Rock failure prediction under true triaxial loading: Evaluation of statistical techniques and machine learning based techniques

Anele Sambo

A research report submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Science in Engineering.

Johannesburg, 2019

DECLARATION

I declare that this report is my own, unaided work. I have read the University Policy on Plagiarism and hereby confirm that no Plagiarism exists in this report. I also confirm that there is no copying nor is there any copyright infringement. I willingly submit to any investigation in this regard by the School of Mining Engineering and I undertake to abide by the decision of any such investigation.

A.N. Sambo — 05 February 2019

(Signature of candidate) — (Date)

At Wits University

ABSTRACT

Understanding the stress changes within the vicinity of an excavation is critical in mining environments and should be accounted for by rock engineers. These changes in stresses have a substantial influence on the strength and deformational properties of rocks and can result to instabilities and rock failure. Rock failure is considered as one of the phenomenon having dire consequences to mining operations. Due to the repercussion of rock failure, rocks are subjected to laboratory testing with the aim of determining the stresses that are likely to result to failure as well as measuring the rock strength. Normally, conventional testing is conducted which focuses on the major and minor principal stress. With further studies, the influence of the intermediate stress was appreciated therefore introducing true triaxial testing which is also used today. These tests are the bases of several failure criterions including Hoek-Brown failure criterion, Mogi 1971 etc. The results retrieved from laboratories are analysed with the aim of determining stresses resulting to rock failure, rock strength as well as rock mass parameters which are normally used in numerical modelling. It has been tradition to use least square regression as a modelling technique. This technique is easy to implement and interpret, however it is a non-robust technique and thus causes inaccuracies when predicting rock failure.

The inaccuracies caused by poor modelling have greater harm than what is expected. Poor modelling in rock engineering can result to the application of inappropriate numerical models, poor support design, rock failure, fatalities and many more. Hence this research focuses on evaluating different modelling techniques. It outlines the capabilities and limitations as well as factors affecting the predictive performance of each technique. True triaxial data of 11 rocks retrieved from scientific journals and books was used. The data included all three principal stresses (major, intermediate and minor) as well as the strength of each rock. In rock failure prediction, a failure criterion is normally chosen. For this research, Mogi's 3D failure criterion utilizing all three principal stresses was used. The laboratory data was modelled using each modelling technique.

The four modelling techniques were divided into two groups' i.e. statistical techniques (least square regression and quantile regression) and machine learning based techniques (logistic regression and k-nearest neighbour).

The statistical techniques made use of regression lines as failure envelopes hence they were strongly affected by indecisive points i.e. points plotting directly on the regression line and can neither be classified as stable nor failing. This had an effect on the performance and accuracy of the model. Furthermore, the least square regression was influenced by outliers as expected. On the other hand, the machine learning based techniques showed a very good predictive performance but were also affected by data pre-processing parameters. Logistic regression was influenced by the cutoff value and the pre-processing stages where the author had to decide on the variables to use since some of the variables were statistically non-significant. K-nearest neighbour was limited by the k-value (neighbourhood values) and the split (ratio between trained dataset and tested dataset). These parameters had an influence on the predictive performance of the models.

The performance of the four techniques were evaluated using different measuring techniques including the coefficient of determination (R^2) for statistical techniques, Receiver Operating Characteristic (ROC) curve for logistic regression and the confusion matrix for all four techniques. Research showed that R^2 is influenced by outliers and the x-axis range and does not give a true reflection of the model's performance. Confusion matrix however is not influenced by any factors hence the accuracies of the models were based on this parameter. K-nearest neighbour had the highest rock failure prediction accuracy followed by logistic regression, then least square regression and lastly quantile regression. Furthermore, the machine learning based techniques were good at identifying and classifying points resulting in failure as they had high sensitivity values while statistical techniques had high specificity meaning they were perfect at predicting stable points. In overall, it was a challenge implementing and interpreting machine learning based models however the rock failure prediction accuracy was higher compared to that of statistical techniques.

DEDICATIONS

This research is dedicated to my family “the Sambos”, my biological father (MJ Sambo) who is no more and the three trapped Lily mine workers (Their lights will remain shining).

ACKNOWLEDGEMENTS

Firstly, I would like to thank God almighty for giving me the strength and ability to complete this research. Without the knowledge granted by him, the research would not have been a success.

I am indebted to my supervisor Prof R. Mitra. His advises and support played a fundamental role to the success of this research. Special thanks are due to my computer scientists Vhahangwele Ramuada and Ivhani Maselesele as well as my statistician friend Tino Sibanda for assisting and clarifying all problems related to predictive models and algorithms. I would also like to express my gratitude to Mr T. Zvarivadza for his kind assistance.

Last but not least, I would like to send my appreciation to my mother (S.M Sambo), father (G.E. Sambo), my two siblings (Portia and Sakhi) and my amazing friends for their enormous support.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ROC	Receiver Operating Characteristic
CSIR	Council for Scientific and Industrial Research
MOSHAB	Mines Occupational Safety and Health Advisory Board
MATLAB	Matrix Laboratory

LIST OF SYMBOLS

σ_1	Maximum/major principal stress
σ_2	Intermediate principal stress
σ_3	Minimum/minor principal stress
σ_{oct}	Octahedral normal stress
τ_{oct}	Octahedral shear stress

CHAPTER 1: INTRODUCTION

This chapter outlines the background of the study, problem statement and research aims and objectives. These components are needed to provide context to the rest of this research.

The background focuses on the change in rock stresses within the vicinity of a mining excavation. It further looks at the significance of rock strength and its correlation to rock failure. Moreover, it provides insight into the true triaxial testing method which is used to determine the three principal stresses i.e. minor, intermediate and major principal stresses subjected to rocks. Lastly, it briefly highlights the various modelling techniques considered and used to model the data collected from true triaxial testing.

The factors discussed in the background are then combined to define the problem statement and research aims and objectives. The significance and layout of the research conclude this chapter.

1.1. Background

When a mining excavation is executed, the state of stress changes from primitive (virgin) to primitive and induced stress. Due to such stress changes, the stresses exposed to the rocks within the vicinity of the final excavation will be a combination of primitive and induced stress - *collectively known as resultant stress*. Such stress changes are critical and should be accounted for by rock engineers, as they are likely to affect the strength of the surrounding rock and lead to instabilities. These instabilities tend to result in rock failures which have dire consequences for mining operations.

Rock failure is considered a very crucial phenomenon in mining and any factors triggering this process are given enough attention. Moreover, a relationship exists between rock failure and rock strength; hence rock strength is considered a fundamental parameter when analysing the mechanical properties of rocks (Bieniawski, 1974).

Traditionally, the strength of rocks is determined in laboratories by conducting one of the two triaxial strength tests i.e. conventional triaxial testing or true triaxial testing. Conventional triaxial testing focuses on two stress directions (major and minor stress) and ignores the intermediate stress direction while the true triaxial testing accounts for all three stresses (Haimson, 2006).

In general, data collected from laboratory testing is analysed using a chosen failure criteria and various modelling techniques. From the test analysis, the principal stresses are determined and with the use of those results, rock strength can be deduced. This then assists in predicting the occurrence of failure and deciding on the type of support and support capacity when designing a mining excavation.

Various modelling techniques are available however; it has become common to use statistically based techniques such as least squares regression, quantile regression etc. These techniques are simple and easy to interpret. The drawback is that some of them including least square regression are strongly influenced by outliers hence, commonly referred to as non-robust and biased techniques. Alternative methods are machine learning based and they include: logistic regression, k-nearest neighbours, neural networks, Bayesian networks etc. These techniques are believed to provide more accurate and robust results compared to statistical based methods however, they require a high amount of data. Both types of techniques are rated based on their accuracy and interpretability. Generally speaking, modelling techniques that produces accurate result are often difficult to interpret.

For the purpose of this study, true triaxial data of 11 rocks is retrieved from various journals and books. This data consists of all three principal stresses i.e. major, intermediate and minor principal stress as well as the strength for each rock. The Mogi 1971 failure criterion is applied in calculating the mean effective normal and octahedral shear stress as it is one of the most important existing true triaxial failure criterions. The latter stresses are modelled using four modelling techniques: least square regression, quantile regression, logistic regression and k-nearest neighbour. On the completion of modelling, each model then predicts the number of stable points and points resulting to rock failure. The capabilities and limitations of the models are identified and the performances of the four models are evaluated using confusion matrix.

1.2. Problem statement

The inaccuracies caused by poor modelling have the ability to cause greater harm than what's expected. From a rock engineering perspective, poor modelling of the stresses affecting rocks can lead to the use of inappropriate numerical models, poor support design, rock failure, mine accidents, fatalities and an overall effect on the production of the mine. This clearly shows the importance of evaluating different modelling techniques in order to identify the influence

of their limitations and abilities on data modelling. It has become a tradition in almost all fields including mining to use least square regression as a modelling technique since it is simple and easy to interpret. This technique is exposed to outliers and classified as a non-robust technique. The limitations of this type of modelling technique are likely to influence the results produced resulting in inaccuracies when analysing and interpreting. This study then explores various predictive modelling techniques (statistical and machine learning based techniques) with the aim of evaluating their capabilities and limitations when predicting rock failure. The four modelling techniques of interest are least square regression, quantile regression, logistic regression and k-nearest neighbour.

1.3. Significance of the project

Rock strength is a critical parameter and of valuable interest to rock engineers as it plays a fundamental role in mine support design. According to Bieniawski (1974), being able to predict the strength behaviour of rocks creates a better understanding and clarification of the effects of different variables and a proper estimate on the design parameters. This study focuses on true triaxial data generated from past journals and books. The retrieved data comprises of the maximum, intermediate and minimum principal stresses as well as the strength of 11 rocks. These rocks are of different origin and type (sedimentary, igneous and metamorphic). Analysing the principal stresses and strength of the 11 rocks broadens our knowledge pertaining to rock properties specifically rock strength of each rock type. For example, the rocks of sedimentary origin have the least rock strength which is expected based on their origin. This is important in rock mechanics as it has an influence on numerical modelling of rock masses as well as the support design. A number of studies on rock failure prediction under true triaxial loading have been carried out by several rock engineers. However, prediction was based on least square regression. This study explores the capabilities and limitations of various modelling techniques originating from different fields: statistics (least square and quantile regression) and computer science (logistic regression and k-Nearest Neighbours). This will improve our modelling skills and the manner in which we treat and interpret data in our daily lives. With improved and better knowledge of modelling, better techniques can be used not only in rock engineering but also in other fields such as mine planning and optimisation or in plant and processing. In terms of data storage, the research will be of beneficial as a database for true triaxial testing will be created and will assist researchers in the future.

1.4. Research aims and objective

The capabilities and limitations of the four modelling techniques are quite fundamental in rock failure prediction. The aim of the research is to evaluate the four techniques by focusing on their capabilities and limitation as these have an influence on their predictive performance and accuracy on rock failure prediction. The popularity of the conventional testing method has led to it having vast information and database. True triaxial testing on the other hand is hardly carried out due to the scarcity of the machines to run this test thus less data is available. This research is also conducted to expand the true triaxial database by storing the data in a folder which people can easily have access to.

The objective of the research is:

- To model true triaxial data retrieved from journals and books, using Mogi's failure criterion whilst changing the modelling techniques. The four modelling techniques of interest are least square regression, quantile regression, logistic regression and k-nearest neighbour algorithm.

1.5. Research layout

The research is made up of five chapters. The first chapter outlines the introduction, problem statement, aims and objective and significance of the research. Within the introduction and problem statement, various aspects are highlighted i.e. the background of stresses around excavation, rock failure, the use of least square regression as a traditional modelling technique and the other modelling techniques.

Chapter two looks at work done previously by other authors. It explains the significance of rock strength, Mogi's 1971 failure criterion, rock strength testing methods and the background of the four modelling techniques (least square regression, quantile regression, logistic regression and k-nearest neighbour. It further highlights python programming language, robustness and biasness of the techniques and the parameters used to evaluate the accuracy of the techniques.

The third chapter outlines the methodology used throughout the research. This chapter explains the steps and procedures taken to ensure that the research's aims and objective are achieved. This includes collection of data from various scientific sources (journals and

books), sorting data in excel, calculating the stresses using Mogi's failure criterion and modelling using the four modelling techniques mentioned above.

Chapter four focuses on the analyses and interpretation of the results obtained from modelling the true triaxial data using the four modelling techniques. The last chapter concludes and summarises all the work carried out in the research and recommendations for future studies. The research layout is summarised in Fig 1.1.

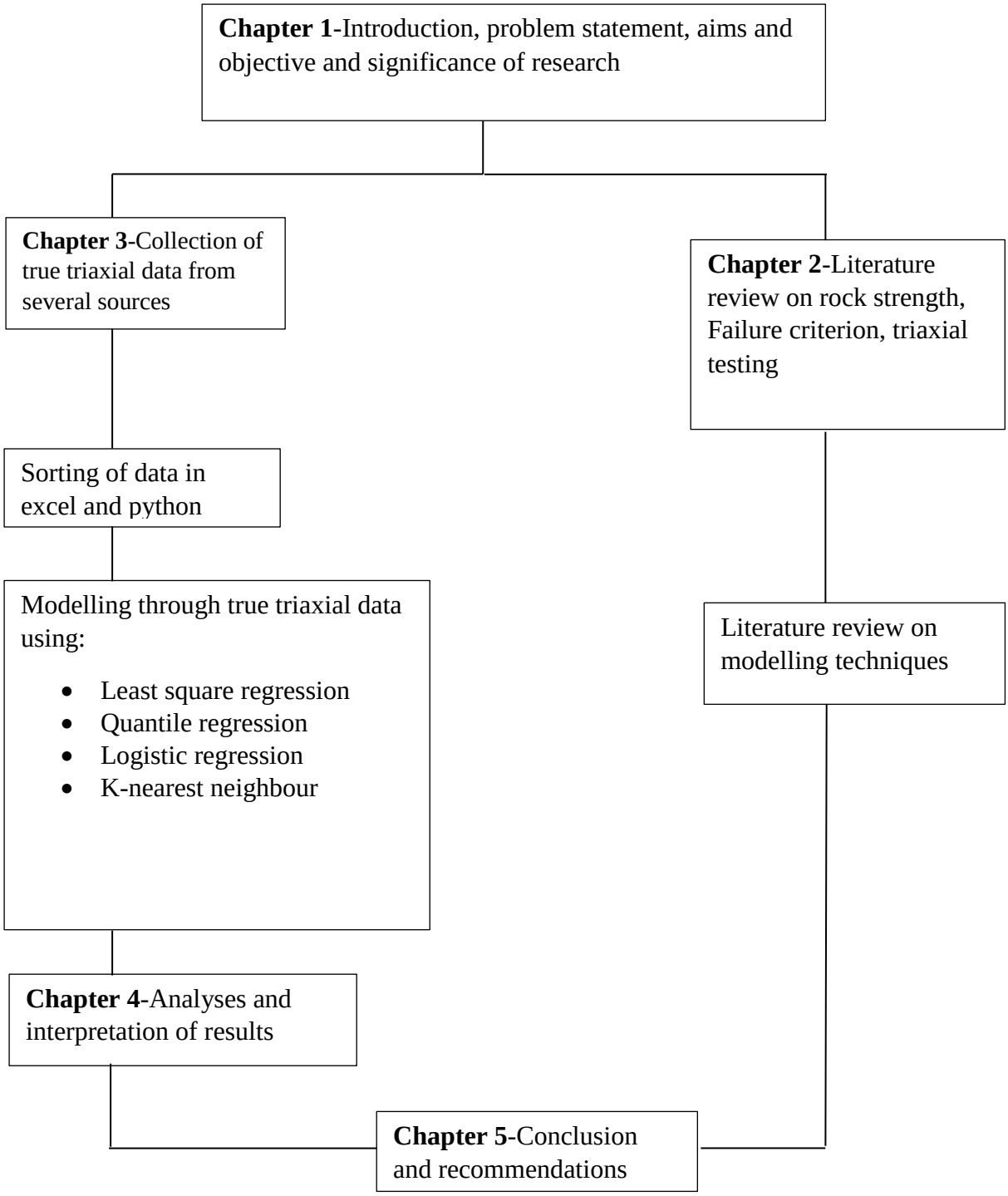

Figure 1.1: A diagram illustrating the research layout of the overall study

CHAPTER 2: LITERATURE REVIEW

This chapter reviews work done and published by several authors such as Bieniawski, Hoek and Brown, Haimson and Chang, Mogi and others. It highlights the significance of rock strength as well as understanding the stresses acting in the vicinity of an excavation in mining environments. It further discusses the two main rock strength testing methods i.e. conventional testing and true triaxial testing. Since rock failure is crucial for this research, it is also discussed in detail and within that section; Mogi's 1971 failure criterion is explained. Lastly, this chapter discusses the background and application of the four predictive modelling techniques i.e. least square regression, quantile regression, logistic regression and k-nearest neighbour as well as the relationship between accuracy and interpretation and biasness and variance of modelling techniques. The measures used to evaluate the performance of each model conclude this chapter.

2.1. Rock Strength

Rock strength is one of the mechanical properties of rock material and forms part of the parameters influencing the behaviour of rocks (Bieniawski, 1974). From a study of rock testing submitted to the rock mechanics division at the Council for Scientific and Industrial Research (CSIR), it was concluded that rock strength (both uniaxial compressive strength and triaxial strength) is the most critical parameter required by the mine and rock engineer. This parameter is crucial and plays an important role in mine and support design hence determining the strength of rocks before engaging in any mining activities is important. According to the Mines Occupational Safety and Health Advisory Board (MOSHAB , 1994), the strength of a rock is controlled by a combination of several factors including geological structures such as faults , joints (planes of weakness), groundwater, change of minerals as a result of exposure to water or air with time and intact compressive strength of the rock. It is also proposed that rock strength is mainly dependent on the direction of the load applied and the amount of load being applied i.e. the volume (MOSHAB, 1994). Based on Hoek et.al. (1980) rock strength is quite significant in mining activities such as drilling, blasting and excavation by machines. This mechanical property can be determined by either performing a conventional or true triaxial test.

