[17] A. G. C. R. S. Sankar Chatterjee, "The longest voyage: Tectonic, magmatic, and paleoclimatic evolution of the Indian plate during its northward flight from Gondwana to Asia," *Gondwana Research,* 2013.

[18] W. M. T. A. M. P. J. A. L. D. E. F. H. F. N. L. K.-F. M. C. M. M. D. P. M. C. Gerstenberger, "Probabilistic Seismic Hazard Analysis at Regional and National Scale: State of the Art," *Reviews of Geophysics,* 2020.

[19] "Earthquake Risk Assessment of Sabah, Malaysia Based on Geospatial Approach," *Noor Suhaiza Sauti, Mohd Effendi Daud, Masiri Kaamin , Suhaila Saha*, 2022.

[20] B. Growley, "Landslide Susceptibility Zonation GIS for the 2005 Kashmir," University Of Montana, 2008.

[21] G. F. &. B. J. Panza, "NDSHA: A new paradigm for reliable seismic hazard assessment," *Engineering Geology,* pp. 275, 105403, 2020.

[22] B. &. N. F. Aslam, " A statistical analysis of the spatial existence of earthquakes in Balochistan: clusters of seismicity," *Environmental*

Earth Sciences, pp. 79(1), 1-11, 2020.

[23] J. A. Nizamani, "Assessment of different hazards and vulnerabilities with sparse data in coastal city of Karachi, Pakistan," *Master's thesis, Middle East Technical University,* 2020.

[24] J. P. A. F. W. S. A. J. P. M. L. S. L. R. J. Z. D. &. H. D. Avouac, "The 2013, Mw 7.7 Balochistan earthquake, energetic strike-slip reactivation of a thrust fault," *Earth and Planetary Science Letters,,* pp. 391, 128-134., 2014.

[25] H. D. V. R. D. S. V. T.-D. D. R. E. A. K. S. F. ,. Z. Z. L. S. S. O. a. P. G. Crowley, "Exposure model for European seismic risk assessment. Earthquake," *Earthquake Spectra,* pp. 252-273, 2020.

[26] S. M. G. F. S. M. A. S. R. A. &. K. S. Ainuddin, "Earthquake Vulnerability Assessment through Spatial Multi-Criteria Analysis," *A Case Study of Quetta City, Pakistan.,* 2022.

[27] S. M. Md. ShaharierAlam, "Multi-dimensional earthquake vulnerability assessment of residential neighborhoods of Mymensingh City, Bangladesh: A spatial multi-criteria analysis based approach," *Journal of Urban Management,* pp. 37-58, 2022.

[28] O. G. H. A. M. S. &. E. M. Hamdy, "Identifying exposure of urban area to certain seismic hazard using machine learning and GIS: A case study of greater Cairo.," *Sustainability,* pp. 14(17), 10722, 2022.

[29] M. I. S. L. T. L. M. &. C. M. Brunner, "Challenges in modeling and predicting floods and droughts: A review.," *Wiley Interdisciplinary Reviews: Water,* pp. 8(3), e1520, 2021.

[30] V. G. C. M. M. G. S. R. &. B.-W. S. Deparday, "Machine learning for disaster risk management.," University Of Twente, 2019.

[31] R. P. B. B. G. A.-A. A. &. S. H. Jena, "Seismic hazard and risk assessment: a review of state-of-the-art traditional and GIS models," *Arabian Journal of Geosciences,* pp. 13(2), 1-21, 2020.

[32] E. V. V. &. P. T. Nazarenko, "Features of Application of Machine Learning Methods for Classification of Network Traffic (Features, Advantages, Disadvantages)," in *International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon),* 2019.

[33] K. &. A. T. Dineva, "SYSTEMATIC LOOK AT MACHINE LEARNING ALGORITHMS–ADVANTAGES, DISADVANTAGES AND PRACTICAL APPLICATIONS,"

International Multidisciplinary Scientific GeoConference: SGEM, pp. 20(2.1), 317-324., 2020.

[34] M. A. E. B. N. G. ,. C. B. Mathieu Allimant, "Seismic hazard assessment in Menton, France: Topographical site effect zoning considering a semi-empirical approach and a Machine Learning scheme," in *Third European Conference on Earthquake Engineering and Seismology*, 2022.

[35] K. C. A. &. S. R. D. Khanzode, "Advantages and Disadvantages of Artificial Intelligence and Machine Learning: A Literature Review," *International Journal of Library & Information Science (IJLIS),,* pp. 9(1), 3, 2020.

[36] D. P. N. O. K. C. E. A. M. J. H. R. E. D. S. A. &. I. C. Kontoni, "Gene Expression Programming (GEP) Modelling of Sustainable Building Materials including Mineral Admixtures for Novel Solutions.," *Mining,* pp. 2(4), 629-653., 2022.

[37] C. Ferreira, Gene expression programming: mathematical modeling by an artificial intelligence, Springer, 2006.

[38] P. &. H. D. Sun, " Regional-scale assessment of earthquake-induced slope displacement considering uncertainties in subsurface soils and hydrogeological condition.," *Soil Dynamics and Earthquake Engineering,* pp. 164, 107593., 2023.

[39] V. T. A. J. H. H. A. &. H. J. Nguyen, "Fragility-based seismic performance assessment of modular underground arch bridges," *Structures* , pp. Vol. 39, pp. 1218-1230, 2022, May.

[40] S. G. &. K. S. Jahromi, "The plurality effect of topographical irregularities on site seismic response," *Earthquake Engineering and Engineering Vibration,* pp. 18, 521-534., 2019.

[41] M. A. H. B. B. Mohammad Momeni, "Damage evaluation of H-section steel columns under impulsive blast loads," *Engineering Structures*, 2020.

[42] M. A. A. F. J. M. F. A. H. &. D. A. F. Khan, "New prediction models for the compressive strength and dry-thermal conductivity of bio-composites using novel machine learning algorithms," *Journal of Cleaner Production,* pp. 350, 131364., 2022.

[43] M. N. A. ,. R. ,. K. A. A. ,. I. U. a. M. I. Mohammed Najeeb Al-Hashem, "Predicting the Compressive Strength of Concrete Containing Fly Ash and Rice Husk Ash Using ANN and GEP

Models," *Materials,* pp. 15(21), 7713., 2022.

[44] S. D. Y. A. E. K. &. U. B. Himmetoğlu, "Green building envelope designs in different climate and seismic zones: Multi-objective ANN-based genetic algorithm," *Sustainable Energy Technologies and Assessments,* pp. 53, 102505., 2022.

[45] A. J. A. D. K. M. H. A. A. A. &. Y. K. Mahdiyar, "Practical risk assessment of ground vibrations resulting from blasting, using gene expression programming and Monte Carlo simulation techniques," *Applied Sciences,* pp. 10(2), 472, 2020.

[46] A. S. E. A. C. A. F. &. G. C. Cevik, "Modeling of the uniaxial compressive strength of some clay-bearing rocks using neural network," *Applied Soft Computing,* pp. 11(2), 2587-2594, 2011.