Title
Statistical Methods to Predict Earthquake Damage to Buildings

Author
Mao, Henan

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Statistical Methods to Predict Earthquake Damage to Buildings

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Statistics

by

Henan Mao

2021

ABSTRACT OF THE THESIS

Statistical Methods to
Predict Earthquake Damage to Buildings

by

Henan Mao
Master of Science in Statistics
University of California, Los Angeles, 2021
Professor Yingnian Wu, Chair

Earthquake is the most destructive hazard in building design; base-isolations, as one effective method mitigating the earthquake hazard, are widely used in the building design. However, simulating the building response under earthquake using the physical-based model is time-consuming and undesirable. Therefore, several statistical methods (linear regression, weighted least square, the decision tree, random forest, and neural network) are applied to predict building responses based on the characteristics of applied earthquakes. After principal component analysis, the Statistical models' prediction matches the simulation data very well, indicating that it is promising to utilize statistical methods in predicting the critical building response under earthquake. These predictions provide insightful guidance to the designer.

The thesis of Henan Mao is approved.

Chad Hazlett

Tao Gao

Yingnian Wu, Committee Chair

University of California, Los Angeles

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Professor Yingnian Wu, for his guidance, encouragement, and ingenuity. This dissertation would not have been possible without his steadfast support. I would like to thank Professors Tao Gao and Chad Hazlett for serving on my dissertation committee and their insightful comments. Special thanks to Yike Xu and Dr.Pengyu Chen for their countless assistance during my study.

CHAPTER 1

Introduction to Base Isolation System and Study Object

Earthquakes, as known, are the most important factor causing building damages. There are many techniques developed to mitigate the earthquake hazard, such as mechanical damping, soft stories, and base isolation. Among all these techniques, base isolation is the most effective and widely adopted method [28]. It can reduce seismic demand in building design and reduce injuries to the occupants caused by nonstructural components. There are many contemporary buildings adopting base isolation systems. For example, Apple park, the headquarters of Apple (Figure 1.1) utilizes 700 base isolators in construction. Another example is the Sabiha Gokcen International Airport (Figure 1.2) finish in 2006, which uses 296 base isolators.

The isolators are set between the superstructure and the ground. The base isolation system can elongate the natural period of the superstructure to avoid resonance caused by the earthquake excitation. Four major types of isolators are currently commonly utilized, which are Elastomeric Rubber Bearings, Roller and ball bearings, Steel spring bearings, and Triple Friction pendulum bearings, shown in Figure 1.3. Among them, the Triple Friction pendulum bearings use state-of-the-art technology and provide the highest performance [28]. The base isolation device is originally developed from the idea of "flexible first story" [30, 25, 17, 22], then Penkuhn [35] proposed the friction pendulum concept, and Zayas et al. created the single friction pendulum device [46]. Later on, the Earthquake Protective Systems Corporation (EPS) developed the "Triple Friction Pendulum (TFP)" bearing (Figure

Figure 1.1: Apple Park, Cupertino, California [45].

Figure 1.2: Sabiha Gokcen International Airport, Turkey [42].

1.4), which is an advanced version of the single friction pendulum bearing.

The TFP bearing can dissipate energy through friction and demonstrate high nonlinearity in dynamic structural analysis. Regarding the TFP bearing's uniaxial behavior, Fenz and Constantinou [9, 12], and Morgan and Mahin [32] capture the five-stage behavior of the TFP bearing by taking the equilibrium of each stage to find the force-deformation relationship (Figure 1.5).

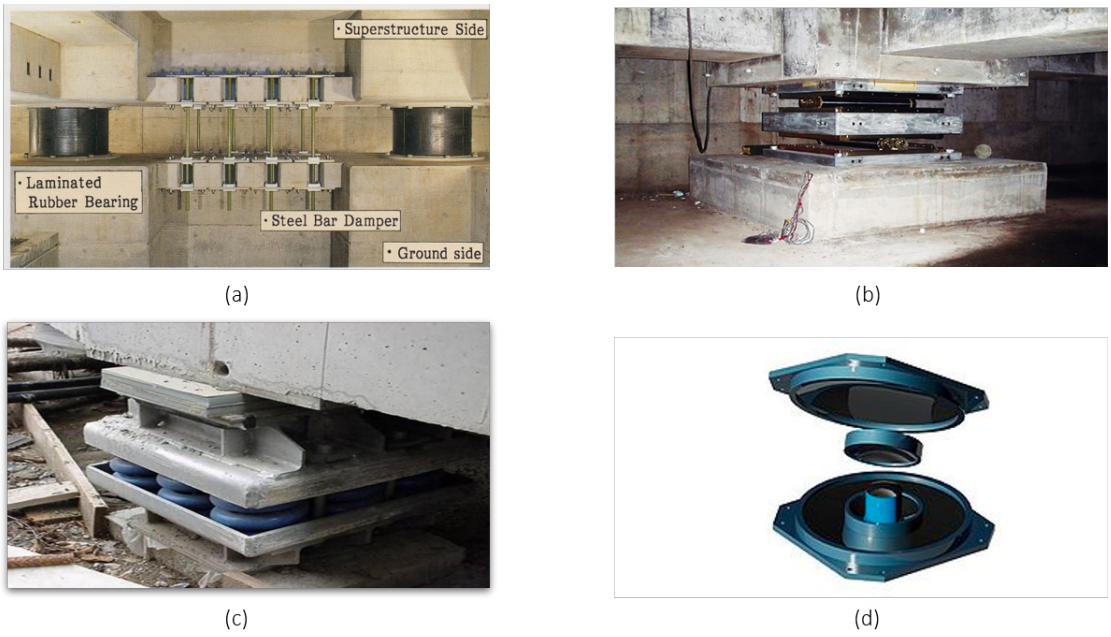

Figure 1.3: Elastomeric rubber bearings (a) [7], Roller and ball bearings (b) [44], steel spring bearings (c) [8], friction pendulum bearings (d) [43].

The biaxial behavior of a TFP bearing is more complex due to the biaxial coupling effect; Becker and Mahin [3] developed a numerical model for TFP bearing based on geometric compatibility. In 2003, Dao et al. [11] implemented a phenomenological TFP bearing model, adopting the idea of a serially connected spring. From another perspective, Mao et al. [29, 28] implemented a physical numerical model in commercial finite element analysis software ABAQUS [41] based on multi-surface plasticity, which is validated through experimental data.

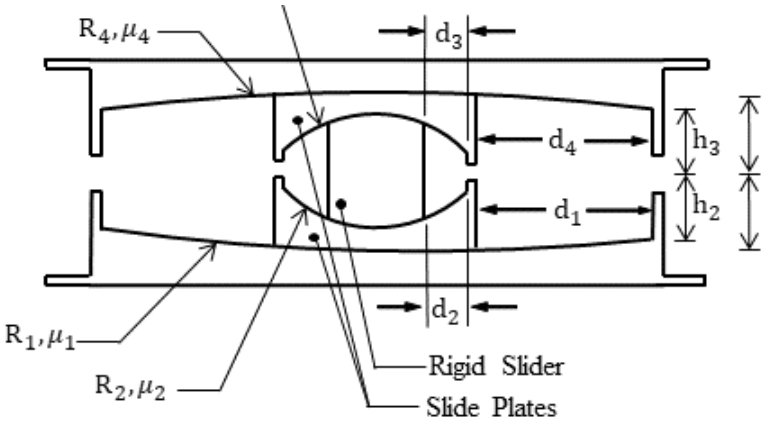

Figure 1.4: Triple Friction Pendulum Bearing section view and basic parameters.

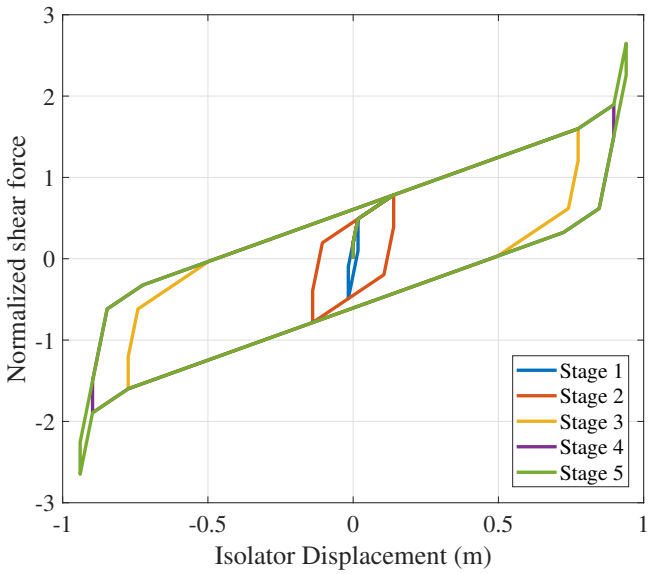

Figure 1.5: The force displacement relationship of TFP bearing.