2.2. Stresses around an excavation

Designing an excavation in rocks diverges from other structural designs because of the nature of the load acting on the system (Brady and Brown, 2004). According to Brady and Brown (2004), a structure designed underground is subjected to initial stress before excavation is carried out. As already stated in the introduction, this form of stress is known as virgin stress. When an excavation is made, the virgin stress of the rocks around the excavation changes (Fig: 2.1). This results to a distribution of stresses and formation of new stress concentration at different localities (Brady and Brown, 2004). The state of stress around the final excavation is then a result of the initial (virgin stress) and induced stress. This is collectively known as the resultant stress. Knowledge of the change in stresses is very fundamental as the stability of an excavation is predicted using these stresses (Brady and Brown 2004). According to Brady and Brown (2004), the local response of rock mass is dependent on the stress distribution component in the volume of the affected rock (Fig: 2.1).

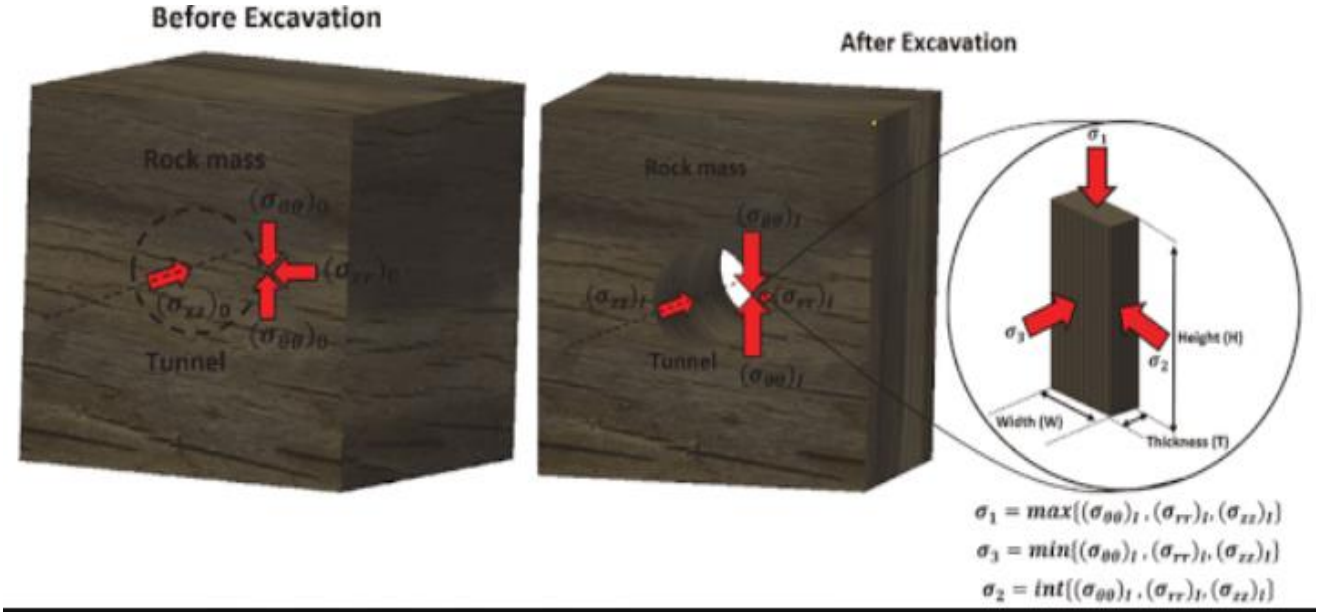

Figure 2.1: A diagram illustrating the different stresses acting before and after an excavation (Akdag, et.al. 2017)

2.3. The effect of intermediate principal stress (σ_2) on rocks

Several authors believe that the intermediate (σ_2) and minor (σ_3) principal stresses are equal hence σ_2 has been ignored for some time. However with information collected from the experiments conducted by Mogi in 1971 to 2006, it was concluded that the stress-strain

characteristics, deformation and failure modes are related to the intermediate and minimum principal stresses, respectively (Kwasniewski et al., 2012). An increase in yield point and coefficient of strain hardening is as a result of an increase in intermediate stress. Fracturing planes in rocks are parallel to the intermediate principal direction while the fracture angle decreases with an increase in the intermediate stress (Kwasniewski et al., 2012). According to Lee and Haimson (2011), the rock strength becomes lower when $\sigma_2=\sigma_3$ than when $\sigma_2>\sigma_3$. When the intermediate stress is greater than the minor stress, the true triaxial strength reaches levels of up to 60% higher than the conventional strength (Lee and Haimson, 2011). Furthermore, when $\sigma_2>\sigma_3$ then the maximum principal stress required to cause failure becomes higher than if $\sigma_2=\sigma_3$ (Adel, et.al. 2007). A limiting shear comparison between Mohr-Coulomb and Mogi's assumption was made. Mogi's theory suggested that the limiting shear is affected by not only the normal stress but also the intermediate principal stress hence the Mohr-Coulomb theory does not apply to true triaxial state stress (Kwasniewski et al., 2012).

2.4. Rock Failure

Under compressional conditions, failure occurs when the shear stress reaches and exceeds a value that is adequate enough to overcome the cohesion and frictional forces of the rock which are responsible for opposing motion (Adel, et.al. 2007). Chang and Haimson (2007) describe failure as a process that takes place when shear stress reaches a critical value which can be explained in terms of the mean effective normal stress.

2.4.1. Failure criterion

With the development of rock strength testing methods, various failure criterions were proposed. Some favoured the conventional method while some preferred the true triaxial method. Some well-known criterions include: Mohr-coulomb, Hoek-Brown, Modified wiebols and cook, Modified lade, Mogi 1967 and Mogi 1971 (Colmenares and Zoback, 2002). These failure criterions define failure differently hence different parameters are required. For example, according to Mohr-Coulomb, the shear strength or cohesion relates to the normal and shear stress during shear failure while Mogi 1971 failure criterion proposes failure as occurring when the energy reaches a critical value that is parallel to σ_2 stress (Colmenares and Zoback, 2002). From this example, it can be concluded that modelling the

same data whilst changing the failure criterions produces different results as illustrated in Figure 2.2.

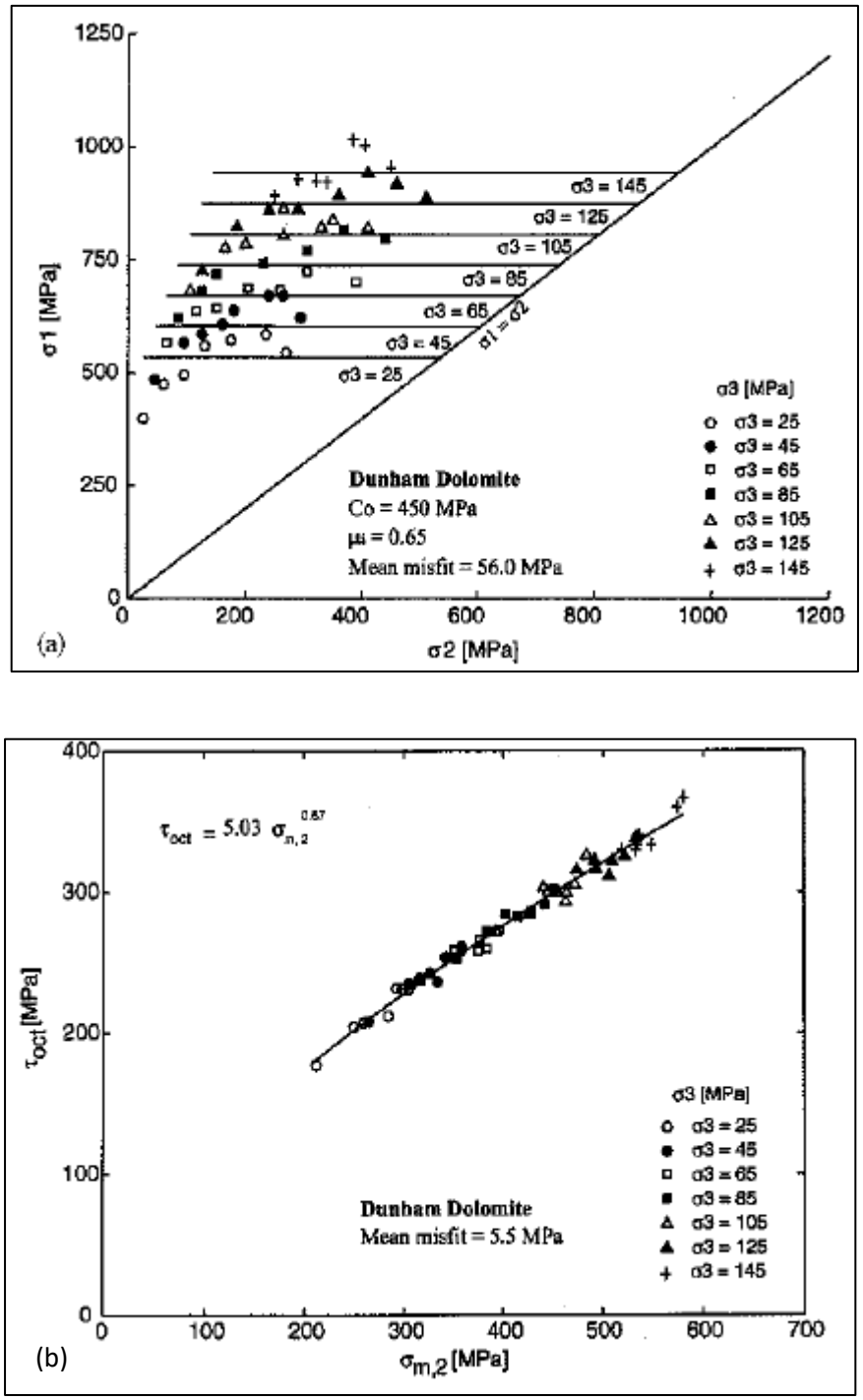

Figure 2.2.: A comparison between data modelled using (a) Mohr-coulomb vs (b) Mogi 1971 failure criterion (Colmenares and Zoback, 2002).

2.4.2. Brittle and ductile failure in rocks

Rocks are exposed to various physical and chemical conditions. Such conditions have an impact on the manner in which rocks fail. For example, low temperatures and confining

pressures are likely to result to brittle failure whereas high temperatures and confining pressures results to ductile failure (Hoek, 1968). Brittle failure is based mainly on the cleavage and intergranular fracture while ductile fracture focuses on the cavitation. In terms of plastic deformation, the bulk plastic deformation parameter allows cracks to initiate as long term stable crack before failure (Hoek, 1968). Based on the differences between the modes of failure, Mogi proposed two criterions i.e. ductile (yield) criterion and brittle criterion (Colmenares and Zoback, 2002). The intermediate stress (σ_2) was taken as the basic parameter when developing these two criterions. From Mogi's observations, brittle failure occurs when the strain energy reaches a certain value. This takes place in the σ_2 direction hence the effective mean is given by:

$$\sigma m,_2 = \frac{\sigma 1+\sigma 3}{2} \quad \text{.....1}$$

Where $\sigma_{m,2}$ is the effective mean acting parallel to σ_2, σ_1 is the maximum principal stress and σ_3 is the minimum principal stress. On the other hand, ductile failure is denoted by equation 2 with σ_2 being the intermediate principal stress. σ_{oct} is the octahedral normal stress with σ_2 as shown in equation 2 . This equation is based on the Nadai (1950) failure criterion.

$$\sigma_{oct}= \frac{\sigma 1+\sigma 2+\sigma 3}{3} \quad \text{.....2}$$

For this research, brittle failure will be taken into consideration therefore, equation 1 will be used.

2.4.3 Mogi 1971 failure criterion

While 2D criterions (criterions ignoring σ_2) became popular, some authors such as Mogi, Nadai and others were curious about the influence of the intermediate stress. This resulted to further conducting of tests on several rocks to check the effect of the intermediate stress. In 1967, Mogi developed a failure criterion which was applicable to brittle fracture and also accounted for σ_2 (Colmenares and Zoback, 2002). In 1971, he generalized the Von Mises theory and came up with an empirical formula that could be applicable to many rocks and give satisfactory results (Colmenares and Zoback, 2002). The criterion consists of the octahedral shear stress and mean effective normal stress (equation 4 and 5). Both stresses are calculated from the three principal stresses derived from true triaxial testing (Colmenares and Zoback, 2002). The octahedral shear stress gradually increases as a function of the mean

effective normal stress within the plane of failure as illustrated in equation 3 (Colmenares and Zoback, 2002).

$$\tau_{oct} = f(\sigma_1+\sigma_3) \quad \text{.....3}$$

F is a non-linear parameter determined experimental, σ_1 and σ_3 are the major and minor principal stresses and τ_{oct} is the octahedral shear stresses (Colmenares and Zoback, 2002). The two equations of Mogi's 1971 failure criteria are given by:

$$\sigma_{m,2} = \frac{1}{2}(\sigma_1+\sigma_3) \quad \text{.....4}$$

$$\tau_{oct} = \frac{1}{3}((\sigma_1-\sigma_2)^2+ (\sigma_2-\sigma_3)^2+ (\sigma_3-\sigma_1)^2)^{0.5} \quad \text{.....5}$$

2.5. Rock strength testing methods

2.5.1. Conventional triaxial testing

A conventional triaxial test is a trial in which the strength and deformability of a specimen are analysed under different confining stresses (Haimsom, 2006). The major (σ_1) and minor principal (σ_3) stresses are controlled while the intermediate stress is ignored. Under compression, it is assumed that the two stresses i.e. σ_2 and σ_3 are always equal. This test is quite popular and it is mainly conducted in laboratories. Furthermore, various failure criterions originated from this testing method including Mohr-Coulomb criterion, Hoek-Brown criterion etc. (Colmenares and Zoback, 2002). These failure criterions are classified as 2D criterions and do not account for the influence of the intermediate stress suggesting that the intermediate and minor stresses are equal (Haimsom, 2006). The conventional strength criterion is generally expressed as:

$$\sigma_1 = f(\sigma_3) \quad \text{.....6}$$

Where σ_1 is the major/ maximum principal stress, σ_3 is the minor/minimum principal stress and f is a parameter dependent on the rock type and it is determined experimental (You, 2010).

2.5.2. True triaxial testing

The conventional triaxial testing gained major popularity however some authors were curious about the influence of the intermediate stress. Authors like Nadai, Mogi and others made an effort to incorporate the effect of the intermediate stress on rock strength. In 1971, Mogi designed a true triaxial testing machine which was aimed at determining the true triaxial strength of rocks (Haimson, 2006). According to Li, et.al. (2012), Mogi's true triaxial testing machine was the first to allow the application of all three principal stresses independently. With the assistance of the machine, it was discovered that the major principal stress at failure is linked to the intermediate principal stress under constant minimum principal stress (Li et.al. 2012). Based on Kwasniewski et al., (2012) experimental work, the true triaxial testing in laboratories is conducted to determine the deformation characteristics and strength behaviour of a specific rock at different in-situ stresses. According to Haimson, (2006), the true triaxial testing machine has the ability to recreate hydraulic pressure and stress environment. In a true triaxial test, all three principal stresses are taken into account i.e. the effect of the intermediate stress is also taken into consideration. Similar to the conventional triaxial testing, several criterions have been developed from data retrieved from true triaxial tests. These are grouped into polyaxial and empirical criterions. Polyaixal criterions include: modified Wiebols and Cook, modified Lade and Drucker–Prager while examples of empirical criterions are Mogi 1967 and Mogi 1971 (Colmenares and Zoback, 2002). According to Colmenares and Zoback (2002), performing polyaxial tests is quite difficult hence true triaxial tests are more preferable.

2.6. Modelling techniques

Modelling is defined as a useful tool in the development and testing of various theories by means of explaining, describing and predicting (Shmueli, 2010). Based on the definition, three types of models are generally used i.e. explanatory, descriptive and predictive modelling. Shmueli (2010) defines explanatory modelling as "An application of statistical models to data for testing casual hypothesis about theoretical constructs". In this type of modelling, specific parameters are measured by independent variables and are expected to have an effect on the dependent variables (Shmueli, 2010). An example of explanatory modelling is regression. He further defines predictive modelling as applying data mining algorithm with the aim of predicting new or future observations. Examples of predictive

modelling are Bayesian, k-nearest neighbour etc. The last model i.e. descriptive modelling, is considered as the most commonly used modelling technique. This model summarises the data structure in a concise way. As opposed to explanatory modelling, the focus of descriptive modelling is aimed at measurable level and not the construct level (Shmueli, 2010). Regression can also be considered as a descriptive model if it is utilized in determining the relationship between the dependent and independent variables rather than being used for prediction (Shmueli, 2010). In general, the modelling process involves various steps which are listed in Fig: 2.3 (Shmueli, 2010).

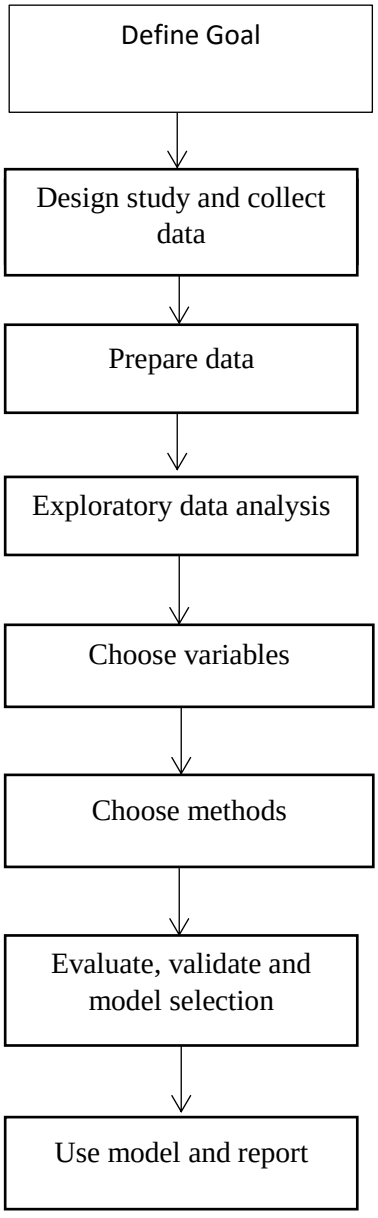

Figure 2.3: The steps involved in the modelling process (Shmueli, 2010)

2.6.1. Statistical techniques

Conditional mean modelling

Regression modelling techniques focuses on the mean of the data hence a function called conditional mean is used to model the data. This group of models are collectively known as "conditional mean models" and are easy to use and interpret (Hao and Naiman, 2007). They have been widely applied in various fields including mining. Despite their popularity, they do not extend to locations that are not within the centre and are strongly affected by outliers. Due to the above mentioned restriction, these models are likely to produce a misleading measure of centre location (Hao and Naiman, 2007). An example of a conditional mean model is the least square regression.

2.6.1.1. Least square regression

Regression analysis is defined as a parametric statistical technique used in investigating and determining the relationship between variables (Montgomery, et.al, 2012). This analysis is conducted to display the relationship between a response variable and predictor variable (O'Neill, 1999). A number of regression techniques exist including least square regression, quantile regression, logistic regression etc. Least square regression is defined as a method that generates a line which is a representation of the real function of data points (O'Neill, 1999). The line is referred to as the "best fit" line and it is given by the following equation:

$$Y=\beta_o+\beta_{1x} \quad \ldots\ldots 7$$

This equation was further modified to equation 8 in order to accommodate for data points that do not fall directly on the straight line.

$$Y=\beta_o+\beta_{1x}+\varepsilon \quad \ldots\ldots 8$$

According to O'Neill (1999), the objective of the least square regression method in data points is to reduce the square of the error distance. Furthermore, this method has the ability to determine trend in data with its accuracy being dependent on the input data. The least square regression method assumes that a near linear data set will be produced as a result of transformation of the data (O'Neill, 1999). This suggests that if data does not follow the above mentioned assumption, then the method will fail to establish the relationship between the variables. As stated already, a best fit line is generated and it is defined based on its

function e.g. in cases where measurements are done on the X coordinates only, the best fit line will be defined as the line that reduces the Y residual squares (York, 1966). If the errors are within the Y coordinates then the best fit line will be explained as the line that minimises the X residual squares (York, 1966). In practical, error occurs in both variables hence two best lines are fitted and the mean is then calculated (York, 1966).

The least square regression method might have gained popularity however it is affected by various factors. Some of these factors are: the arrangement of the x values as they affect the slope and the outliers which have an influence on the overall modelling of data (Montgomery, et.al. 2012). Figure 2.4 is an illustration of a scatter plot modelled using least square regression. In general, when prediction is carried out using least square regression, the size of the error becomes the parameter of interest. This is analysed relative to the true or observed value and called percentage error.

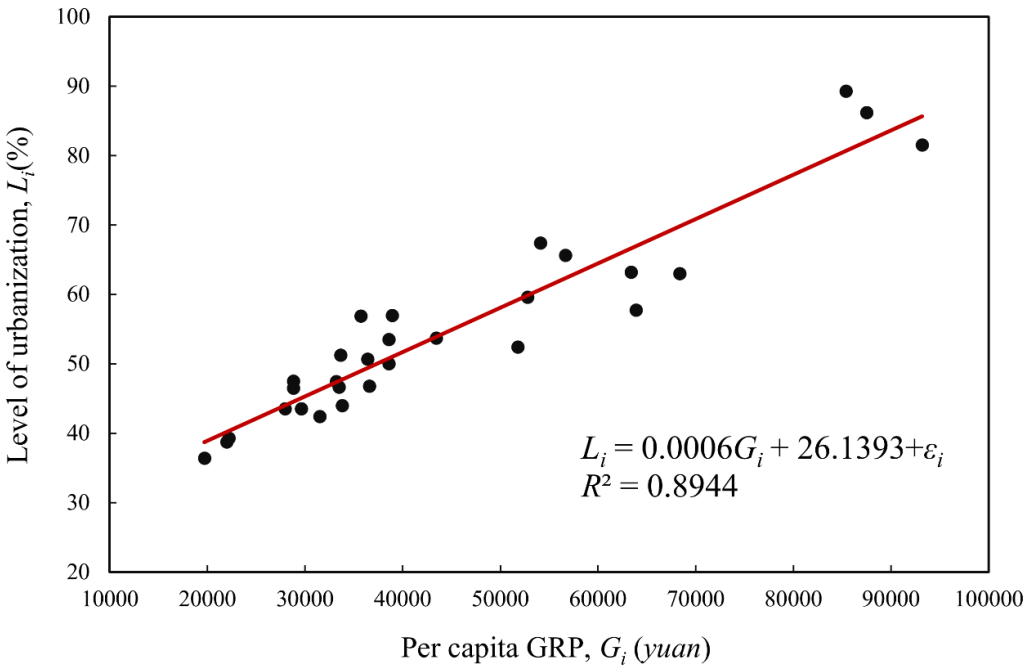

Figure 2.4: An example of data modelled using least square regression (Chen, 2016).

2.6.1.2 Quantile regression

Quantile regression was introduced in 1978 by Koenker and Bassett as a non-parametric tool that models conditional quantiles as function of predators (Koenker and Bassett, 1978). Quantile is a term that generalises other terms such as quartiles, quintiles, decile and percentile. This form of regression is regarded as an extension of the linear regression model with the difference being the adjustments in the conditional mean of the variables (Hao and

Naiman, 2007). In linear regression, the adjustments are more focused on the conditional mean, while in quantile regression the changes in conditional quantiles are specified (Hao and Naiman, 2007). With quantile regression, it is possible to model any position of the distribution since any quantile can be used. Multiple quantiles (Fig, 2.5) can be modelled suggesting that the effect of predictors to the response distribution is understood clearly (Hao and Naiman, 2007). The shape of the conditional distribution can be characterised by the central location and equally spaced conditional quantiles (Hao and Naiman, 2007). A quantile regression equation equivalent to equation 8 of the least square regression is given by:

$$Y_1 = \beta_o^P + \beta_1^P x + \mathcal{E}_1^P \quad \text{.........} \quad 9$$

Where 0<p<1 is explained as the proportion of the population with scores lying below the quantile p. The conditional p^{th} quantile is described by β_o^P and β_1^P.

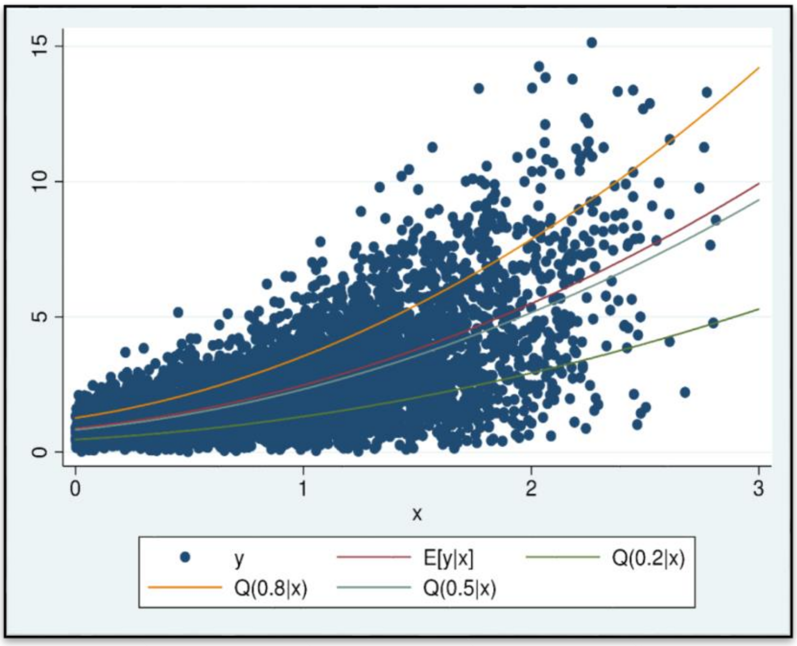

Figure 2.5: A graph illustrating modelling using quantile regression (Drukker, 2016).

2.6.2. Machine learning based techniques

Machine learning is defined as an intrinsic part of artificial intelligence that enables the computer to learn data directly without using any assumed equations or models (Escober and Morales-Merendez, 2018). Various machine learning methods exist however the two popular

methods are supervised and unsupervised machine learning. According to Brownlee (2016), supervised learning uses labelled data to train the algorithm i.e. the desired output for an input is already known. Unsupervised learning works with unlabelled data i.e. the corresponding outputs of a specific input are unknown (Brownlee, 2016). Logistic regression is a supervised machine learning technique whilst k-nearest neighbour is an example of unsupervised learning machine.

2.6.2.1. Logistic regression

Logistic regression is explained as a parametric technique used to model the impact of independent variables on a binary dependent variable focusing on the probability of it occurring in one of its two classifications (Fleiss, et.al., 1986). Logistic regression can be either classified as a statistical or machine learning technique. According to Austin and Steyerberg (2012), it is generally fitted for 3 reasons i) to determine the relationship between 2 variables ii) to determine the independent variables of an outcome iii) to estimate or predict the probability of an event occurring. Logistic regression deals with probabilities that are between 0 and 1 (Fleiss, et.al., 1986). In general, one represents the occurrence of an event and zero represents the absence (Pojlman and Leitner, 2003). In reference to this study research, logistic regression results have two possible outcomes i.e. stable or fail. This suggests that it is possible to predict if a certain point will lead to failure or not given the true triaxial data and rock strength.

The logic behind logistic regression is that the dependent variable only works with values between 0 and 1 while the predicted values focuses on the mean proportion of the independent variables (Fleiss, et.al., 1986). As much as the probability is represented by values between 0 and 1, the linear line can extend towards the negative as a result of a decrease in the independent values. On contrary, an increase in the same values can result to the line extending above 1. This suggests that one would end up with probability values of less than 0 or above 1 (Fleiss, et.al., 1986). This then violates the logic that the probability is represented by values between 0 and 1. Logistic regression is also known to be influenced by pre-processing of the data. This includes the cut-off value and variables chosen to be statistically significant. From an economic description, the cut-off value is defined as the point in which decision markers decide on whether to approve a loan application or not (Soureshanji and Kimiagari, 2012). In the context of this research, a cut-off value would be a point in which the model decides on whether to classify the points as failing or stable.

2.6.2.2. K-nearest neighbour

K-nearest neighbour is defined as a non- parametric machine learning method with strong statistical properties (Eskandarinia, et.al, 2010). It is characterised as flexible and capable of adapting to irregular feature space. It is mainly used for classification but can also be used in regression problems. K-nearest neighbour works by predicting and classifying an outcome of an unknown point by using the outcomes of the neighbours in the training set (Alkasassbeh et.al, 2015). Generally, it favours the majority over the minority class thus classifying according to the majority class. Figure 2.6 explains the principle under which k-nearest neighbour functions. From this figure, 7 neighbouring points were used to predict the unknown point hence k=7. Majority of the class are class 2 thus the point is predicted to be of class 2. The k value is considered as a fundamental parameter in k-nearest neighbour since it determines the number of neighbouring points to use when predicting and classifying the unknown point. Another important parameter in k-nearest neighbour is the distance between the known points and the unknown point. This distance can be calculated using various distance metrics including Euclidean distance, cosine, chi square, minkowsky etc. (Hu, et.al, 2016). The last important parameter in k-nearest neighbour is the pre-processing method. This includes the split value used to differentiate between training set and testing set data. The split has an influence on the accuracy of the model. In total, 3 parameters mainly influence the performance of k-nearest neighbour i.e. k value, distance metric and split value (ratio between training set and testing set (Min Ma, et.al., 2014).

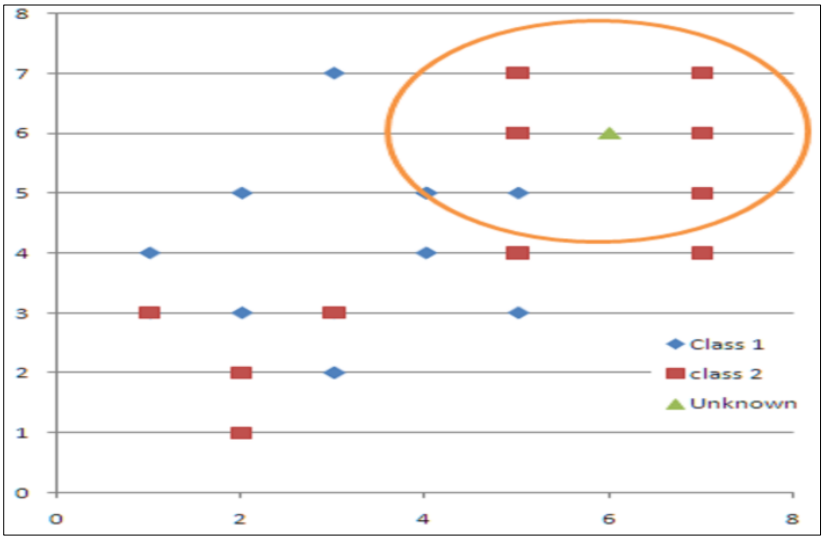

Figure 2.6: An illustration of the principle behind k-nearest neighbour (Alkasassbeh, et.al. 2015)

According to Oneata (2011), this technique does not require any prior-knowledge but works with the data currently present. It also has the ability to model complex functions possessing non-linear decision boundary by utilising local assumptions (Oneata, 2011). The drawback of this method is its exposure to the ad hoc choice of same metric when dealing with heterogeneous data (Yao and Ruzzo, 2004). Another disadvantage of the k-nearest neighbour is its dependence on the geometric distance in order to determine the resemblance and variation between objects without taking any statistical consistency of the data (Abedallah and Shimshoni, 2012).

As stated already, several distances are used in order to reach the final output. For this research, the Euclidean distance will be used. It is defined as the distance between 2 points in space and computed by determining the root of square differences between two points (Mulak and Talhar, 2013). This distance is given by equation 10:

$$\text{Dist. (A, B)} = \sqrt{\frac{\sum_{i=1}^{m}(x_1 - y_1)^2}{m}} \quad \text{.....10}$$

Where A and B are the vectors (A= x_1, $x_2 \ldots x_m$ and B=y_1, $y_{2 \ldots} y_m$) and m is a dimensionality of feature space.

2.7. Algorithm

2.7.1. Programming language

According to Zelle (2004), a program is a set of instructions that commands the computer on what to do. A programming language is then defined as a code used in writing instructions that will be followed by the computer (Zelle, 2004). Three main groups of programming languages are mainly used i.e. procedural, declarative and object-oriented. Before 1990's programming languages were regarded as procedural. This implied that the program uses a set of instructions provided by the programmer (Akbligic and Akinci, 2009). The next group of language to be introduced was the declarative language. This language is not affected by the manner in which instructions are ordered. Lastly, the object-oriented programming language was introduced. An object oriented language allows the user to split the problem into objects. Several programming languages are classified into these three groups. For example, Matrix Laboratory (MATLAB) is classified as a procedural language while python can be both procedural and object-oriented.

2.7.2. Python programming

Python is one popular program used worldwide. This program consists of a sequence of statements for which the python interpreter can execute. It is a dynamically typed language (Zelle, 2004). Two types of languages are associated with programming i.e. low level language (machine language) and high level language (human language). Two methods are also available for the translation of human language to machine language. These methods are compiling methods and interpretation methods. A compiler uses high level language of another program and translates it into a program of the same level as the machine language while an interpreter stimulates a computer that has knowledge of the high level language (Zelle, 2004).

2.8. Robustness

The validity of estimation method is based on assumptions and these assumptions are considered to be fundamental in statistics (Fox and Weisberg, 2011). This leads to the introduction of robustness. According to Buhai (2004) robustness is crucial throughout statistics. An estimate is said to be robust if useful information is produced regardless of some assumptions being applicable or not (Fox and Weisberg (2011). The effectiveness of a modelling technique can also be evaluated using robustness. For example, least square regression behaves badly particular when the distribution is not normal suggesting that it is non-robust. On the other hand, quantile and logistic regression estimates are robust to impureness of the response observation however sensitive to design observation contamination (Buhai, 2004). Two approaches are used to handle non-robust estimates i.e. by removing influential observatory or using a robust regression (a criterion not affected by outliers) (Buhai, 2004).

2.9. Accuracy vs interpretability

The performance of a model is explained based on its accuracy in predicting event occurrences (Plate, 1999). Some techniques can perfectly model data suggesting an accurate prediction of events while some are poor modellers resulting to an incorrect prediction of events. The major drawback with accurately modelling techniques is their difficulty in interpretation (Plate, 1999). This suggests that an inverse relationship exists between accuracy and interpretations i.e. modelling techniques with high accuracy are difficult to

interpret while those with poor accuracy are quite easy to interpret (Plate, 1999). It can be concluded that the factor of interest should be highlighted at the initial stages of a project i.e. it should be decided on whether accuracy or interpretability is more significant for a specific problem as this plays a role in the process of selecting a model.

2.10. Bias vs variance trade off

When analysing models, prediction error is mainly the parameter of focus. Within this error, two components are of interest i.e. errors as a result of bias and errors caused by variance. These two errors are important in evaluating models and avoiding over and under fitting of models (Scott, 2012). According to Scott (2012), bias and variance can be explained in three ways i.e. conceptual, graphical and mathematically. From a conceptual perspective, bias errors result from the difference between predicted values and true values while errors due to variance are explained as the differences of prediction models in a given data point (Scott, 2012). The differences causing bias errors are a result of assumptions made by the model while variance errors are caused by the specifics made by the training dataset. In a graphical definition, a bull-eye diagram is used to explain bias and variance. Figure 2.7 is an example of a graphical definition of bias and variance. In this figure, a combination of high and low bias and variance is explained using the centre of target as a model that predicts correct values. According to Scott (2012), moving away from the centre suggests poor prediction and points closer to the centre are a reflection of good prediction.