In order to achieve the high-performance requirement of the base-isolated building, non-linear history structure dynamic analysis is required, then using peak floor acceleration as a significant indicator to decide the damage state of the building. This traditional method is creating the high fidelity finite element model of the building and performs the dynamic analysis to obtain the acceleration records of the floor, which is time-consuming.

This study proposed several statistical methods (linear regression, weighted least square,

the decision tree, random forest, and neural network) to predict the floor accelerations. Therefore, the peak floor acceleration can be treated as the response matrix in this study. On the other hand, there are four essential features to characterize earthquake excitations. They are Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), spectrum acceleration at the first structural period (SaT1), and spectrum displacement at the first structural period (Sd). Thereafter, the design matrix can be constructed by these four features. Several statistical methods (linear regression, weighted least square, the decision tree, random forest, and neural network) are utilized to predict the response of the building in the following chapters and compared with the experimental results. The accurate prediction results can provide insightful guidance for the building designer.

CHAPTER 2

Data Set

As mentioned in the introduction, the dynamic response of the base-isolated structure is complex due to the nonlinearity caused by the TFP bearing. Moreover, conducting the real experimental test is expensive and impractical. Besides, the experimental data is rare and difficult to obtain. Therefore, numerical simulation data is utilized in this study. Mao's [28] numerical model for TFP bearing is adopted due to the stability of this numerical model. The finite element numerical model of the prototype building was created based on the experimental data conducted at E-Defense [11], which is shown in Figure 2.1. In order to ensure the numerical model can represent the real building's characteristics, the modal analysis is performed. The modal analysis results are shown in Table 2.1. The maximum difference between the finite element model and the E-Defense building is 7%, which indicates that the finite element model can capture the characteristics of the E-Defense building accurately and effectively.

The nonlinear dynamic analysis is performed on the E-Defense building using 3189 earthquakes obtained from the PEER NGA-West2 database [34]. After one week of numerical analysis, the peak floor accelerations are obtained. Due to the complexity of the numerical model, a few simulations did not converge at the end, causing the NAN and zeros in the response variable. Since they do not provide useful information, they are removed from the raw data set. The distributions of the features and responses in the revised data set are shown in Figure 2.2 and Table 2.2.

According to HAZUS [18], the initial nonstructural damage happens when the maximum

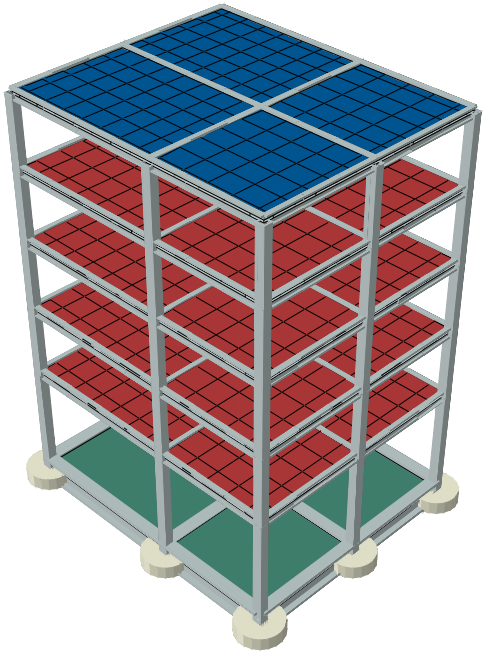

Figure 2.1: The finite element model of E-Defense building.

Table 2.1: Modal analysis of the E-Defense building.

Mode No.	Experiment period T (sec)	Abaqus model period T (sec)	Error (%)
$1^{st}y$ direction mode	0.677	0.674	0.52
2^{nd} y direction mode	0.211	0.223	5.21
3^{rd} y direction mode	0.113	0.119	4.71
1^{st} x direction mode	0.652	0.658	0.96
2^{nd} x direction mode	0.204	0.220	7.18
3^{rd} x direction mode	0.112	0.117	4.56
1^{st} vertical mode	0.142	0.146	2.8
1^{st} rotational mode	0.472	0.475	0.65

Table 2.2: Data description

	Max. floor Acc.	PGA	PGV	SaT1	Sd
mean	2.488	4.034	0.406	0.174	0.041
std	1.702	4.295	0.520	0.306	0.051
min	0.012	0.050	0.001	0.000	0.000
25%	1.209	1.281	0.090	0.026	0.011
50%	2.119	2.649	0.217	0.070	0.025
75%	3.425	5.128	0.507	0.192	0.052
max	10.038	39.036	4.453	3.853	0.678

floor acceleration is beyond 2 m/sec^2; the building will collapse when the acceleration is beyond 20 m/sec^2. Most of the peak floor accelerations (Mac_acc) are below 6 m/sec^2, shown in Figure 2.2, which is reasonable since this study only does not include the collapse buildings. The pairwise relationships of features and responses are shown in Figure 2.3. The plots in the first column demonstrate that all the features have a positive relationship with the response. This is expected because the severer the earthquake is, the higher the maximum floor acceleration will be. Besides, other plots in Figure 2.3 and the correlation plot shown in Figure 2.4 indicate that there is also a clear positive correlation between each pair of features. The correlation between SaT1 and PGV is as high as 0.86. The highly correlated features may cause the multicollinearity problem. To solve this, Principal Component Analysis is applied and will be discussed in the next chapter.

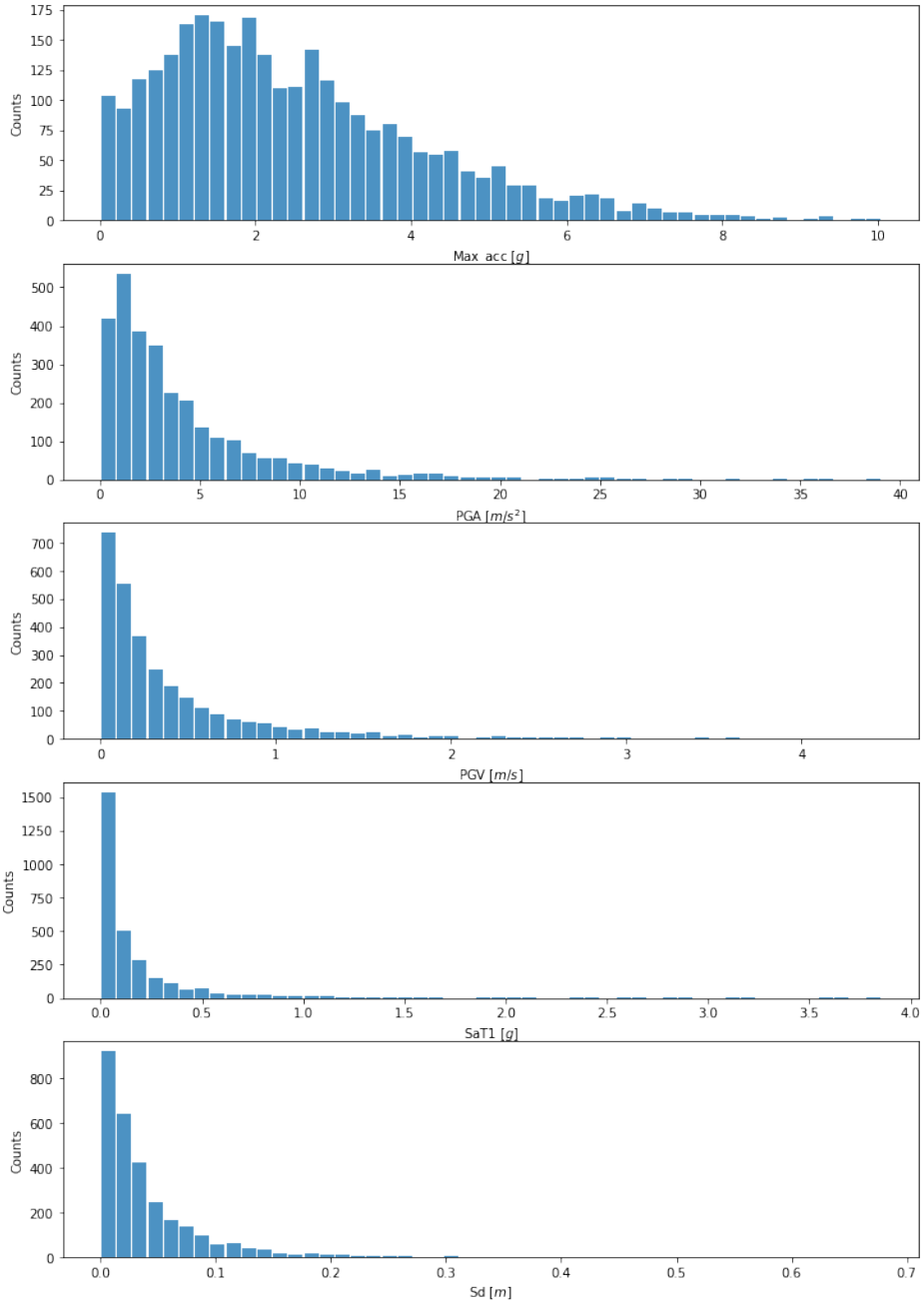

Figure 2.2: The distributions of the features and responses.