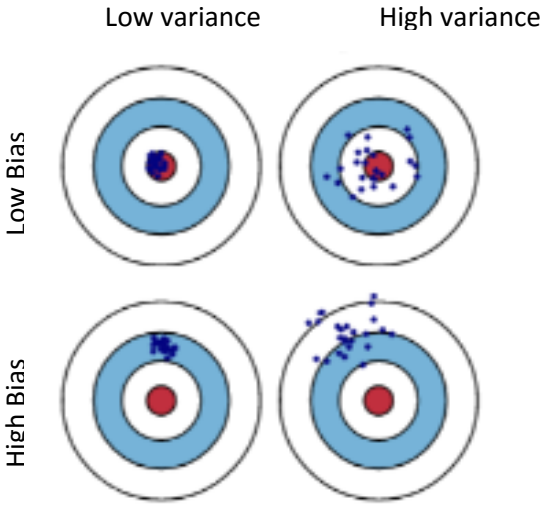

Figure 2.7: An example of a graphical illustration of bias and variance (Scott, 2012)

Lastly, bias and variance can be explained mathematically as an expected squared prediction error at a point X and this is given by equation:

$Err(x) = E[(Y - \hat{f}(x))^2]$ $Err(x)$

$= E[(Y - \hat{f}(x))^2]$ 11

This can be broken down into bias and variance and given as:

$Err(x) = bias^2 + variance^2 +$ irreducible error.....12

Irreducible errors are those that cannot be reduced by modelling hence they are normally ignored (Scott, 2012).

According to Scott (2012), when the variance increases, the biasness decrease and this is how the bias-variance tradeoff works (Fig: 2.8). Managing bias and variance is equivalent to over and under-fitting. Based on Scott (2012), variance is increased while bias is decreased relative to the model complexity (Fig: 2.8). With an increase in parameters, the model becomes more complex and the variance becomes the parameter of interest while the bias drops (Scott, 2012). Scott (2012) further explains bias as having a negative first order derivative while variance has a positive slope as seen in Figure 2.8.

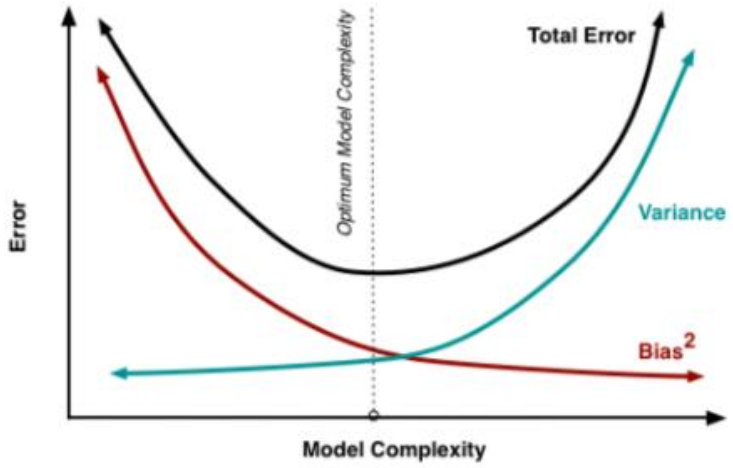

Figure 2.8: A graph showing the bias-variance trade-off (Scott, 2012)

2.11. Evaluation of model performance

2.11.1. Coefficient of determination (R^2)

In regression modelling, the coefficient of determination (R^2) is commonly used as a measure for goodness of fit. According to Hagquist and Stenbeck (1998), R^2 focuses on the amount of reduction of the prediction error relative to the potential error present. It measures the relationship between the model and the observed data and it is considered as a sample specific measure just like other correlation coefficients (Hagquist and Stenbeck, 1998). Its value differs between various samples as a result of changes in the variance between different samples. In R^2, a baseline is provided such that a 1 represents a perfect fit while a 0 is a no fit (Hagquist and Stenbeck, 1998). A downfall of using R^2 is that it increases for each increased variable added to the model. The R^2 measure is given by the following equation:

$$R^2 = \frac{\sum(Y_1 - Y_2)^2}{\sum(Y - Y_2)^2} \quad \text{..... 13}$$

Where Y_1 is the observed, Y_2 is the predicted and Y is the mean value.

2.11.2. Receiver-Operating Characteristic (ROC) curve

When the performance of logistic regression is evaluated, 2 parameters are assessed i.e. the calibration and discrimination. Calibration focuses on the agreement amongst the observed and predicted values whilst discrimination is based on the ability of the model to differentiate between the two outcomes (stable and failure) (Austin and Steyerberg, 2012). Calibration is assessed using a c-statistic parameter meanwhile a ROC curve is used for discrimination (Austin and Steyerberg, 2012).

According to Florkowski (2008), the ROC curve functions as a trade between sensitivity/true positive rate (points failing and predicted to fail) and specificity/false positive rate (points stable and predicted to be stable). The closer the curve is to the left of the graph, the more accurate the test is as observed in Figure 2.9. The area under the ROC curve is known as the AUC and also used as a measuring tool. In 2001, Tape evaluated and classified the performance of a logistic regression test using the AUC as follows:

0.5-0.6- Fail

0.6-0.7- Poor

0.7-0.8- Fair

0.8-0.9- Good

0.9-1- Excellent

This suggests that a perfect test has an area of 1 while 0.5 represents a worthless test. If the AUC is less than 0.5, then the test cannot be used for prediction. Some authors refer to such test as a guess work.

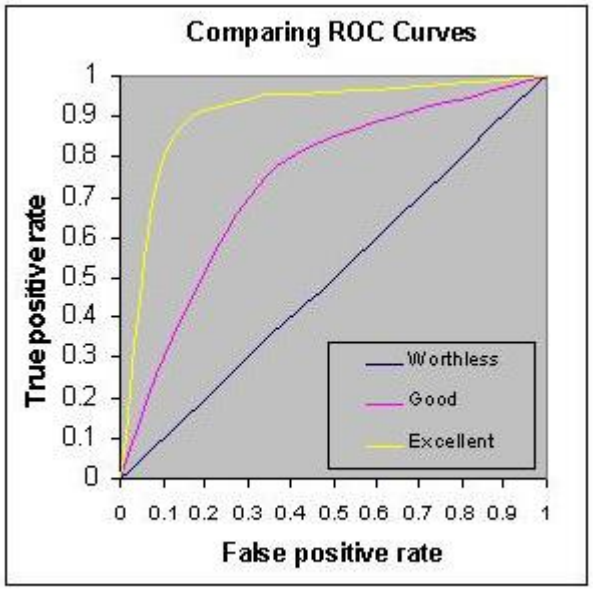

Figure 2.9: A graph showing different examples of ROC curves (Tape, 2001)

From the logistic regression results, the coeffb, se, p-value and exp (b) are also given. The coeffb and exp (b) evaluates the effect of changing the predictor on the final outcome. A value closer to zero suggests a small effect (Tape, 2001). The SE coeff (standard error) is used to explain the variation obtained when the tests are conducted over and over again using the same population. The smaller the standard error, the more precise the estimate is. The p-value is used to assess whether a specific variable is statistically significant or not. When a p value of a variable is less than the significance level, it is suggested that the variable is statistically significance (contributes to the event occurrence) while p greater than the significance level is not statistically significant (Ranganathan, et.al. 2017).

2.11.3. Confusion matrix

A confusion matrix is a system used to evaluate the performance of a model technique. It is made up of the actual and predicted data and uses a table (Table 2.1) as an evaluating measure (Visa et.al. 2011). In Table 2.1, four alphabets are used as numbers representing positive and negative values which are either correctly or incorrectly classified. Number "a" represents true negatives (points predicted correctly as failing) while "d" represents true positives (stable points predicted correctly). Number "c" represents false negatives (stable points but incorrectly classified as failing) and "b" represents false positive (failing points misclassified as stable) (Visa, et.al. 2011). According to Parikh, et.al. (2008), a number of terms are derived from the confusion matrix. This includes the accuracy, sensitivity, specificity, error rate and precision. This research will focus on accuracy, sensitivity and specificity.

Accuracy is described as the accuracy of the correctly classified percentage and based on Table 1, it is given by:

$$\text{Accuracy}=\frac{a+d}{a+b+c+d} \quad \text{.....14}$$

$$=\frac{Number\ of\ correctly\ classified}{total\ number\ of\ cases}$$

A good model should have fewer b (misclassified as positive) and c (misclassified as negative) points. The higher the misclassification of points, the lower the accuracy and performance of the model.

Table 2.1: An example of a confusion matrix for a classification problem with two classes (Visa, et.al. 2011)

	Predicted Negative	Predicted positive
Actual negative	a	B
Actual positive	c	D

Sensitivity and specificity are two critical parameters when evaluating the ability of a model to correctly classify points. They are mainly used in clinical tests where patients are tested for

a certain disease and classified either as positive or negative (Lalkhen and McCluskey, 2008). Sensitivity focuses on the ability of the test to correctly identify an event occurrence. In the context of this study, sensitivity would refer to the ability of the model to correctly classify failing points as failing. A test with 100% sensitivity implies that all failing points are classified as failure while 80% sensitivity suggests that 80% of the failing points can be classified as failing while 20% are classified as stable. On the other hand, specificity refers to the ability of the test/model to correctly classify an event not occurring. For this research, specificity focuses on the ability of the model to correctly classify stable points as stable. According to Lalkhen and McCluskey (2008), a higher sensitivity is important when the model is used to identify and classify serious phenomenon which requires immediate solving. Since rock failure is critical in mining and requires immediate attention, the sensitivity for all models should be higher than the specificity.

Both sensitivity and specificity are influenced by various parameters including the cut-off value, k value, split etc. Sensitivity and specificity are given by the following equations:

$$Sensitivity = \frac{Number\ of\ true\ positives}{Number\ of\ true\ positives + Number\ of\ false\ negatives} \quad 15$$

= Probability of a predicting a failing point provided that the point is failing

$$Specificity = \frac{Number\ of\ true\ negatives}{Number\ of\ true\ negatives + Number\ of\ false\ positives} \quad 16$$

=Probability of predicting a stable point provided that the point is stable

2.12 Conclusions

In order to thoroughly understand and achieve the aims and objective of the research, extensive literature review was carried out on important concepts related to the study. From views of several authors, it was concluded that rock strength is a fundamental mechanical property of rocks. This rock property is affected by external factors such as stresses acting in the vicinity of excavations in rocks. The stresses comprises of three principal stresses i.e. major, intermediate and minor stress. It has been questionable for some time on whether the intermediate stress has an influence on rock failure hence the influence of the intermediate stress was reviewed in this chapter. A conclusion was drawn that the intermediate stress does have an influence on rock failure. Rock failure and failure criterion are also important

components in rock engineering and governed by various factors including the strength of rocks. Two rock strength testing methods are generally used i.e. conventional triaxial testing and true triaxial testing. The conventional triaxial testing ignores the intermediate stress whilst true triaxial testing accounts for the intermediate stress. Results retrieved from rock strength testing are modelled using a specific failure criterion and modelling technique (Mogi 1971 in the context of this research). It has become standard to use statistical techniques (least square regression) when modelling rock strength testing results. Various modelling techniques i.e. quantile regression and machine learning based techniques (logistic regression and k-nearest neighbour) have been reviewed. It was noted that these techniques have abilities and limitations that influences the predictive performance as well as accuracy of the techniques. For example, least square regression is highly affected by outliers while k-nearest neighbour is influenced by the k value and pre-processing (split). The ability of each technique in predicting failure is evaluated using several parameters and systems. They are collectively evaluated using a confusion matrix system.

CHAPTER 3: METHODOLOGY

This chapter outlines the various steps taken in order to achieve the aims and objectives of this research. It discusses the procedures carried out from the initial stages i.e. data collection right through to the last stage which is analysing and interpretation of results. Processes such as data sorting, visualization and modelling are also discussed in detail.

3.1. Data collection

Several authors have conducted experimental tests with the aim of determining the strength of different rocks. A number of authors have conducted conventional tests while some have opted for true triaxial tests. The results from those tests were expressed as principal stresses. With the use of the conventional testing method, the major and minor principal stresses were the end product whilst in true triaxial testing, the results comprises of all three principal stresses. As stated already, the research focuses on true triaxial testing implying that all three stresses are taken into account. To fulfil the aim and objectives of the research, the data used for predictive modelling was retrieved from two published sources i.e. scientific journals and books. Another important parameter was the true triaxial strength of each rock and was obtained from the true triaxial book by Kwasniewski, et.al. (2017). Data from eleven rocks that have undergone true triaxial testing was used for this study (appendix A, Table 1). These data is an end product from experiments conducted by:

- Mogi (1971) -Dunham dolomite, solnhofen limestone, yamaguchi marble, mizuho trachyte, orikabe monzonite, inada granite, monazuru andesite
- Haimson and Chang (2000, 2002)-KTP amphibolite and westerly granite
- Takahashi and Koide (1989)- Shirahama sandstone and Yuubari shale

3.2. Sorting and visualizing

The results retrieved from journals and books were sorted in excel in the form of σ_1, σ_2 and σ_3. Since eleven rocks were used, eleven tables were produced each having four columns (three columns for the principal stresses and one for the true triaxial strength of the rock) (Table 3.1 generated from Appendix A, Table 1). The Mogi 1971 failure criterion was used to calculate the mean effective normal stresses and octahedral shear stresses using equation 4 and 5 as stated in the literature review. On completion of sorting the data, a scatter plot was

then used to visualize the relationship between the mean effective normal stress and octahedral shear for statistical techniques. Meanwhile, machine learning techniques modelled data directly from the table i.e. a graphical plot was not required.

Table 3.1: An illustration of the sorting of data (Solnhofen limestone) retrieved from journals and books (Mogi, 1971).

σ_1	σ_2	σ_3	Rock Strength
310	0	20	310
397	20	20	310
417	51	20	310
413	92	20	310
453	165	20	310
460	206	20	310
465	233	20	310
449	40	40	310
446	40	40	310
46	80	40	310
499	113	40	310
530	193	40	310
547	274	40	310
535	315	40	310
473	60	60	310
517	87	60	310
537	102	60	310
530	113	60	310

3.3. Modelling

The pre-processing stage in the modelling of data varies per model. For this research, four modelling techniques were used. These are: the least square regression, quantile regression, logistic regression and k-nearest neighbour. The idea behind modelling is to determine the relationship between the explanatory variable and response variables with the aim of using that relationship in the prediction of rock failure. The two regression techniques were constructed in excel using XLSTAT tool. For least square regression, a line of best fit was fitted in the data while quantile regression made use of three quantiles (0.25, 0.5 and 0.75). For logistic regression, IPython was used to define the data (Appendix B, Table 2). When the data was well defined, real statistics tool in excel was used to run the model. For the last

model (k-nearest neighbour), coding was constructed in python. An example is shown in Figure 3.1. The overall algorithm is shown in Appendix C (Fig.3)

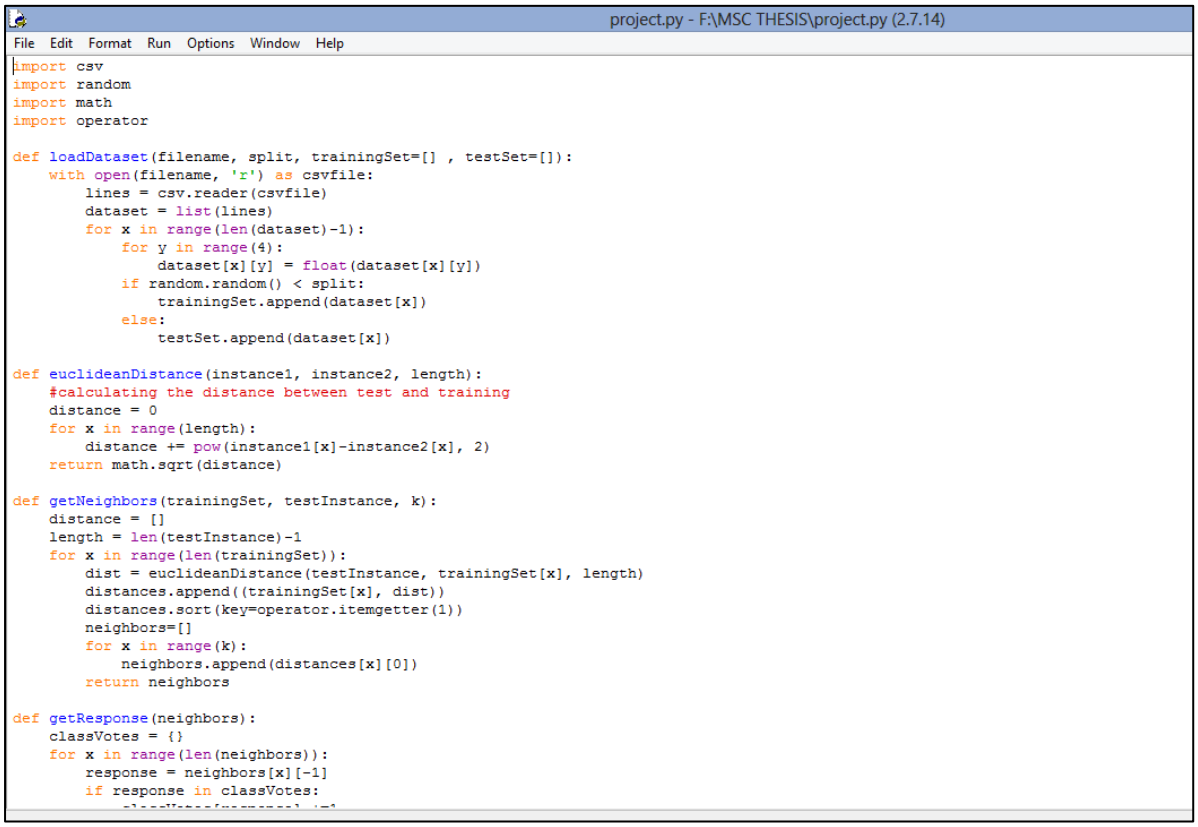

Figure 3.1: An example of a coding in progress using python language

3.4. Analysis

When modelling was complete, the data was then analysed against each modelling technique. The points that would result to rock failure were identified in each individual rock. The predictive performance of each technique was evaluated. Various techniques use different parameters to evaluate the performance of a model. For regression techniques, a measure called R^2 was used as a measure tool. In logistic regression, an ROC curve was used in assessing its performance. To determine the accuracy of each model, a confusion matrix table was constructed in which two important parameters were also determined i.e. the sensitivity and specificity of each model. In the confusion matrix, the points were classified as observed and predicted to be stable, observed and predicted to fail, observed to fail but predicted to be stable and observed to be stable but predicted to fail. The observed is the state of the point in reality whilst the predicted is the estimation made by the predictive model.