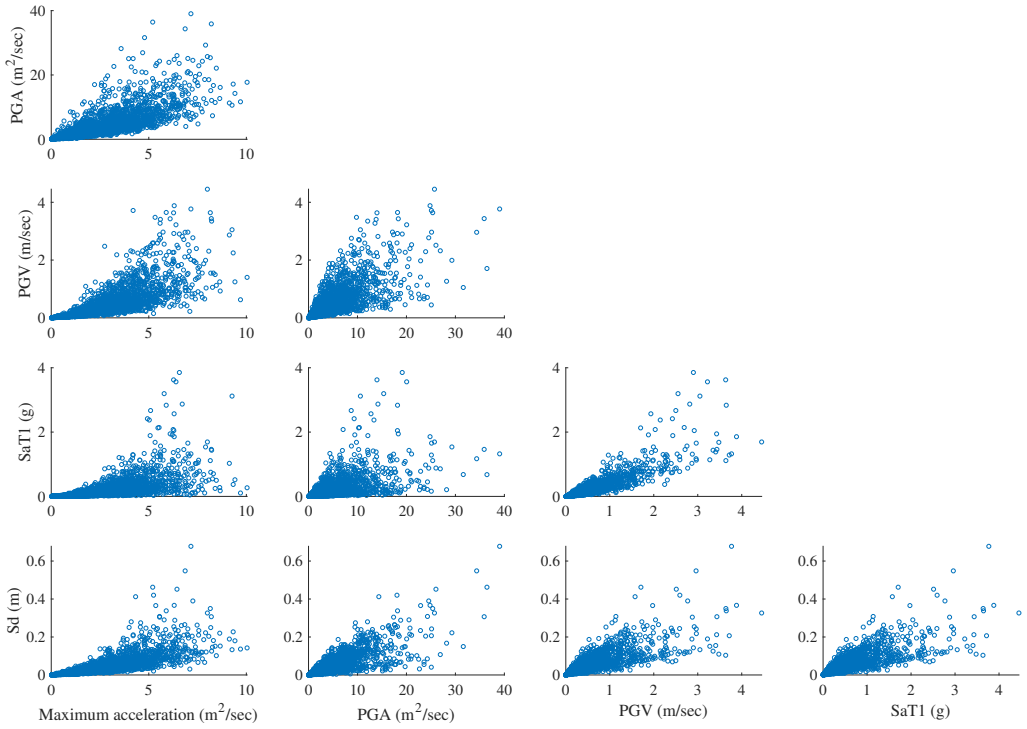

Figure 2.3: The pairwise scatter plot of features and response.

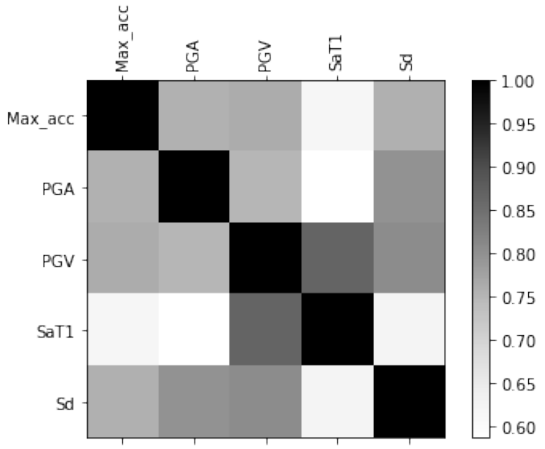

Figure 2.4: The correlation of features and response.

CHAPTER 3

Method and Results

3.1 Linear Regression

Linear regression method is the most widely used method to predict the numeric response. The design matrix is composed of a $n \times p$ matrix $\mathbf{X} = (x_{ij})$, and the response is a $n \times 1$ vector $\mathbf{Y} = (y_i)$. The model can be expressed in the following form:

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_j + \epsilon_i, \tag{3.1}$$

for $i = 1, ..., n$, where $\epsilon_i \sim \mathrm{N}(0, \sigma^2)$ independently for $i = 1, ..., n$. The Equation 3.1 can be written in the matrix form as following:

$$\mathbf{Y} = \mathbf{X}^T\beta + \epsilon. \tag{3.2}$$

The least squares estimate of β is

$$\hat{\boldsymbol{\beta}} = \mathbf{arg\,min}_{\boldsymbol{\beta}} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|_{\boldsymbol{\ell_2}}^{\mathbf{2}} = (\mathbf{X}^{\boldsymbol{T}}\mathbf{X})^{-\mathbf{1}}\mathbf{X}^{\boldsymbol{T}}\mathbf{Y}. \tag{3.3}$$

In order to evaluate the goodness of the model, the Mean Squared Error (MSE) (Eq.3.4) and the Mean absolute error (MAE) (Eq.3.5) are used. The Mean Squared Error can penalize the large prediction errors, but is susceptible to data with various outliers. This disadvantage can be mitigated by evaluating the Mean absolute error.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \tag{3.4}$$

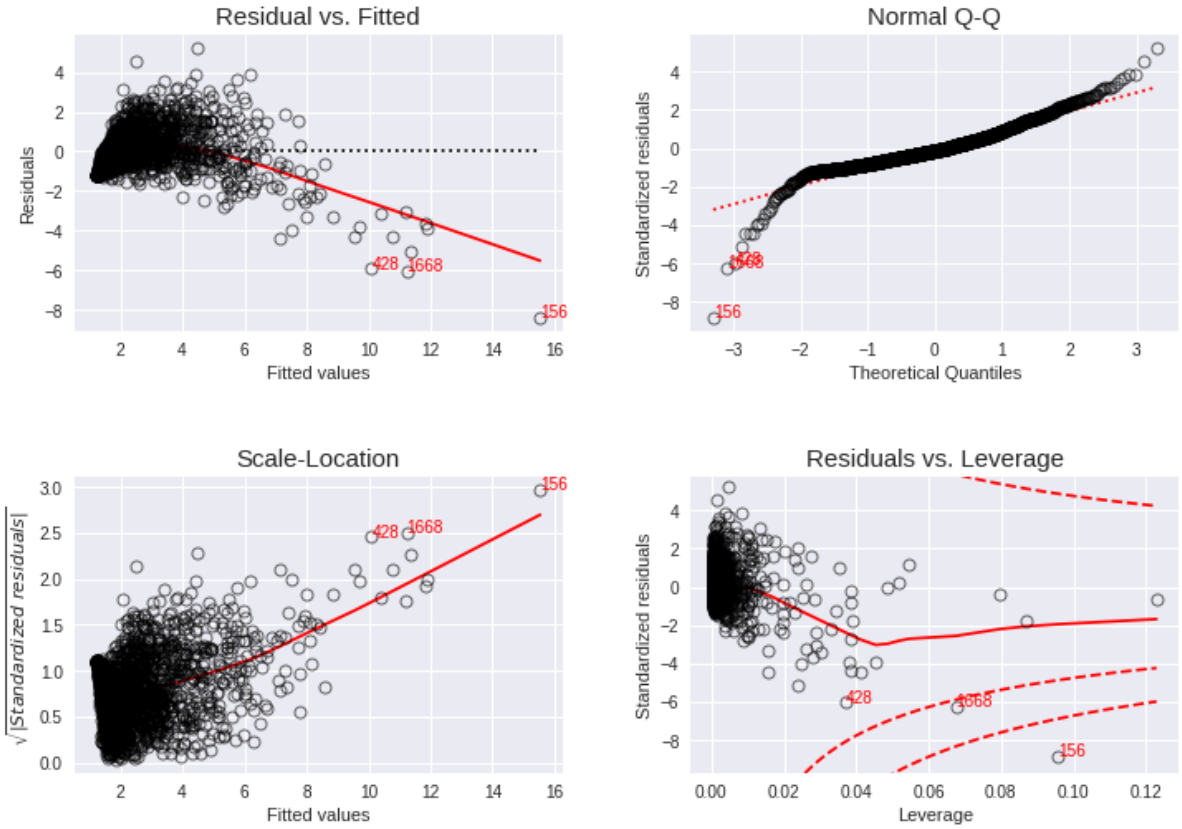

Figure 3.1: The linear regression results.

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \tag{3.5}$$

3.2 Linear Regression Results

The simulation data has been split into train set (70%) and test set (30%), then the linear regression analysis has been performed, the MSE and MAE are 0.819 and 0.681, separately. The detailed results are shown in Figure 3.1. The residual is not evenly distributed, which indicates not a good fit. On the other hand, the normal Q-Q plot demonstrates that the error is not satisfying the normal distribution assumption. Besides, the heteroscedasticity is observed through the Residue v.s. leverage plot.

Table 3.1: The multicollinearity of features.