3.5 Conclusions

The dataset used for modelling was collected from scientific journals and books. During collection of the dataset, it was noted that not a lot of true triaxial database is available for public access as compared to conventional triaxial database. Hence this research will add on the information available for public use. The retrieved data was modelled using all 4 modelling techniques to predict rock failure. The statistical techniques were based on statistical parameters whilst machine learning based techniques utilized algorithms.

CHAPTER 4: RESULTS AND INTERPRETATION

It is a general trend to use a scatter plot when analysing the relationship that exists between two variables. Although a scatter plot assist in visualizing data, it however does not necessary show the relationship between the two variables hence the application of modelling is necessary. As mentioned, different modelling techniques such as least square regression, quantile regression, logistic regression, k-nearest neighbour etc. are used. Each technique models data differently and this has an influence on the final output as well as the accuracy of the model. Majority of the statistical based techniques require the use of a scatter plot when modelling data. Machine learning techniques on the other hand do not require graphical plots. The effectiveness of a modelling technique is evaluated based on its ability to accurately predict an outcome of an event i.e. points resulting to rock failure in the context of this research.

In 1971, Mogi developed an empirical 3D criterion after conducting experiments in a number of rocks. This criterion was further supported by several authors including Colmenares and Zoback (2002). According to Colmenares and Zoback (2002), three types of functions can be used in fitting data when applying Mogi's criterion i.e. the power law, linear and second-polynomial. When Mogi conducted his experiment, three principal stresses were the end products. From the three stresses, he then calculated the octahedral shear and mean effective normal stress. These two latter stresses are the bases of the Mogi criterion for brittle failure in rocks. The relationship between the two stresses can be explained using the functions already mentioned above. According to Colmenares and Zoback (2002), the functions are used as lines of best fit and in rock engineering, they are considered as failure envelopes. The lines of best fit (failure envelopes) are used as a classifying parameter i.e. the area below the envelope is considered as the area of stability and the area above it represents the region of failure. Whatever lies in the failure envelope is considered as indecisive. Modelling techniques that utilize the line of best fit follow the failure envelop method when predicting failure. On the other hand, modelling techniques that do not use a graphical plot make use of an additional parameter i.e. rock strength to model data. In this chapter, the results of the eleven rocks modelled using the four modelling techniques are analysed and interpreted.

4.1. Statistical techniques

In this part of the research, the results from the eleven rocks were individually explained. For the regression modelling, a table and graphical format were used to explain the results. The graph outlined a parameter that describes the rate of change of the octahedral shear stress with respect to the mean effective normal stress. This parameter is known as the gradient/slope. On the other hand, the table illustrated an R^2 value which is a parameter that measures the relationship between the observed data and the fitted model. As explained previously, when R^2 approaches 1, a perfect fit between the model and data exists and if it approaches 0, then no relationship exist. A rule of thumb proposed by Moore, .et.al. (2013) was used when analysing R^2. This divided R^2 in to different categories ranging from very weak to strong. When $R^2 < 0.3$, it is considered as very weak whilst $0.3 < R^2 < 0.5$ is fairly weak. When R^2 is between 0.5 and 0.7, it is said to be moderately weak and when greater than 0.7, it is interpreted as strong (Moore, et.al. 2013).

The graphical representation was used as failure estimation and consists of least square as well as quantile regression lines. In rock engineering, these regression lines are a representative of a failure criterion. Points lying below the line of best fit are considered as stable and those above the line represents failure. This representation was further used in comparing the total number of points failing when the least square regression and quantile regression was used. It should be noted that the 0.25 quantile could not be modelled due to the nature of the data, hence two quantiles (0.5 and 0.75) are represented.

4.1.1. Solnhofen limestone

Figure 4.1 compares the nature in which the least square regression and quantile regression model data. The least square is shown by the red line while the quantile regression is illustrated by the 2 lines with different shades of blue. The 0.5 quantile is represented by the light blue line while the 0.75 quantile is shown by the dark blue colour. Giving reference to Figure 4.1, the three models show variation in the gradient. The least square regression line has a steeper gradient followed by the 0.5 quantile, then the 0.75 quantile. Furthermore, the least square line intersects the 0.5 quantile line at a mean effective normal stress of 290 MPa and octahedral shear stress of 180 MPa. It also intersects the 0.75 quantile line at a mean effective normal stress of 330 MPa and octahedral shear stress of 230 MPa. It is observed that

the 0.5 and 0.75 quantile intersect at a mean effective normal stress of 360 MPa and octahedral shear stress of 250 MPa. The two lines then merge and form a single slope after these stresses.

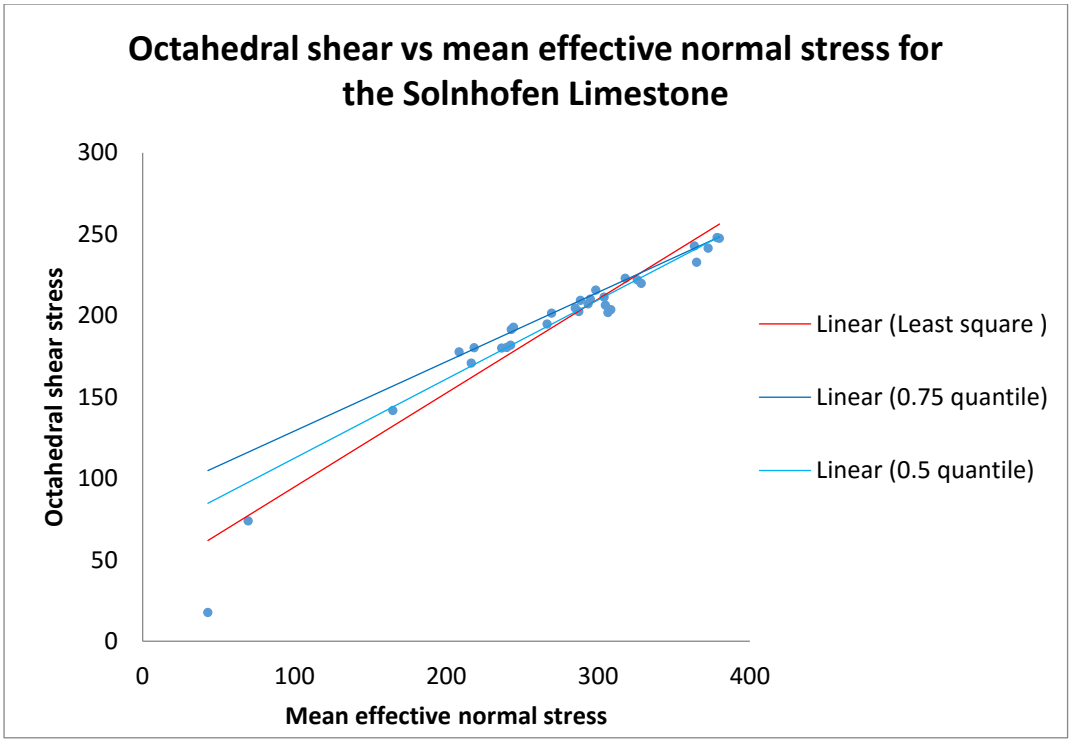

Figure 4.1: Relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Solnhofen Limestone data.

As stated already, the relationship between the observed data points and the fitted model is evaluated using R^2. Table 4.1 shows the R^2 values for the models fitted in the Solnhofen limestone data. The R^2 for all models are above 0.7. With the use of Moore et.al. (2013) rule of thumb, the three models perfectly fit the data.

Table 4.1: The R^2 values evaluating the relationship between the observed Solnhofen limestone data and the fitted models.

Model type	R^2
Least square regression	0.933
0.50 quantile	0.948
0.75 quantile	0.971

Failure prediction

Failure prediction is determined using the regression lines in Fig 4.1 as failure envelops. These failure envelopes are used to estimate the number of points lying either in the stable or failing region. For quantile regression, the number of points plotting in the stable region are calculated as the number of failing points of the previous quantiles e.g. if 3 points failed in the 0.5 quantiles, these points will be considered as stable points for the 0.75 quantile. In some instances, the points lie exactly in the regression lines which are then grouped as indecisive points i.e. they can neither be predicted as stable nor failing. For each rock, three rock states will be of focus i.e. stable, failing and indecisive. Figure 4.2 is a bar graph plotted from estimations made in Figure 4.1. This graph shows that at 0.5 quantile, 9 points lie in the stable region and 6 points plot in the failure region. Furthermore, 7 points plot in the regression line and they are indecisive. In the upper quantile i.e. 0.75 quantile, 6 points are stable, 5 points are failing and 4 points fall under the undecided group. With the use of the least square regression model, it is predicted that 14 points will be stable, 15 will fail and 2 are indecisive. Pertaining to indecisive points, quantile regression shows the highest as 11 points are indecisive whilst only 2 points are indecisive when least square regression is used.

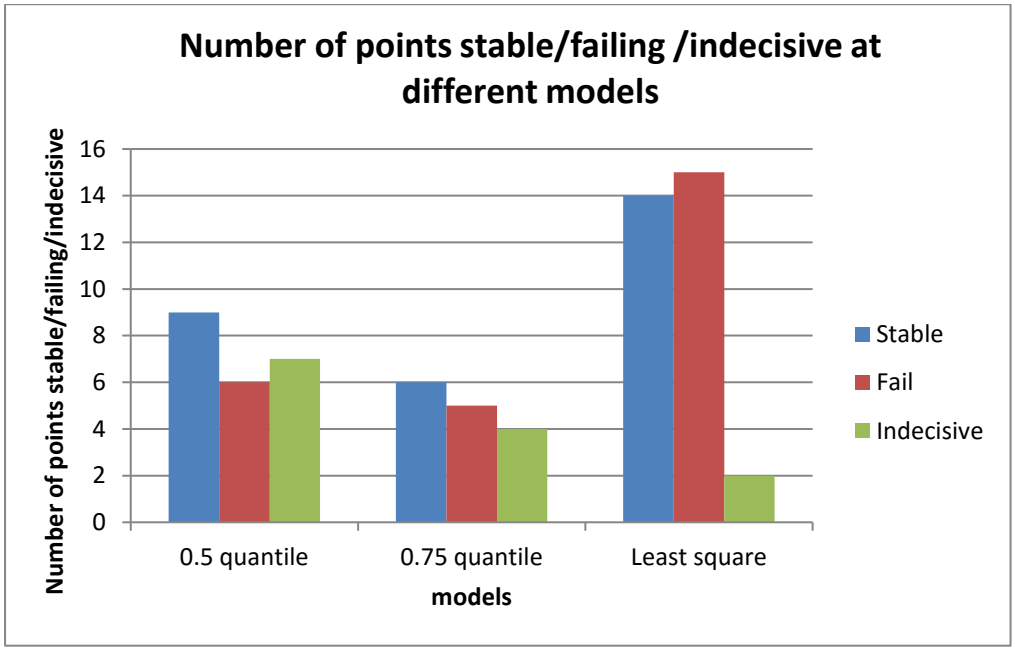

Figure 4.2: The number of points classified as stable/failing/indecisive using both statistical techniques for the Solnhofen Limestone data

Figure 4.3 is a graph comparing failure prediction using the least square regression and quantile regression modelling technique. From this figure, the least square regression predicts

a total number of 15 points to result to rock failure while the quantile regression estimates 11 points. In quantile regression, this was determined by summing the points failing at 0.50 and 0.75 quantiles.

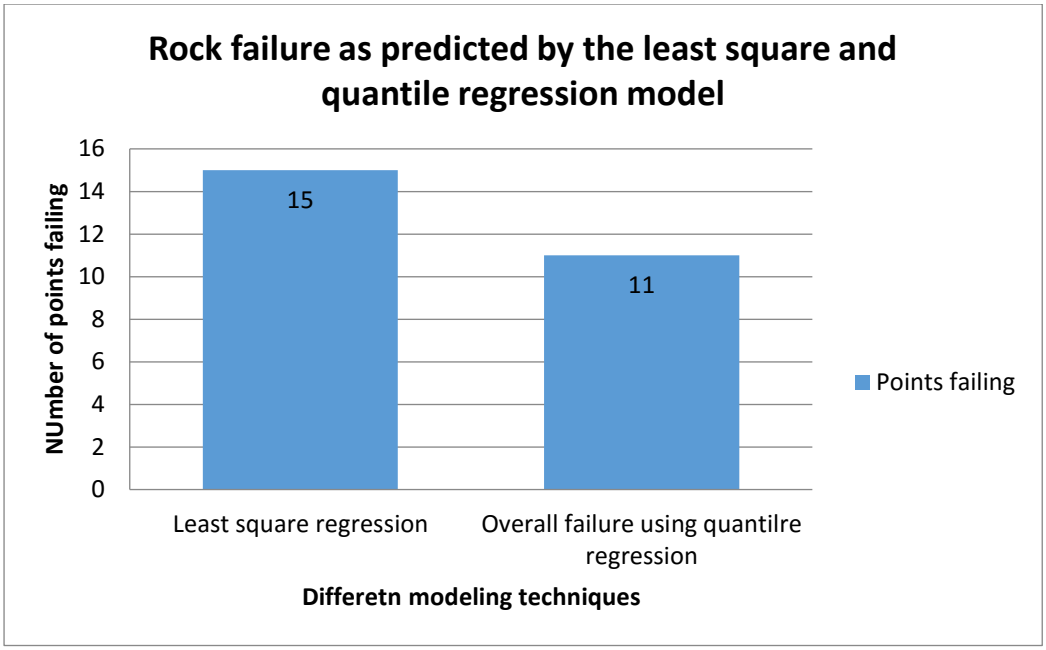

Figure 4.3: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Solnhofen Limestone.

4.1.2. Dunham Dolomite

In Figure 4.4, the behaviour of all models is almost the same since their gradients are fairly similar. Initially, the regression lines of all three models are parallel to each other. The least square regression and 0.5 quantile line then intersect at a mean effective normal stress and octahedral shear stress of 400 MPa and 250 MPa, respectively. These lines merge and form a single gradient. Parallel lines suggest a similar gradient, meaning the gradient of the least square regression and the two quantiles is the same in the lower stresses.

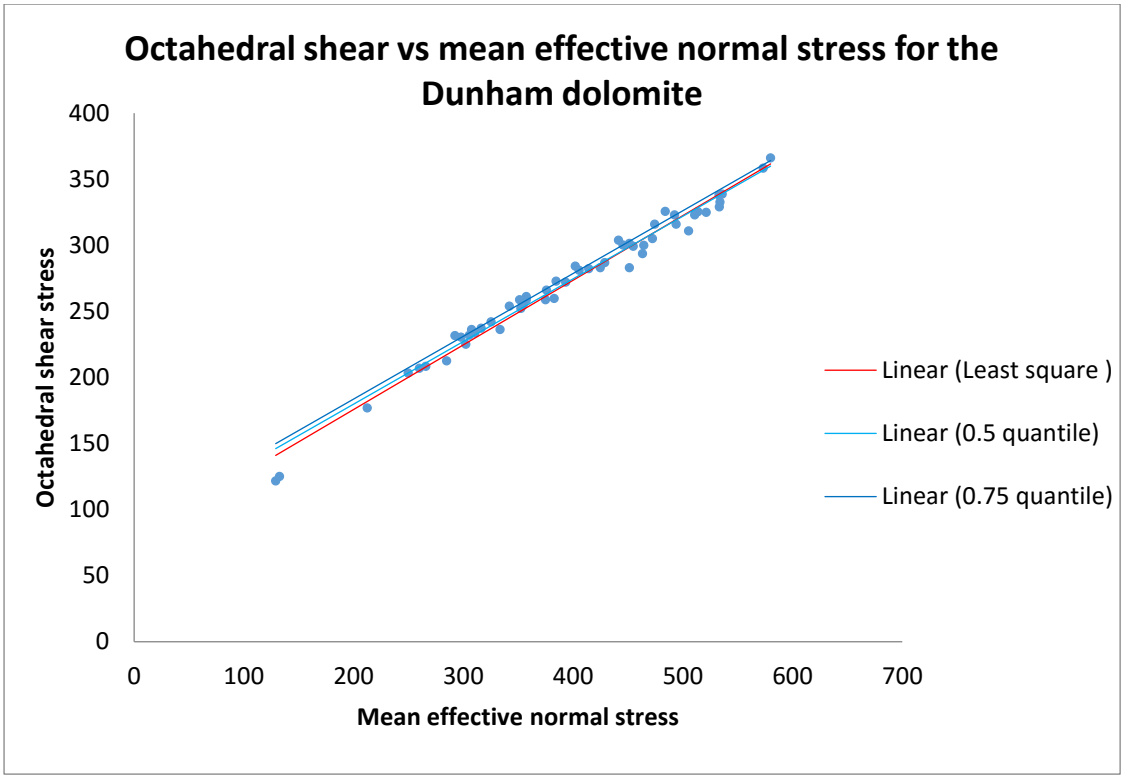

Figure 4.4: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Dunham Dolomite data.

From Table 4.2, it can be observed that the 0.50 quantile and the least square regression have similar R^2 values of 0.970, whilst the R^2 value of the 0.75 quantile is 0.986. In overall, the R^2 values of all the models are above 0.7 suggesting a strong perfect fit. This implies that a linear relationship exists between the points.

Table 4.2: R^2 values evaluating the relationship between the observed Dunham Dolomite data and the fitted models.

Model type	R^2
Least square regression	0.970
0.50 quantile	0.970
0.75 quantile	0.986

Failure prediction

Based on Figure 4.5, more points are stable at the 0.5 quantile (17) while 6 points are predicted to fail. A higher failing rate is observed in the 0.75 quantile as 10 points are estimated to fail while 6 are stable. On the other hand, the least square predicts a failure of

about 28 points while 17 points are stable and 9 indecisive. In total, 21 points are indecisive for the quantile regression.

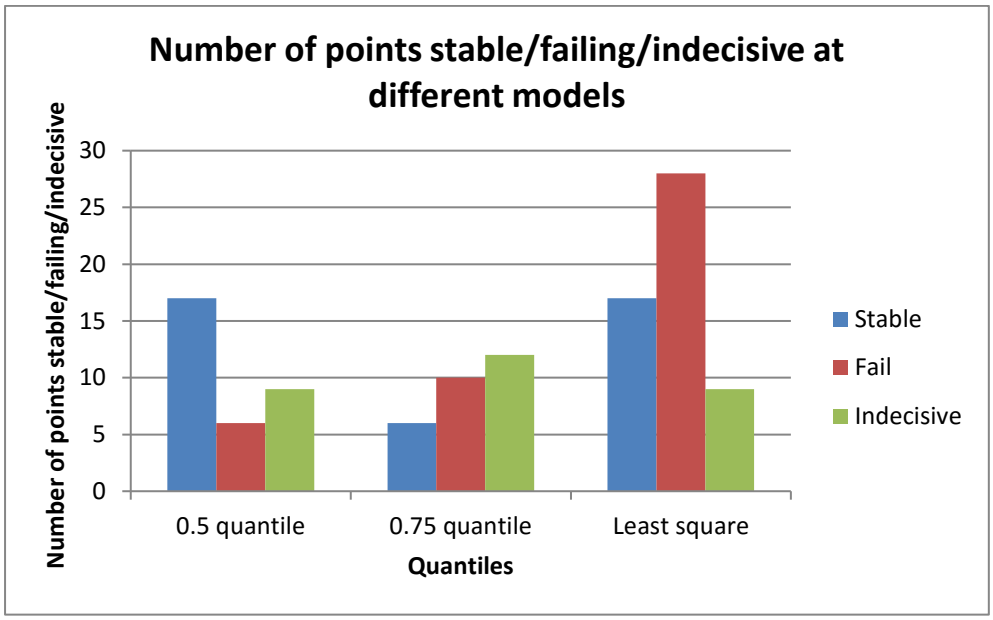

Figure 4.5: The number of points classified as stable/failing/indecisive using both statistical techniques for the Dunham Dolomite data.

Figure 4.6 shows the overall number of points failing predicted by both modelling techniques. The least square regression predicts that 28 points will lead to rock failure while quantile regression estimates 16.

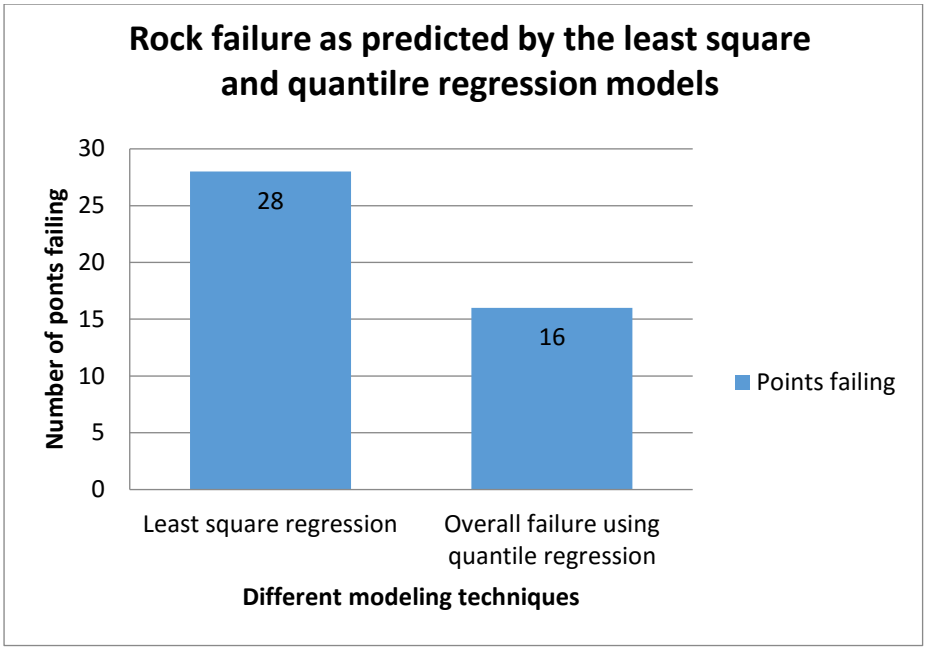

Figure 4.6: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Dunham Dolomite

4.1.3. Yamaguchi Marble

With reference to Figure 4.7, the 0.5 quantile and least square regression lines overlap throughout. Furthermore, the overlapping lines are almost parallel to the slope of the 0.75 quantile. This suggests that the gradient of all three models is the same. Pertaining to the nature of the slope, it can be observed that the 0.75 quantile has a gentle slope compared to the slope of the least square and 0.5 quantile lines.

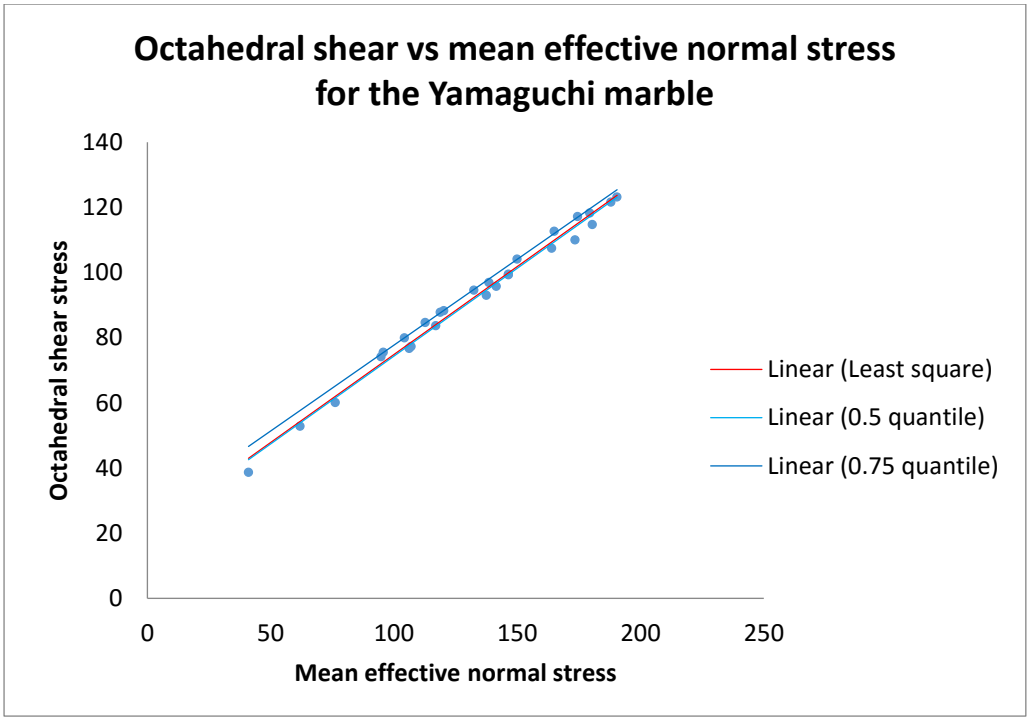

Figure 4.7: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Yamaguchi Marbles data.

The overall R^2 values correlating to each modelling technique are above 0.7 as seen in Table 4.3. This indicates a strong perfect fit meaning the data is behaving in a linear manner. The 0.5 quantile has the lowest R^2 value while the least square regression has the highest value.

Table 4.3: R^2 values evaluating the relationship between the observed Yamaguchi Marbles data and the fitted models

Model type	R^2
Least square regression	0.988
0.50 quantile	0.973
0.75 quantile	0.987

Failure prediction

Figure 4.8 shows that 2 points are likely to fail in the 0.5 whilst only 1 point is estimated to fail in the 0.75 quantile. The number of points plotting in the stable region for both the 0.5 and 0.75 quantile are 9 and 2, respectively. The points plotting directly on the regression line and thus cannot be categorized are quite high in the 0.75 quantile. On the other hand, the least square estimates that 12 points are stable and 11 points are resulting in failure. In overall, 14 points are indecisive for the quantile regression and 3 are indecisive in the least square

regression. This implies that more than 50% of the points can neither be classified as stable or failing when using the quantile regression.

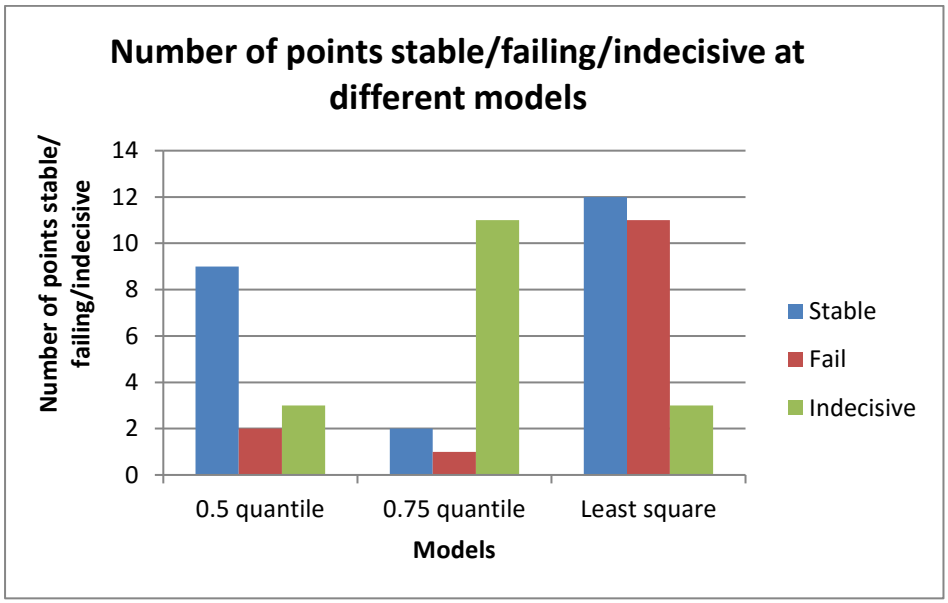

Figure 4.8: The number of points classified as stable/failing/indecisive using both statistical techniques for the Yamaguchi Marble data.

Figure 4.9 compares the number of points resulting to failure using both the least square regression and the quantile regression prediction models. The least square predicted 11 points as failing while the quantile regression estimated a total of 7 failing points.

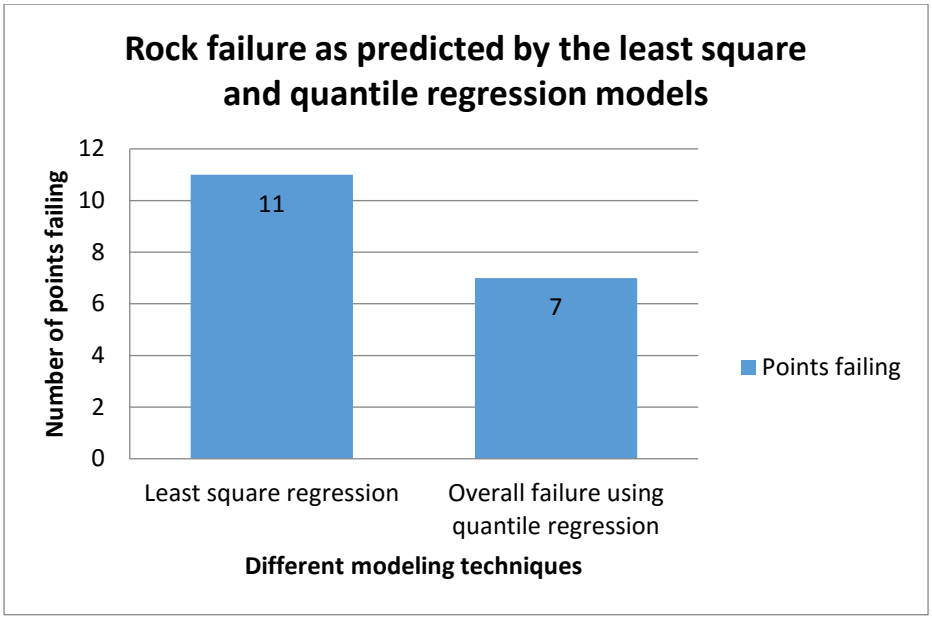

Figure 4.9: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Yamaguchi Marble.

4.1.4. Mizuho Trachyte

Figure 4.10 shows the comparison between the least square regression and quantile regression models. Initially, the 0.5 quantile and 0.75 quantile lines overlap and diverge at a mean effective normal stress of 250 MPa. It is noted also that the least square regression line intersect the 0.5 quantile at 240 MPa and 140 MPa, and further intersects the 0.75 quantile line at a mean effective normal stress and octahedral shear stress of 300 MPa and 180 MPa, respectively. The least square regression line is observed to be steeper as compared to the 0.5 and 0.75 quantile lines.

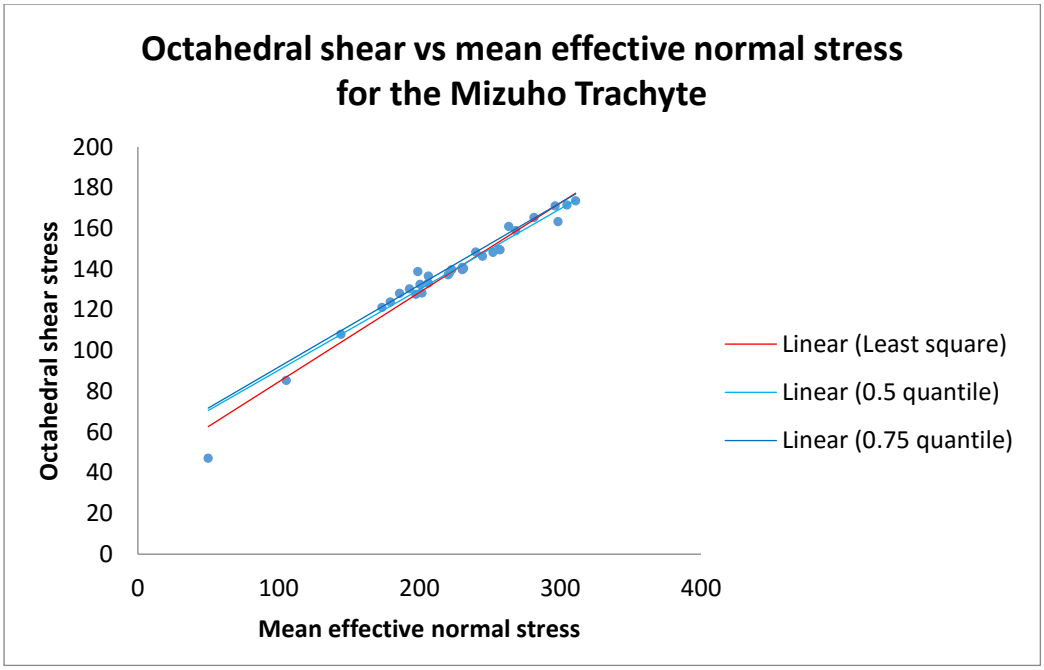

Figure 4.10: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Mizuho Trachyte data.

All models show an R^2 within the range of 0.7 -1.0 suggesting that the models are a perfect fit and can be used to determine the relationship between the octahedral shear and mean effective normal stress. The 0.5 quantile regression has the lowest R^2 (0.963) while the 0.75 quantile has the maximum R^2 (0.982).

Table 4.4: R^2 values evaluating the relationship between the observed Mizuho Trachyte data and the fitted models

Model type	R^2
Least square regression	0.964
0.50 quantile	0.963
0.75 quantile	0.982

Failure prediction

When analysing Figure 4.11, it is observed that most points are predicted to fail at 0.75 quantile (5) while 2 points are failing at the 0.50 quantile. Pertaining to indecisive points, the 0.5 quantile has the highest points as compared to the 0.75 quantile. The least square on the other hand predicts that 12 points are stable and 11 points are failing. In total, more points are indecisive when the quantile regression (13) model is used compared to the least square regression with 8 points.

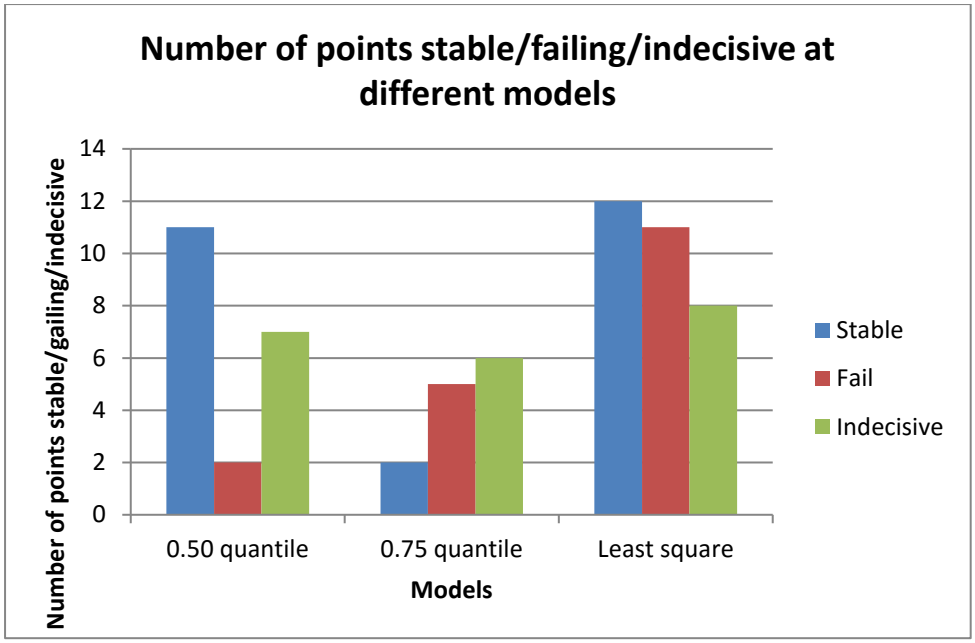

Figure 4.11: The number of points classified as stable/failing/indecisive using both statistical techniques for the Mizuho Trachyte data.