Features	Const	PGA	PGA	SaT1	Sd
VIF Factor	1.9	3.1	8.0	4.1	4.1

The multicollinearity is checked by evaluating the Variance Inflation Factor (VIF) [23]. The VIF result is shown in Table 3.1, typically, multicollinearity is considered as high if $\text{VIF}(\hat{\beta}_i) > 10$. The maximum VIF comes from the PGV, which is 8.0. Therefore, multicollinearity is not the key issue.

3.3 Principal Component Analysis (PCA)

Principal component analysis was first proposed by Karl Pearson [33] and Harold Hotelling [21]. PCA is widely used for dimensionality reduction by projecting the dataset onto new axes, then use fewer components to represent the original dataset, and the correlations of new components are zeros since new axes are orthogonal to each other. Because high correlations of the design matrix are observed in the previous chapter, the principal component analysis is performed in this section.

The first step of PCA is centralizing the design matrix (X),i.e., $\sum_{i=1}^{n} x_{ij}/n = 0$. All the component X_i need to be represented by new basis system Q, let $\mathbf{Z}$ be the design matrix in Q. Therefore, $X_i = QZ_i$. Each column of (X) is orthogonal to each other by definition. Let $\lambda_j = \|\mathbf{Z}_j\|^2/n = \sum_{i=1}^{n} z_{ij}^2/n$, $\mathbf{Z}^T\mathbf{Z} = \Lambda = \text{diag}(\lambda_1, ..., \lambda_p)$. The Q and $\mathbf{A}$ can be obtained by power method through Equation 3.6

$$\mathbf{X}^T\mathbf{X} = Q\mathbf{Z}^T\mathbf{Z}Q^T = Q\Lambda Q^T \tag{3.6}$$

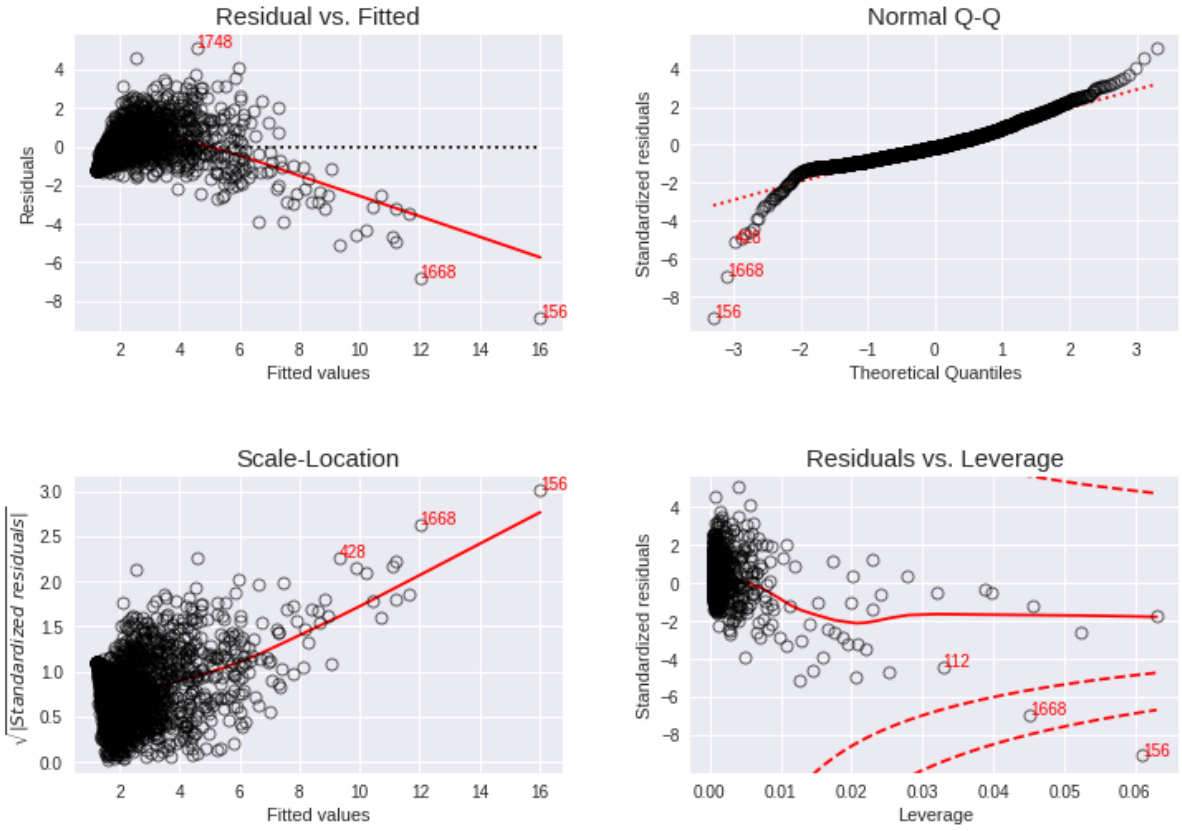

Figure 3.2: The linear regression results after PCA.

3.4 PCA Results

After performing the principal component analysis, Figure 3.3 explains the correlation between design matrix are zeros after transforming to the new basis system. Additionally, Figure 3.4 demonstrates that the first two principal components can explain 92.5% of the total variance. Therefore, the four features can be reduced to two principal components in the following study.

The linear regression was applied to the new design matrix again, utilizing only the first two PCs. Based on the results of Figure 3.2, the improvement after adopting PCA is not significant; the error term violates the normal distribution assumption, and outliers still exist. Therefore, in order to obtain a better prediction, an advanced model is required.

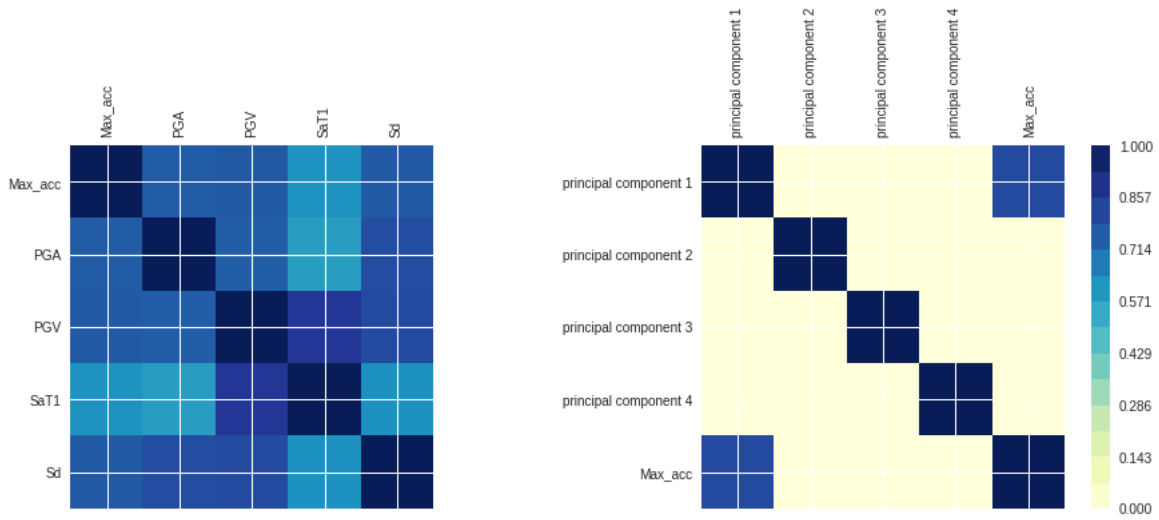

Figure 3.3: The correlation of features and response before PCA (left), after PCA (right).

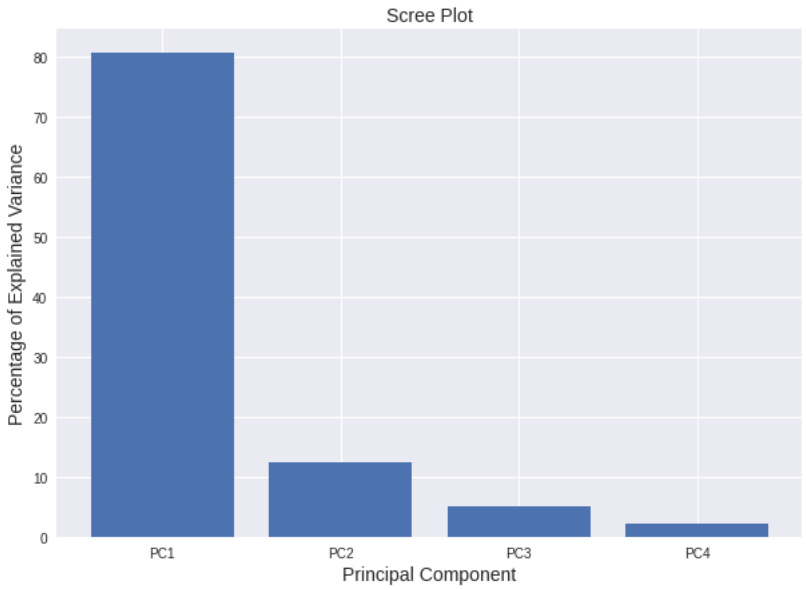

Figure 3.4: Percentage of variance (information) for each by PC.

3.5 Weighted Least Squares (WLS)

The weighted least squares model is designed for the situation that variances of the observations are unequal distributed (heteroscedasticity). The coefficient of the WLS model can be obtained by minimizing the weighted sum of squares.

$$\arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} w_i \left| y_i - \sum_{j=1}^{m} X_{ij}\beta_j \right|^2 = \arg\min_{\boldsymbol{\beta}} \left\| W^{\frac{1}{2}}(\mathbf{y} - X\boldsymbol{\beta}) \right\|^2. \tag{3.7}$$

Where W is the diagonal matrix, and w_i is the weight of the i^{th} observation. Because the weight matrix W is unknown, the Iteratively Reweighed Least Squares (IRLS) algorithm [6] is adopted in this study. The first step of the IRLS algorithm is initialing the parameters as following equations.

$$\begin{aligned}
\mu^{(0)} &= Y \\
\eta^{(0)} &= g(\eta^{(0)}) \\
z^{(0)} &= \eta^{(0)} + \left(Y - \mu^{(0)}\right) \cdot \frac{d\eta^{(0)}}{d\mu^{(0)}} \\
w_i^{(0)} &= \frac{1}{\widehat{\mathrm{Var}}^{(0)}\left(z_i^{(0)}\right)} = \frac{1}{\widehat{\mathrm{Var}}^{(0)}\left(Y_i\right) \cdot \left[g'\left(\mu_i^{(0)}\right)\right]^2}
\end{aligned} \tag{3.8}$$

The parameters are updated through the following equations.

$$\begin{aligned}
\beta^{(1)} &= \left(X^T W^{(0)} X\right)^{-1} X^T W^{(0)} z^{(0)} \\
\eta^{(1)} &= X^T \beta^{(1)} \\
\mu^{(1)} &= g^{-1}\left(\eta^{(1)}\right)
\end{aligned} \tag{3.9}$$

Finally, check whether the residue satisfies the requirement; otherwise, go back to the second step, e.g., $\left\|\beta^{(k+1)} - \beta^{(k)}\right\|_2 < 10^{-5}$.

After applying the IRLS algorithm, the MSE and MAE are obtained as 0.927 and 0.645, separately. These results are similar to the linear regression; in other words, no significant improvement was observed by using weighted least squares. Figure 3.5 suggests that the

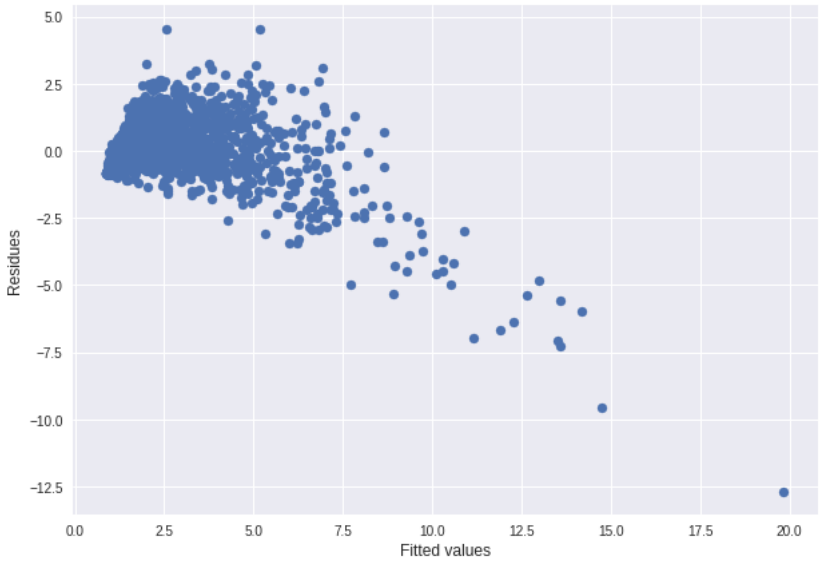

Figure 3.5: Residual v.s. fitted by IRLS algorithm.

residual distribute unevenly, which indicates that the issue exists linear regression still has not been solved.

3.6 Decision Tree

The decision tree is a flow-chart-like model, which is demonstrated in Figure 3.6; each branch represents intermediate results, the path represents the decision rule, all the decisions made at the leaf node. The decision tree model is simple to interpret and versatile. It works even with not enough data.

The decision tree model is applied to predict the maximum floor acceleration in this study. In order to avoid overfitting, cross-validation [1, 39, 40] process is applied. The train data has been split into five folds; for each fold of data, the maximum depth of the tree varies from 1-50. This procedure will help decide the best choice of the maximum depth of the tree. Figure 3.7 and Figure 3.8 suggest that the best choice of the maximum depth of the tree is 4. After obtaining this parameter, MSE and MAE are evaluated by the test

data. Figure 3.9 demonstrates the residual distribution, MSE and MAE equal to 0.508 and 0.506 separately. These results indicate a notable improvement in the predicting accuracy comparing to the weighted least squares. The MSE and MAE decrease by 45.0% and 21.6%, separately. However, the residual distribution plot (Figure 3.9) indicates that the error distributes unevenly; the additional investigation is needed.

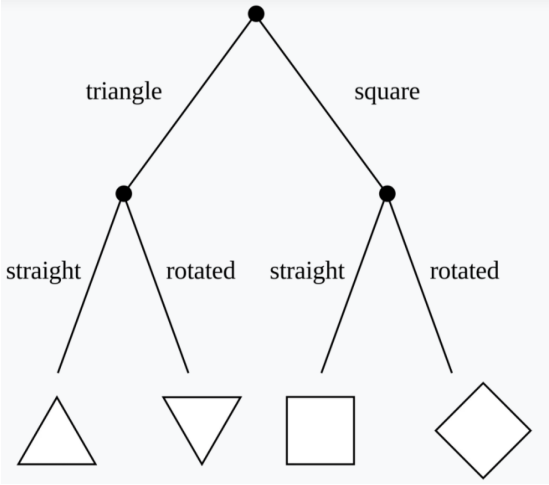

Figure 3.6: Simple Decision Tree [10].

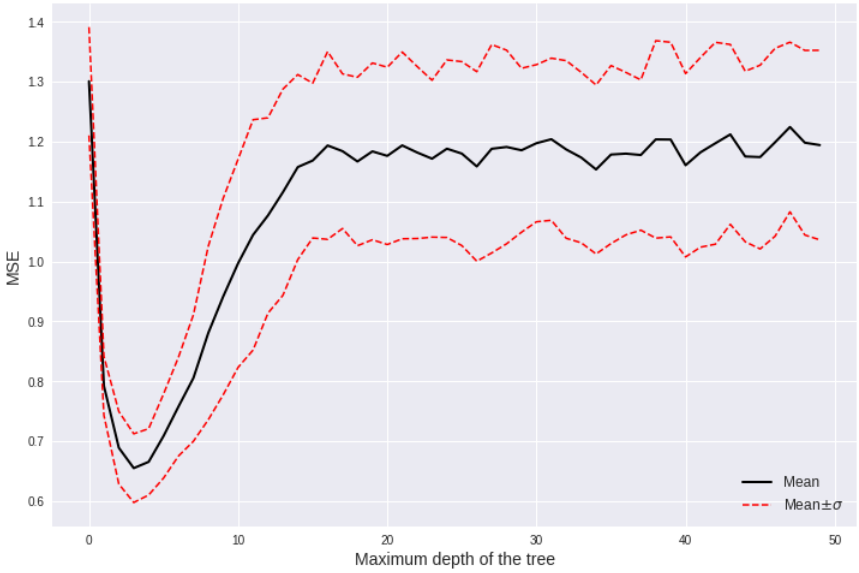

Figure 3.7: MSE v.s. the maximum depth of the tree.

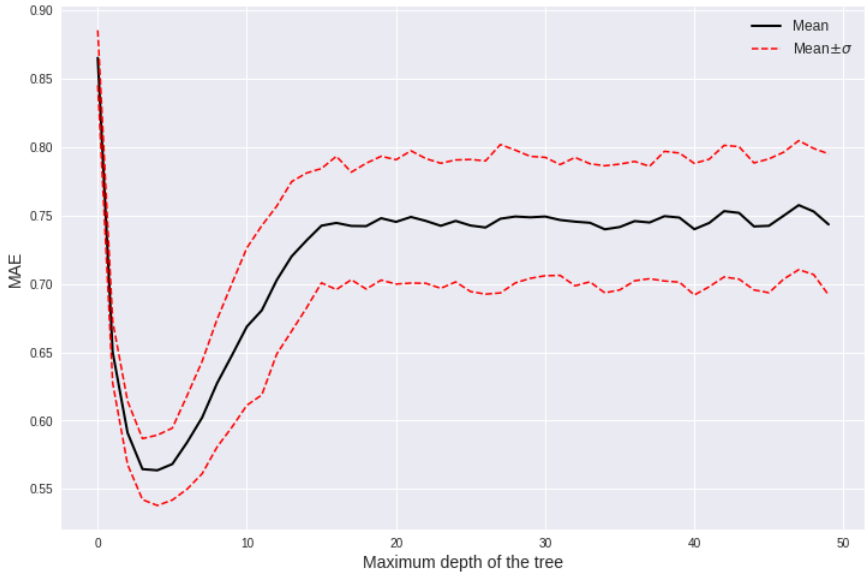

Figure 3.8: MAE v.s. the maximum depth of the tree.