Figure 4.12 compares the overall number of failing points predicted by the least square and the quantile regression model. The least square has estimated that 11 points out of 31 will result to the rock failing while quantile regression has predicted 7 points to cause rock failure.

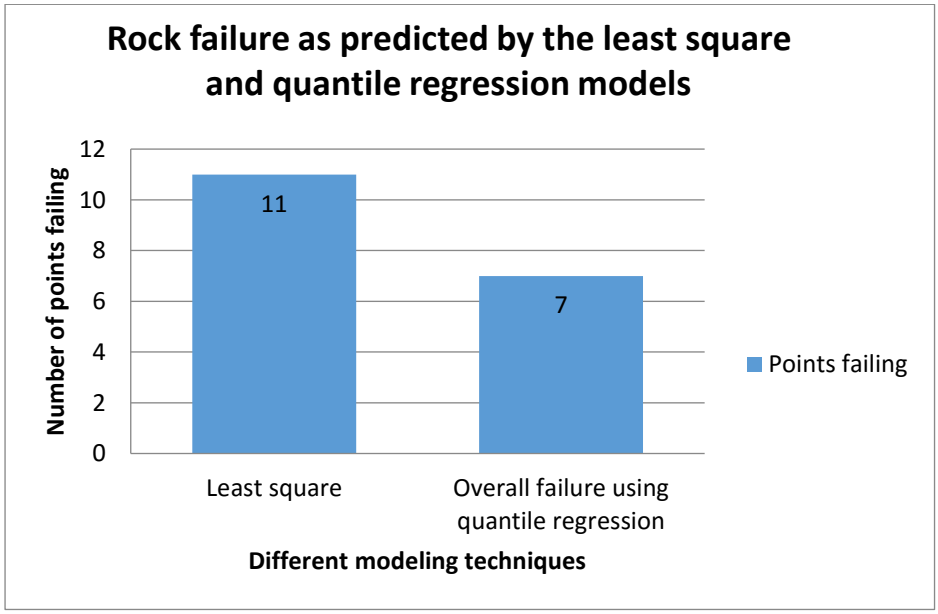

Figure 4.12: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Mizuho Trachyte

4.1.5. Orikabe Monzonite

Figure 4.13 shows the fitted lines of best fit for the data of the Orikabe Monzonite rock. The 0.5 and 0.75 quantile lines are quite close and parallel to each other. At a mean effective normal stress and octahedral shear stress of 700 MPa and 480 MPa, all three regression slopes merge and form a single gradient. Initially, the least square shows a steep slope as compared to the two quantiles.

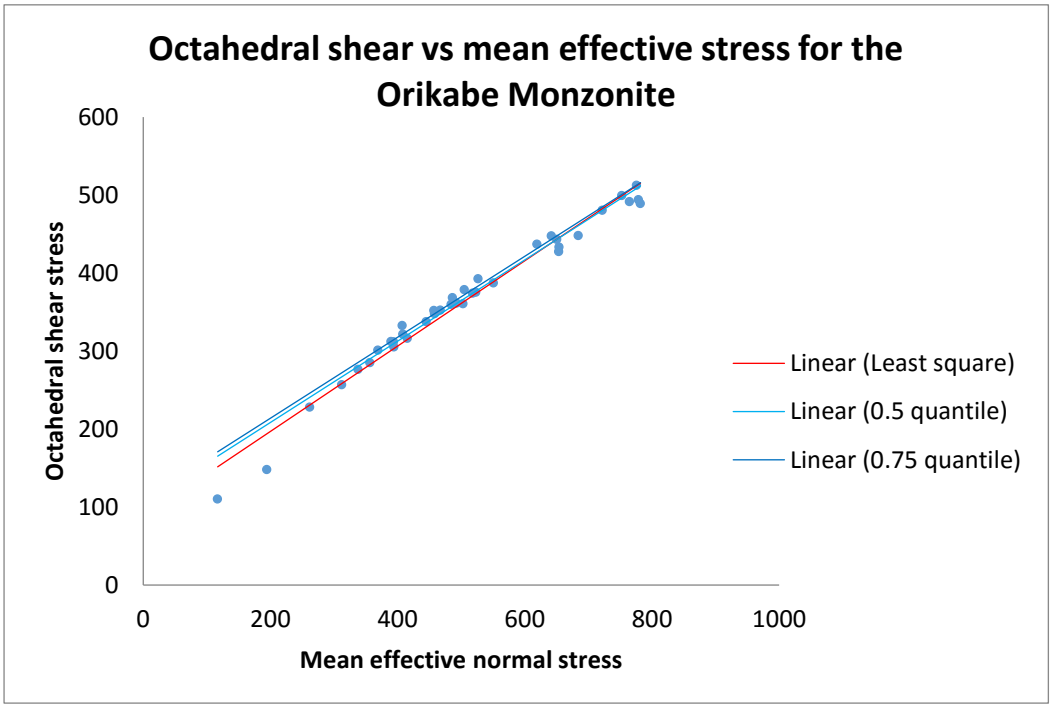

Figure 4.13: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Orikabe Monzonite data.

The R^2 values for all the models are greater than 0.7. This implies that the models can be used to predict failure with greater confidence. The 0.5 quantile has the minimum value (0.962) while the 0.75 has the maximum value (0.982). The R^2 value of the least square regression lies in between that of the 0.5 and 0.75 quantiles.

Table 4.5: R^2 values evaluating the relationship between the observed Orikabe Monzonite data and the fitted models

Model Type	R^2
Least square regression	0.974
0.5 quantile	0.962
0.75 quantile	0.982

Failure prediction

Figure 4.14 shows the number of points stable, failing and indecisive when least square and quantile regression modelling techniques are applied. From this figure, it is observed that the quantile regression classified most points as being stable as a total of 20 points were estimated to be stable when using this model. Pertaining to failure prediction, a total of 12 points are likely to result in rock failure. The least square regression estimated a higher failing rate as 16 points are predicted to cause rock failure while 9 points are classified under the stable group. In terms of indecisive points, the least square has 13 whilst quantile regression shows 11 points to be neither stable nor failing.

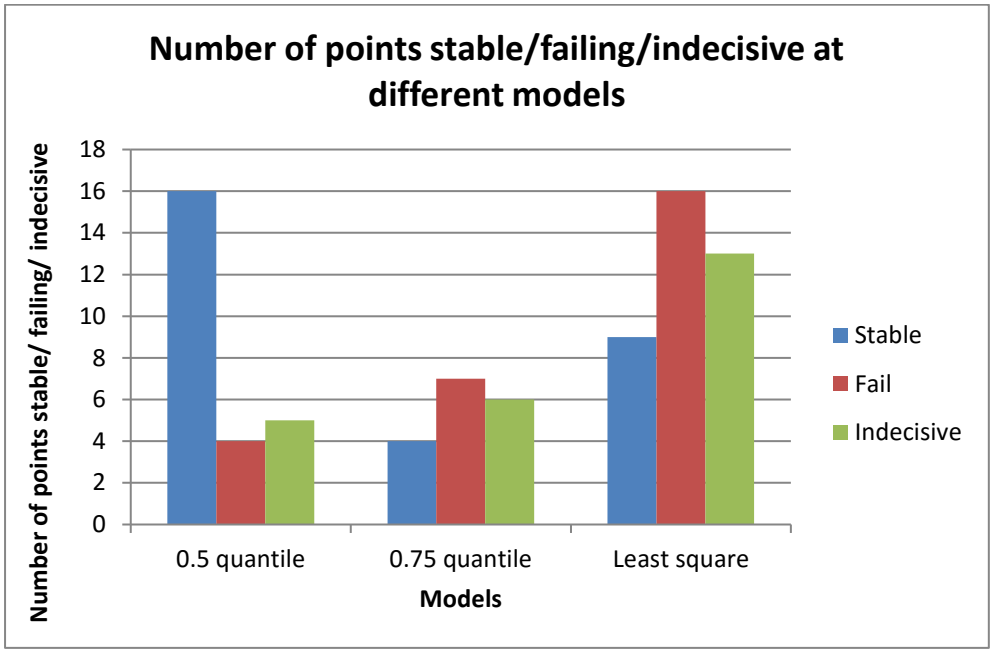

Figure 4.14: The number of points classified as stable/failing/indecisive using both statistical techniques for the Orikabe Monzonite data.

In overall, the least square has estimated more points to result to rock failure than the two quantiles as 16 points are predicted to fail compared to the 12 points estimated by the two quantiles. This is shown in figure 4.15.

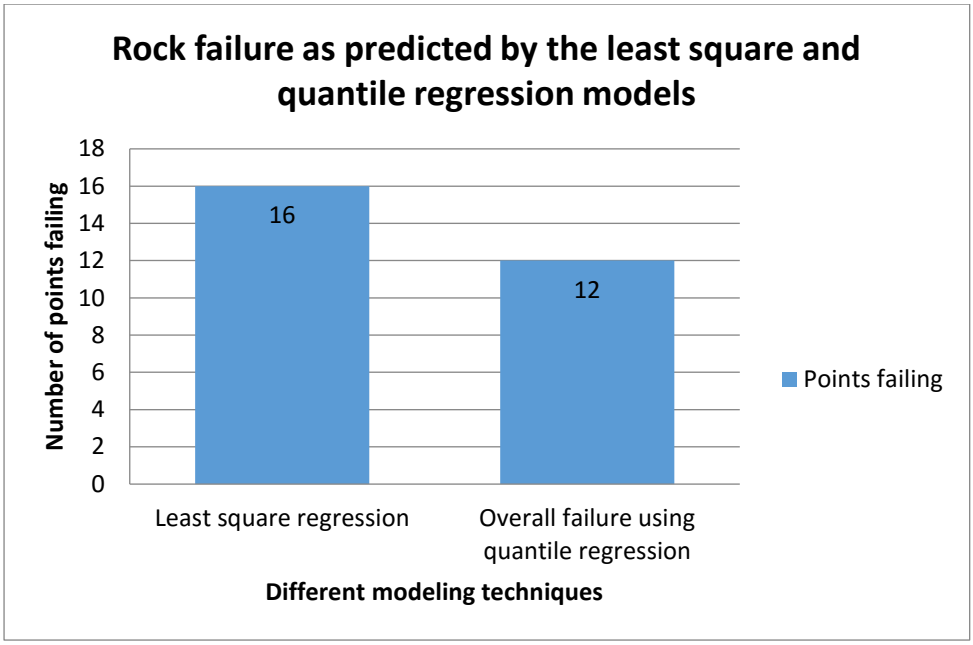

Figure 4.15: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Orikabe Monzonite.

4.1.6. Manazuru Andesite

Figure 4.16 shows modelling of the Manazuru Andesite data using least square and quantile regression. It is observed that the three models lie parallel and very close to each other. This suggests that the gradients of the three models are the same as they are parallel.

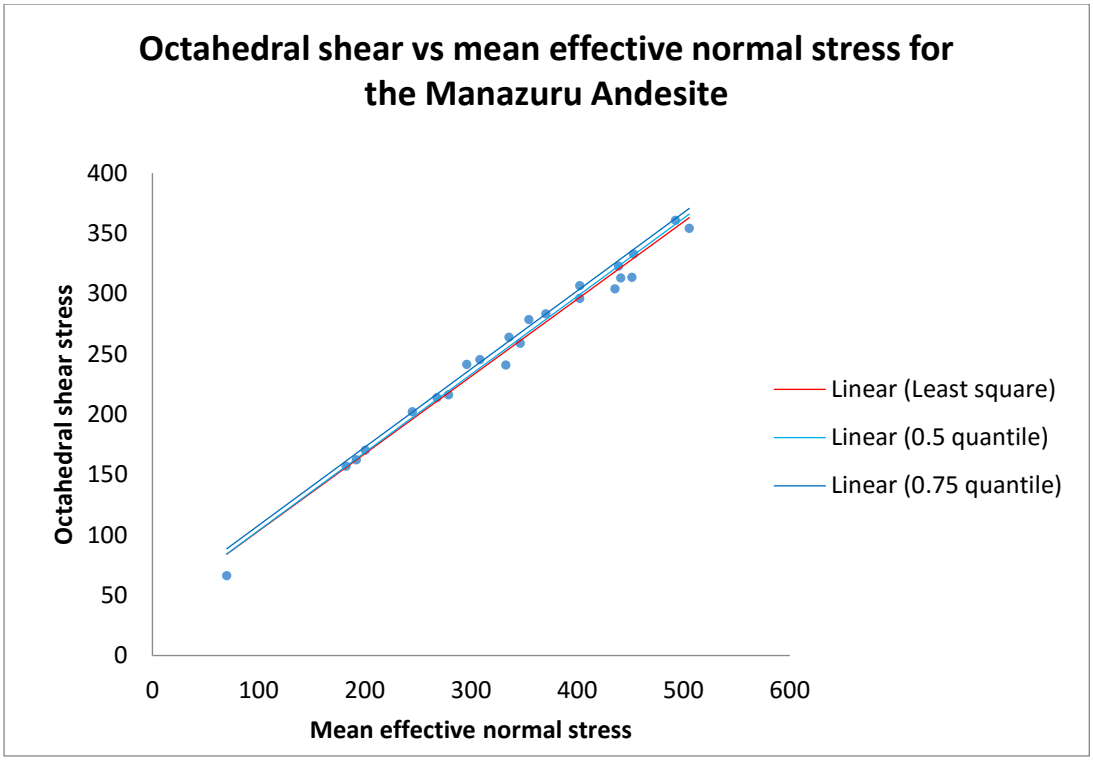

Figure 4.16: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Manazuru Andesite data.

The R^2 values for the models range between 0.968 and 0.985 as shown in table 4.6. These values are greater than 0.7 and according to Moore, et.al. (2013) they indicate a good fit. This implies that the data can be predicted using a linear function such as least square or quantile regression. From Table 4.6, the perfect fit is achieved when modelling the data using the 0.75 quantile as it has a maximum R^2 value of 0.985.

Table 4.6: R^2 values evaluating the relationship between the observed Manazuru Andesite data and the fitted models

Model Type	R^2
Least square regression	0.984
0.5 quantile	0.968
0.75 quantile	0.985

Failure prediction

According to Figure 4.17, three points are predicted to fail in the 0.5 quantile whilst majority of the points are stable (8). In the upper quantile, 3 points are predicted to be stable, 5 points classified under the failure region and 3 points plot exactly on the line and cannot be

categorized. For the least square regression, 30% of the points are indecisive, 35% are estimated to fail and a further 35% of the points belong to the stability region.

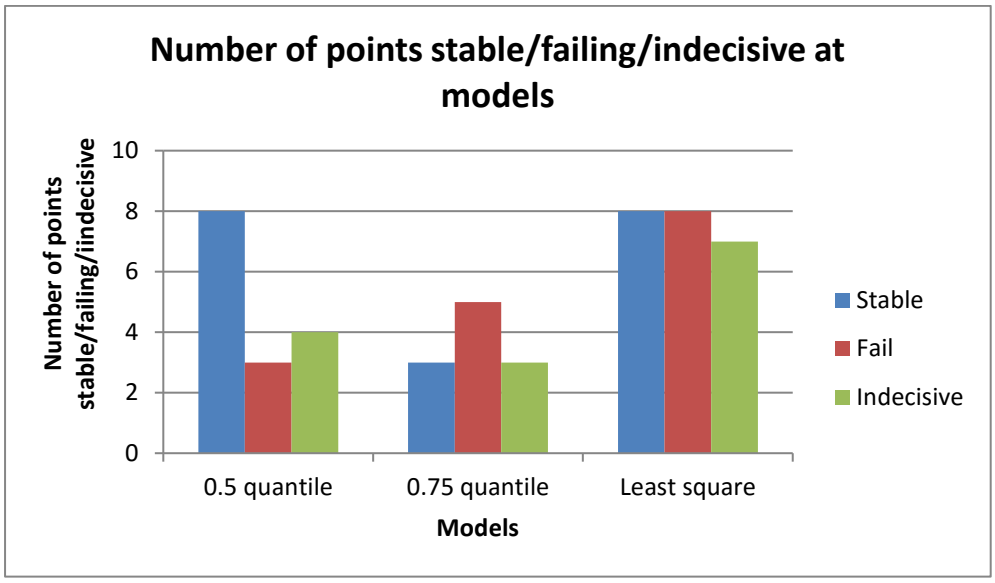

Figure 4.17: The number of points classified as stable/failing/indecisive using both statistical techniques for the Manazuru Andesite data.

Figure 4.18 shows the failure estimation using the two modelling techniques. Both the least square regression and quantile regression predicted that 35% of the points are likely to result to rock failure.