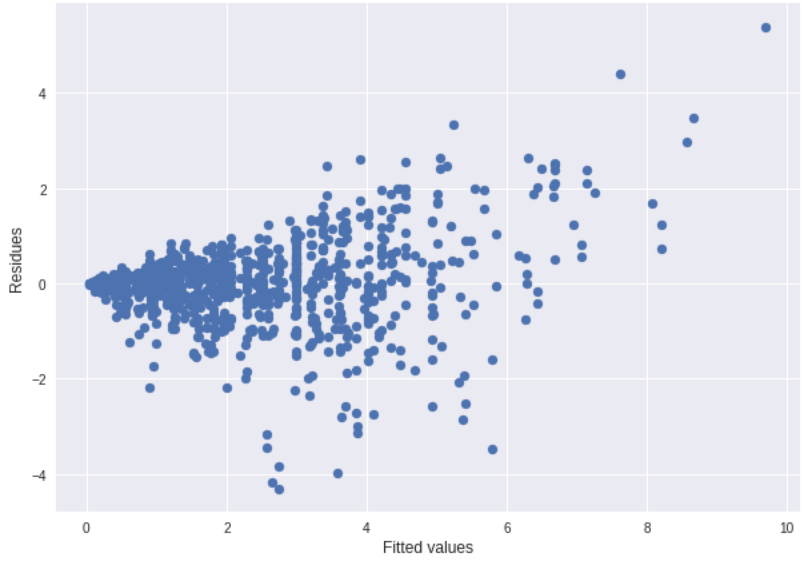

Figure 3.9: Residual v.s. fitted by Decision Tree.

3.7 Random Forest

Random forest was first proposed by Tin Kam Hoin 1995 [19, 20] and extended by Leo Breiman [4] and Adele Cutler [26]. Random forest is a statistical method based on the deci-

sion tree algorithm, and the random forest algorithm predicts the average values of multiple decision trees. This average process makes random forest less susceptible to overfitting issues, which happens in decision tree [15]. Random forests perform better than decision tree in general, but the performance of random forest also depend on data characteristics [36, 27].

Similarly, the five-fold cross-validation procedure is applied to the random forest algorithm. In order to estimate the best value of the tree number, the tree number varies from 1-50 for each fold of the analysis. After performing 250 analysis, the relationship between error and the number of the tree is shown in Figure 3.10 and Figure 3.11. They indicate 48 trees will achieve the highest accuracy, then plug the tree number back to the random forest algorithm, MSE, and MAE equal to 0.576 and 0.533 separately. The MSE and MAE are close to the decision tree; this could cause by the characteristics of the data. The random forest also can predict the peak floor acceleration with high accuracy. The residue distribution plot (Figure 3.12) has a similar issue with the decision tree, which is not surprising since the decision tree is a type of decedent of the random forest.

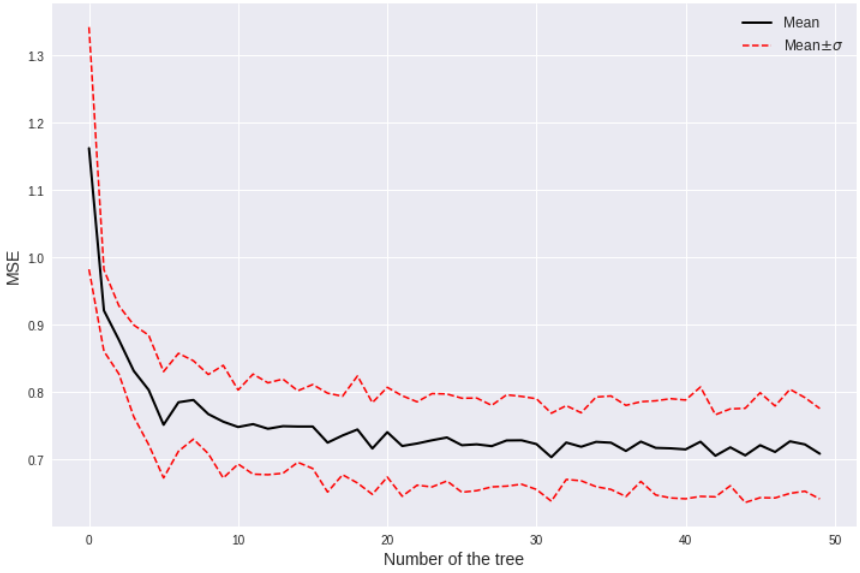

Figure 3.10: MSE v.s. number of the tree.

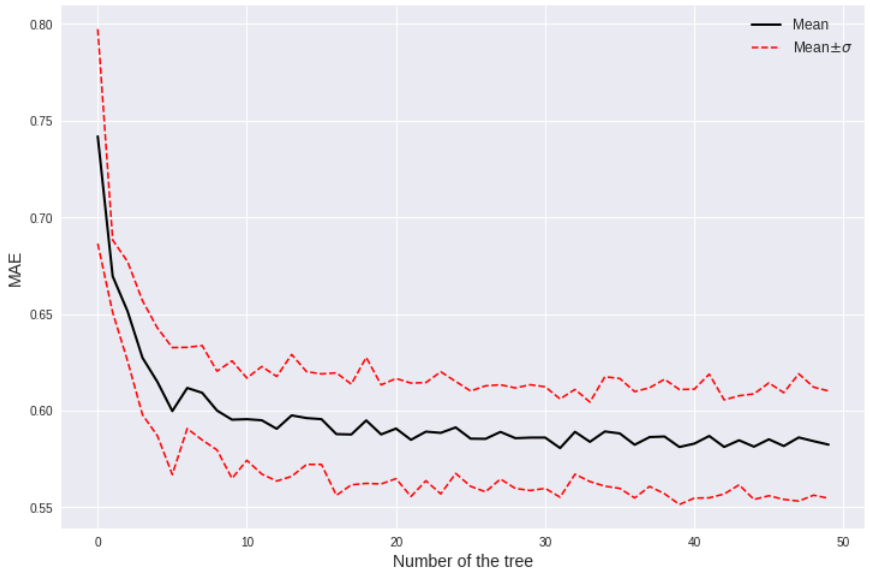

Figure 3.11: MAE v.s. number of the tree.

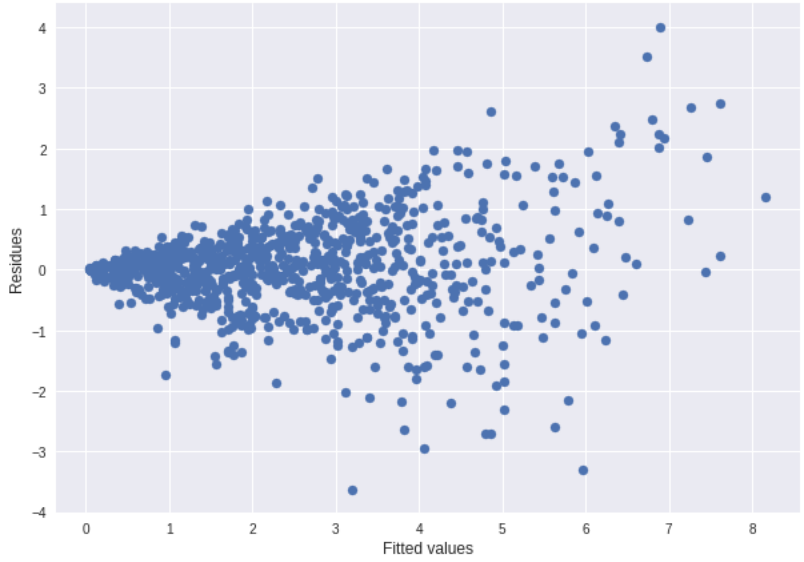

Figure 3.12: Residual v.s. fitted by Random Forest.

3.8 Neural Network

The theory of neural networks was proposed by Alexander Bain [2], and William James [24] independently, then validated by C. S. Sherrington's experiment [38]. In 1943, the compu-

tational model based on mathematics was created by McCulloch and Pitts [31]. Figure 3.13 demonstrates a simple one hidden layer neural network. The inputs shown on the left side are the features; the hidden layers transfer features to the other space that the hyperplane can easily separate. Each layer can be expressed by Equation 3.10 and Equation 3.11.

$$h_l = f_l(s_l) \tag{3.10}$$

$$s_l = W_l h_{l-1} + b_l, \text{for } l = 1, ...L, \tag{3.11}$$

where f_l denotes the activation function for the hidden layer, such as the logistic function, sigmoid function and the rectified linear unit function (ReLU). $h_0 = X$, which is the design matrix, W_l and b_l are the weight matrix and bias vector respectively. $\frac{\partial L}{\partial W_{l,k}}$, $\frac{\partial L}{\partial b_l}$ and, $\frac{\partial L}{\partial W_l}$ which are necessary for updating the neural network, can be obtained through chain rule.

$$\begin{aligned}
\frac{\partial L}{\partial h_{l-1}^\top} &= \sum_{k=1}^{d} \frac{\partial L}{\partial h_{l,k}} \frac{\partial h_{l,k}}{\partial s_{l,k}} \frac{\partial s_{l,k}}{\partial h_{l-1}^\top} \\
&= \sum_{k=1}^{d} \frac{\partial L}{\partial h_{l,k}} f_l'\left(s_{l,k}\right) W_{l,k} \\
&= \frac{\partial L}{\partial h_l^\top} f_l' W_l
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
\frac{\partial L}{\partial W_{l,k}} &= \frac{\partial L}{\partial h_{l,k}} \frac{\partial h_{l,k}}{\partial s_{l,k}} \frac{\partial s_{l,k}}{\partial W_{l,k}} \\
&= \frac{\partial L}{\partial h_{l,k}} f_l'\left(s_{l,k}\right) h_{l-1}^\top,
\end{aligned} \tag{3.13}$$

$$\frac{\partial L}{\partial W_l} = f_l' \frac{\partial L}{\partial h_l} h_{l-1}^\top. \tag{3.14}$$

$$\frac{\partial L}{\partial b_l} = f_l' \frac{\partial L}{\partial h_l}. \tag{3.15}$$

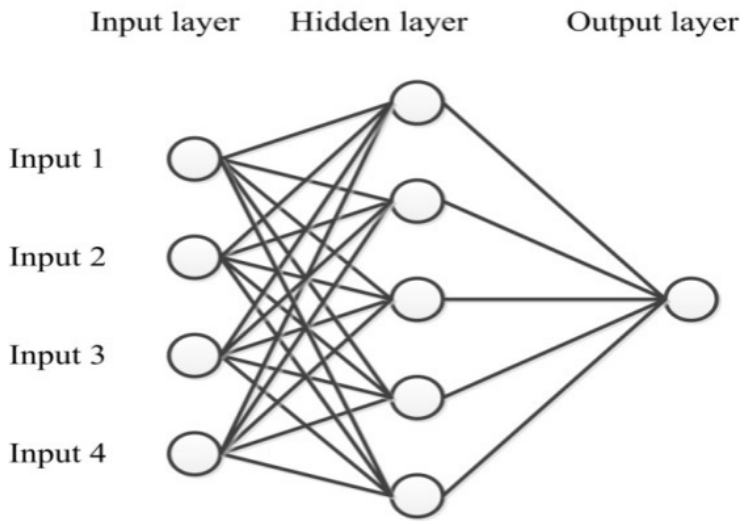

Figure 3.13: Neural networks

3.9 Neural Network Results

Similar to the decision tree and random forest, cross-validation has been applied to the neural network also. The number of hidden layers varies from 1-50 in each fold. Figure 3.14 and Figure 3.15 show that MSE and MAE vary with the number of hidden layers; the minimum error happened when 46 hidden layers were adopted. In addition, the choice of solver and activation function also affect the performance of the neural network. The analysis utilizing different solves and activation functions are performed. The results, shown in Table 3.2 and Table 3.3 suggest that the Broyden–Fletcher–Goldfarb–Shanno (BFGS) [14, 5, 13, 16, 37] method with Tanh activation function achieve the highest accuracy. After finalizing those parameters, then applying these parameters into the neural network model, the MSE and MAE are 0.456 and 0.483, respectively; the residual distribution is shown in Figure 3.16.

The MSE and MAE for all models are summarized in Table 3.4, all statistical models predict the peak floor acceleration accurately and efficiently. Among the five statistical models, the neural network achieves the highest accuracy. Comparing to the decision tree, the MSE and MAE decrease by 10.2% and 4.6%, separately. These results indicate that the

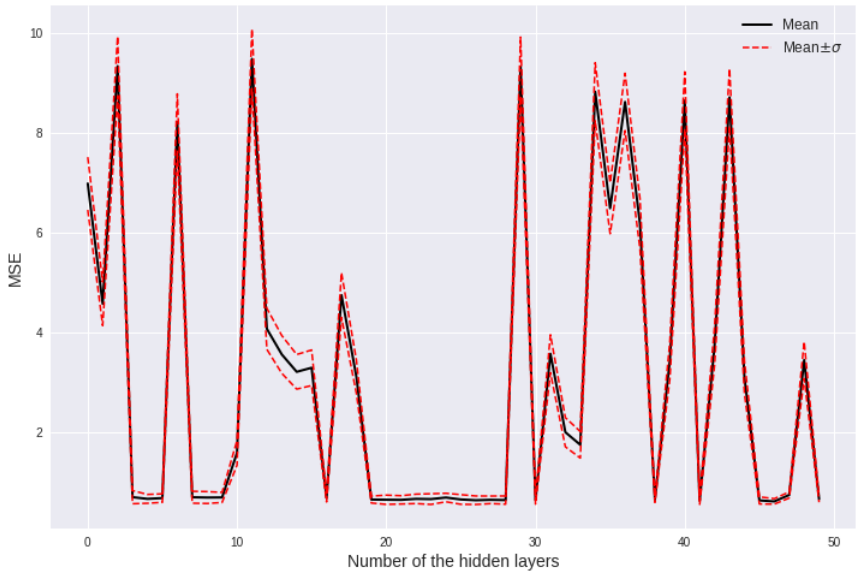

Figure 3.14: MSE v.s. number of the hidden layers.

building response under earthquakes can be predicted by statistical models accurately.

Regarding the residual distribution plot (Figure 3.16), similar issues are observed in all other statistical models, which could cause by the characteristics of the data. In order to solve the uneven residual distribution issue, other advanced statistical models are required.

Table 3.2: The MSE of neural network.

Solver \ Activation	Logistic	Tanh	Relu
lbfgs	0.456	**0.456**	0.486
sgd	0.646	0.480	0.471
adam	1.349	1.261	0.463

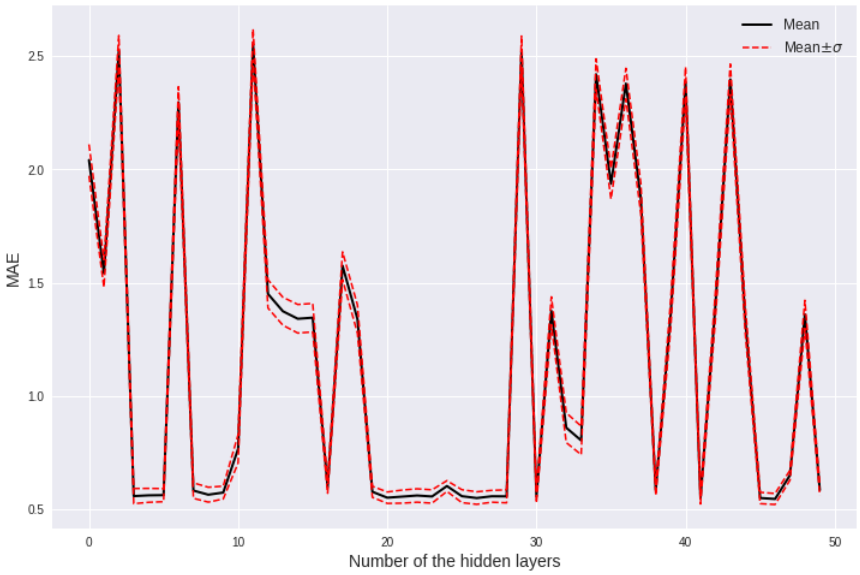

Figure 3.15: MAE v.s. number of the hidden layers.

Table 3.3: The MAE of neural network.

Solver \ Activation	Logistic	Tanh	Relu
lbfgs	0.487	**0.483**	0.484
sgd	0.621	0.510	0.500
adam	0.762	0.691	0.498

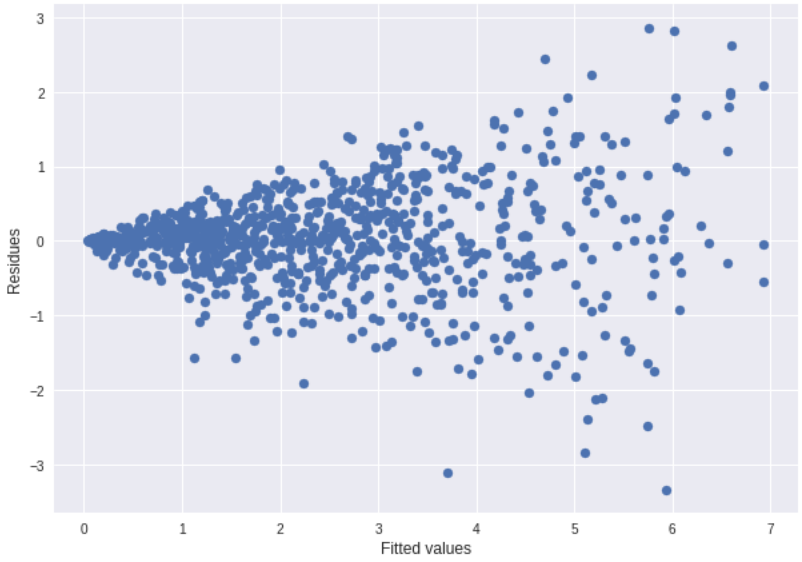

Figure 3.16: Residual v.s. fitted by Neural Networks.

Table 3.4: The MSE & MAE of all models.

Method \ Error type	MSE	MAE
Linear regression	0.819	0.685
Weighted least square	0.927	0.645
Decision tree	0.508	0.506
Random forest	0.577	0.533
Neural network	**0.456**	**0.483**

CHAPTER 4

Conclusions and Suggested Future Works

This study aims at predicting the maximum floor acceleration based on the characteristics of earthquakes. Linear regression result implies the existence of heteroscedasticity. After PCA analysis, the improvement of the linear regression utilizing two principal components is negligible; the weighted least square results are similar to the linear regression; therefore, three advanced models are adopted. The decision tree predicts better results than the two previous linear models. However, the MSE and MAE estimate from the random forest is higher than the decision tree, which indicates that the data characteristics affect the performance. The most accurate model generates from the neural network.

All the statistical models reasonably estimate the floor acceleration; the neural network is best among all the statistical models. It is promising to adopt statistical methods in real building design to estimate earthquake damage efficiently and accurately.

The distribution of residuals indicates that no one model can explain the simulation results perfectly; the residual increases when the predicted value increase. One reason causing this issue is the complexity of the building; there are only four features adopted in this study, which is not enough for this complex situation. The second issue comes from the parameter tuning of the statistical methods. The parametric study of the parameters for each model is not exhausted, so further investigation is needed. Additionally, this study does not investigate other physical-based statistical methods which could better predict the maximum floor acceleration.

REFERENCES

[1] David M Allen. The relationship between variable selection and data agumentation and a method for prediction. *technometrics*, 16(1):125–127, 1974.

[2] Alexander Bain. *Mind and body: The theories of their relation*, volume 4. D. Appleton, 1873.

[3] Tracy C Becker and Stephen A Mahin. Experimental and analytical study of the bi-directional behavior of the triple friction pendulum isolator. *Earthquake Engineering & Structural Dynamics*, 41(3):355–373, 2012.

[4] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[5] Charles George Broyden. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970.

[6] Rick Chartrand and Wotao Yin. Iteratively reweighted algorithms for compressive sensing. In *2008 IEEE international conference on acoustics, speech and signal processing*, pages 3869–3872. IEEE, 2008.

[7] CivilDigital. Elastomeric rubber bearings. `https://civildigital.com/base-isolation-system-outline-on-principles-types-advantages-applications/elastomeric-rubber-bearings/`, 2021. [Online; accessed 15-July-2021].

[8] CivilDigital. Spring isolators. `https://civildigital.com/base-isolation-system-outline-on-principles-types-advantages-applications/spring-isolators/`, 2021. [Online; accessed 15-July-2021].

[9] Michalakis Constantinou, Anoop Mokha, and Andrei Reinhorn. Teflon bearings in base isolation ii: Modeling. *Journal of Structural Engineering*, 116(2):455–474, 1990.

[10] corporate finance institute. Simple decision tree. `https://corporatefinanceinstitute.com/resources/knowledge/other/decision-tree/`, 2021. [Online; accessed 2015].

[11] Nhan D Dao, Keri L Ryan, Eiji Sato, and Tomohiro Sasaki. Predicting the displacement of triple pendulum™ bearings in a full-scale shaking experiment using a three-dimensional element. *Earthquake Engineering & Structural Dynamics*, 42(11):1677–1695, 2013.

[12] Daniel M Fenz and Michael C Constantinou. Spherical sliding isolation bearings with adaptive behavior: Experimental verification. *Earthquake engineering & structural dynamics*, 37(2):185–205, 2008.

[13] Roger Fletcher. A new approach to variable metric algorithms. *The computer journal*, 13(3):317–322, 1970.

[14] Roger Fletcher. Practical methods of optimization john wiley & sons. *New York*, 80:4, 1987.

[15] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[16] Donald Goldfarb. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24(109):23–26, 1970.

[17] Norman B Green. Flexible first-story construction for earthquake resistance. *Transactions of the American Society of Civil Engineers*, 100(1):645–652, 1935.

[18] Hazus-MH FEMA. *Multi-hazard loss estimation methodology earthquake model*, 2003.

[19] Tin Kam Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE, 1995.

[20] Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20(8):832–844, 1998.

[21] Harold Hotelling. Analysis of a complex of statistical variables with principal components. *J. Educ. Psy.*, 24:498–520, 1933.

[22] Lydik S Jacobsen and Robert S Ayre. Experimentally determined dynamic shears in a sixteen-story model. *Bulletin of the seismological society of America*, 28(4):269–311, 1938.

[23] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[24] William James. *The principles of psychology*, volume 1. Cosimo, Inc., 2007.

[25] Lee H Johnson, Edward J Bednarski, Merit P White, Paul L Kartzke, Howard G Smits, and HA Williams. Discussion of "johnson on flexible first-story construction". *Transactions of the American Society of Civil Engineers*, 100(1):653–669, 1935.

[26] Andy Liaw and M Wiener. Documentation for r package randomforest. *PDF). Retrieved*, 15:191, 2013.

[27] S Madeh Piryonesi and Tamer E El-Diraby. Using machine learning to examine impact of type of performance indicator on flexible pavement deterioration modeling. *Journal of Infrastructure Systems*, 27(2):04021005, 2021.

[28] H Mao, H Darama, and E Taciroglu. A consistent omnidirectional nonlinear hysteretic response model for triple friction pendulum bearings. In *Proceedings of the 17th World Conference on Earthquake Engineering*, Sendai International Center, Sendai, Japan, 2020.

[29] H Mao, Z Ni, H Darama, and E Taciroglu. Consistent omnidirectional extension of uniaxial nonlinear hysteretic response models. In *Proceedings of the 11th National Conference on Earthquake Engineering*, Sendai International Center, Los Angeles, California, 2018.

[30] RR Martel. The effects of earthquakes on buildings with a flexible first story. *Bulletin of the Seismological Society of America*, 19(3):167–178, 1929.

[31] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[32] Troy A Morgan and Stephen A Mahin. The optimization of multi-stage friction pendulum isolators for loss mitigation considering a range of seismic hazard. In *The 14th World Conference on Earthquake Engineering*, pages 11–0070. Citeseer, 2008.

[33] K Pearson and LIII On Lines. Planes of closest fit to systems of points in space, london edinburgh dublin philos. mag. *J. Sci*, 2(11):559–572, 1901.

[34] Peer. Peer strong motion database, 2011.

[35] A. L. K. Penkuhn. Three point foundation for building structures, 1967. US Patent No. 3347002.

[36] S Madeh Piryonesi and Tamer E El-Diraby. Role of data analytics in infrastructure asset management: Overcoming data size and quality problems. *Journal of Transportation Engineering, Part B: Pavements*, 146(2):04020022, 2020.

[37] David F Shanno. Conditioning of quasi-newton methods for function minimization. *Mathematics of computation*, 24(111):647–656, 1970.

[38] Charles S Sherrington. Experiments in examination of the peripheral distribution of the fibers of the posterior roots of some spinal nerves. part ii. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 190:45–186, 1898.

[39] Mervyn Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*, 36(2):111–133, 1974.

[40] Mervyn Stone. An asymptotic equivalence of choice of model by cross-validation and akaike's criterion. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):44–47, 1977.

[41] Dassault Systémes. *ABAQUS*. Cambridge, MA, 1985.

[42] Earthquake Protection Systems. Earthquake protection systems (eps) qualifications. http://www.ci.vallejo.ca.us/common/pages/DisplayFile.aspx?itemId=120632, 10 2014. [Online; accessed 15-July-2021].

[43] Earthquake Protection Systems. Friction pendulum bearings. https://www.earthquakeprotection.com/products, 2020. [Online; accessed 15-July-2021].

[44] Andrew Tarantola. How buildings stay up when the earth shakes. https://www.gizmodo.com.au/2011/08/how-buildings-stay-up-when-the-earth-shakes/, 8 2011. [Online; accessed 10-July-2021].

[45] Mahin Zayas. Continued functionality standard seismic isolator standard. https://www.earthquakeprotection.com/standards, 2 2021. [Online; accessed 15-July-2021].

[46] VA Zayas, SS Low, and SA Mahin. The fps earthquake resisting system. experimental report no. ucb/eerc 87/01, eerc. *University of California, Berkeley*, 1987.