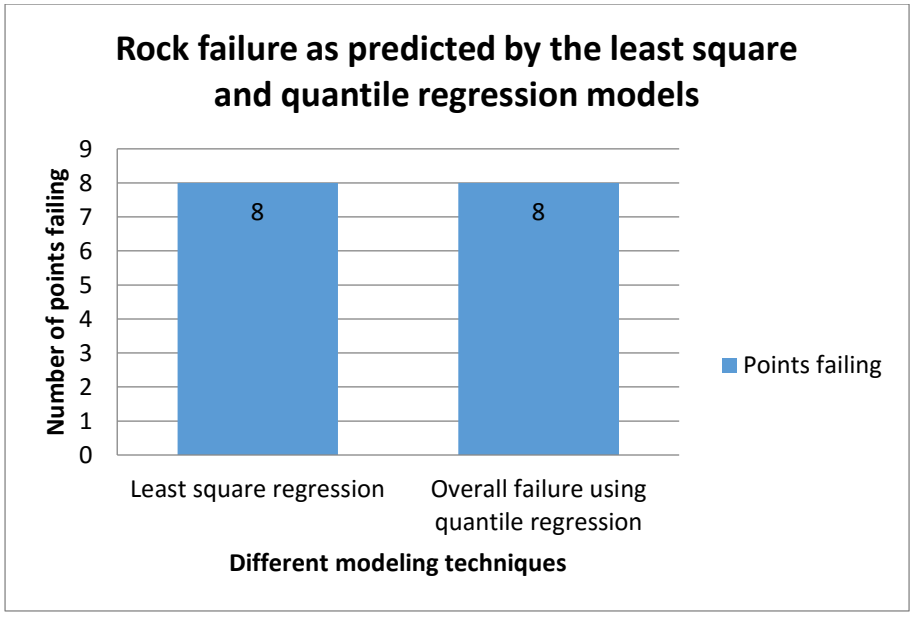

Figure 4.18: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Manazuru Andesite.

4.1.7. Inada Granite

From Figure 4.19, the change in the behaviour of the gradients between the lower and upper mean effective normal stresses is observed. Within the lower stresses, the gradient of the least square and 0.5 quantile regressions are parallel to each other. At around a mean effective normal stress of 620 MPa and octahedral shear stress of 450 MPa, the three gradients merge and form a single gradient. It is also noted that the least square has a steeper gradient followed by the 0.5 quantile whilst the 0.75 quantile has the least steep gradient.

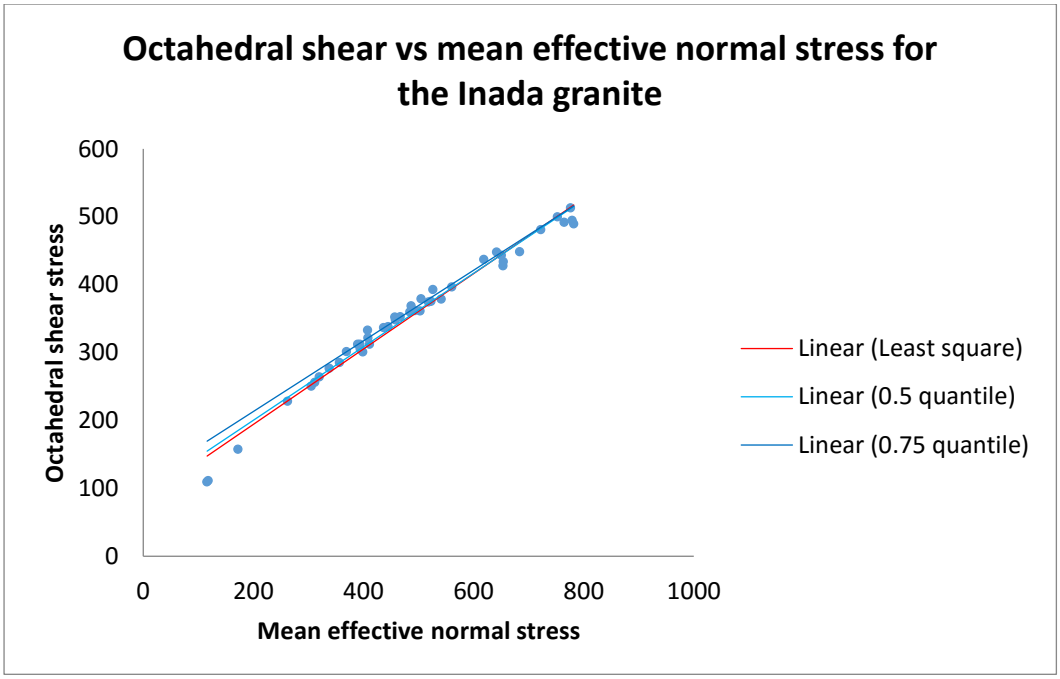

Figure 4.19: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Inada Granite data.

Table 4.7 explains the nature in which the models fit the observed data. For all models, the coefficient of determination values are above 0.7 implying that these models are a perfect fit and estimation can be carried out with about 95% confidence. The 0.5 quantile has the least R^2 while the 0.75 quantile has the highest value.

Table 4.7: R^2 values evaluating the relationship between the observed Inada Granite data and the fitted models

Model type	R^2
Least square regression	0.979
0.5 quantile	0.962
0.75 quantile	0.981

Failure prediction

From Figure 4.20, it can be observed that majority of the points are stable in the 0.5 quantiles compared to the 0.75 quantile. Within this quantile, 4 points are predicted to fail and 11 are indecisive. In the 0.75 quantile, 8 points are neither stable nor failing as they plot directly on the regression line while 7 are failing and 4 are stable. Giving reference to the least square regression, about 16 points are estimated to result to rock failure while 8 are stable and 20 are indecisive. In overall, the quantile regression has 19 indecisive points while the least square has 20 points. This implies that when using the quantile regression, 43% of the points cannot be classified as either stable or failing whilst for least square regression, 45% of the points are indecisive. Predicting rock failure of this rock is quite a challenge as majority of the points plot directly on the regression line.

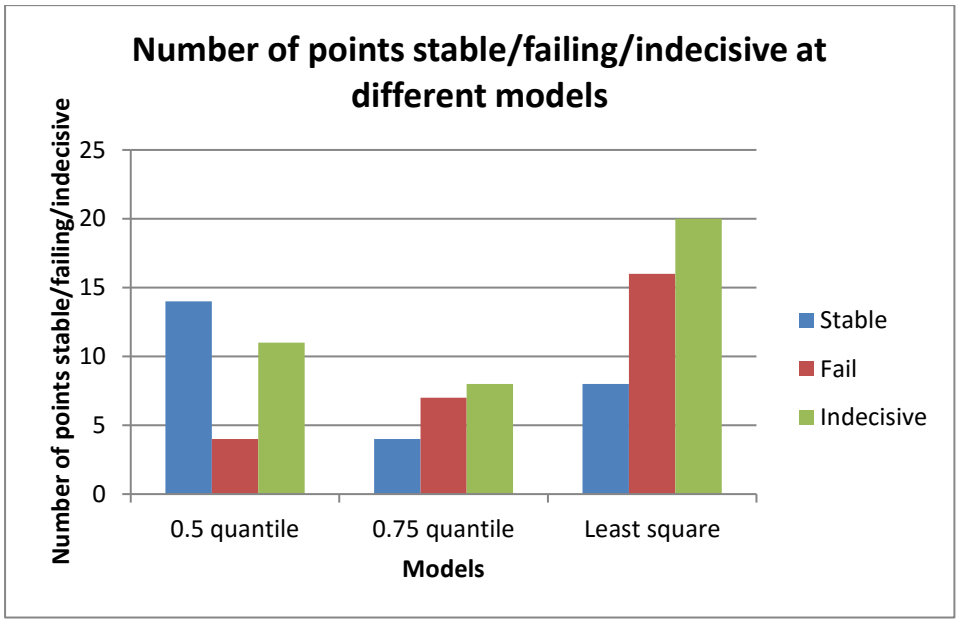

Figure 4.20: The number of points classified as stable/failing/indecisive using both statistical techniques for the Inada Granite data.

Figure 4.21 shows the overall number of points estimated to fail. The least square regression predicts that 16 points will fail while the quantile regression estimates a failure of about 11 points.

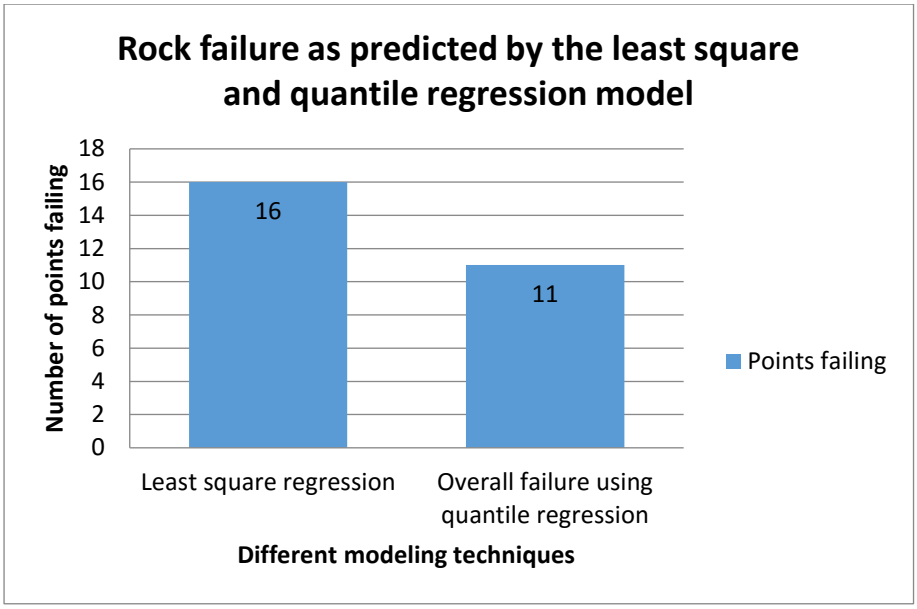

Figure 4.21: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Inada Granite

4.1.8. KTB Amphibolite

A comparison amongst the regression models is observed in Figure 4.22. It can be noted that the 0.5 quantile, the least square and 0.75 quantile lines overlap from 100 MPa to 380 MPa mean effective normal stress. A change is observed after 300 MPa as the 0.75 quantile line is separated from the least square and the 0.5 quantile regression models. It is also observed that the 0.75 quantile has a steeper slope as compared to the other two models.

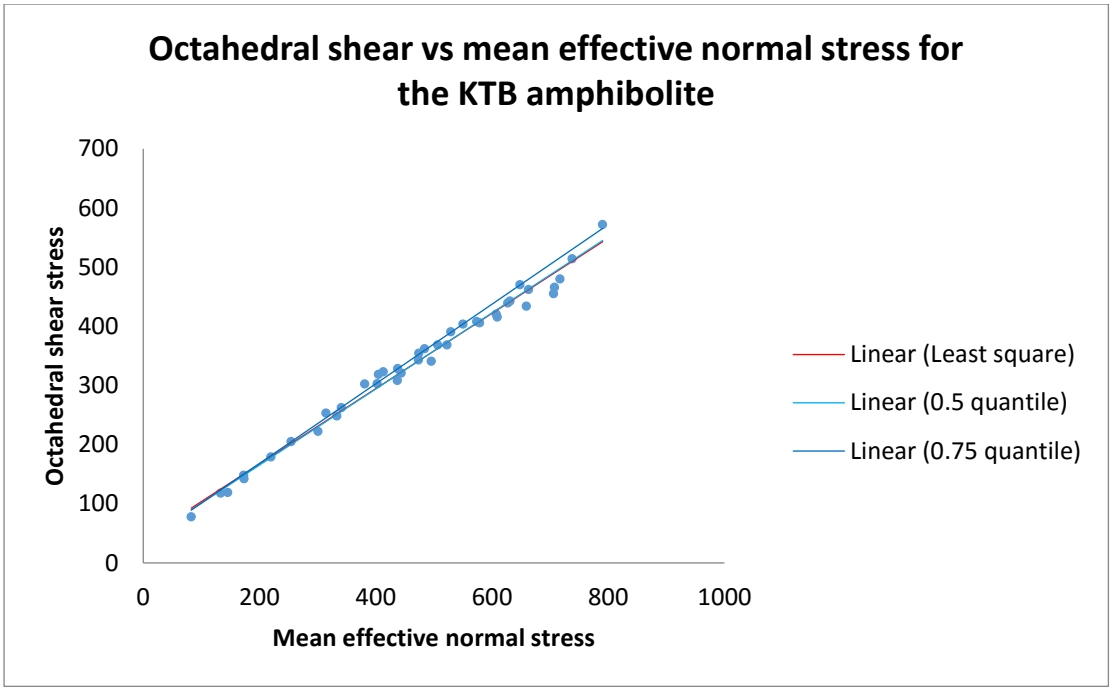

Figure 4.22: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the KTB Amphibolite data.

Table 4.8 shows the R^2 values of the different modelling methods. Based on this table, the three models fit the data perfectly as the R^2 values are greater than 0.7. The least square regression however fits the data much better than the other models and can be used with 95% confidence level.

Table 4.8: R^2 values evaluating the relationship between the observed KTB Amphibolite data and the fitted models

Model type	R^2
Least square regression	0.987
0.5 quantile	0.965
0.75 quantile	0.982

Failure prediction

As three rock classification states are used, Figure 4.23 illustrates the classification of each point according to the various statistical methods. Based on the graph, 4 points are estimated to result to failure in the 0.5 quantile regression while 12 points are stable. The 0.75 quantile estimated a higher failure rate as 8 points are likely to cause rock failure whilst 4 points are stable. On the other hand, the predictions made by the least square regression model are as

follows: 13 stable, 15 failing points and 12 are indecisive points. In total, it is impossible to classify 13 points when using quantile regression i.e. they are indecisive.

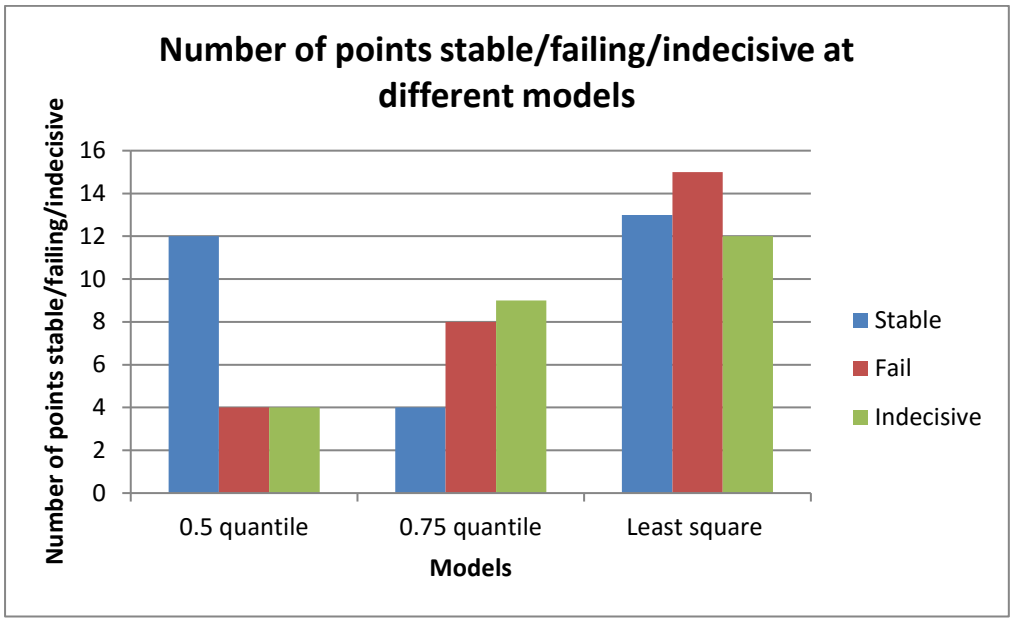

Figure 4.23: The number of points classified as stable/failing/indecisive using both statistical techniques for the KTB Amphibolite data.

Figure 4.24 compares the overall failure prediction using both models. The quantile regression technique predicts 12 points to fail while the least square estimates a failure of 15 points. In terms of percentages, the least square estimates a 43% failure while the quantile regression predicts a 30% failure.

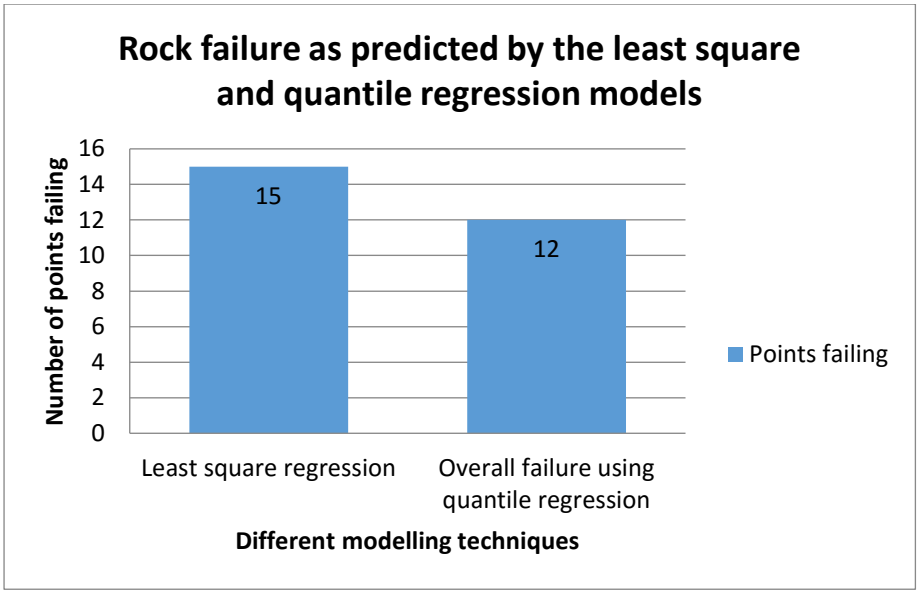

Figure 4.24: A comparison of failure prediction as estimated by the least square regression and quantile regression for the KTB Amphibolite

4.1.9. Westerly Granite

Figure 4.25 illustrates the modelling of the Westerly Granite data using the least square regression and quantile regression models. These models are represented by regression lines (gradients). The behaviour is similar to that of the KTB Amphibolite as the three models overlap in the lower mean effective normal stresses and split at around 360 MPa mean effective normal stress and 200 MPa octahedral shear stress. They then become parallel after splitting. The gradients show a minor deviation from each other. Moreover, the regression lines of this rock are steeper compared to those of the other rocks.

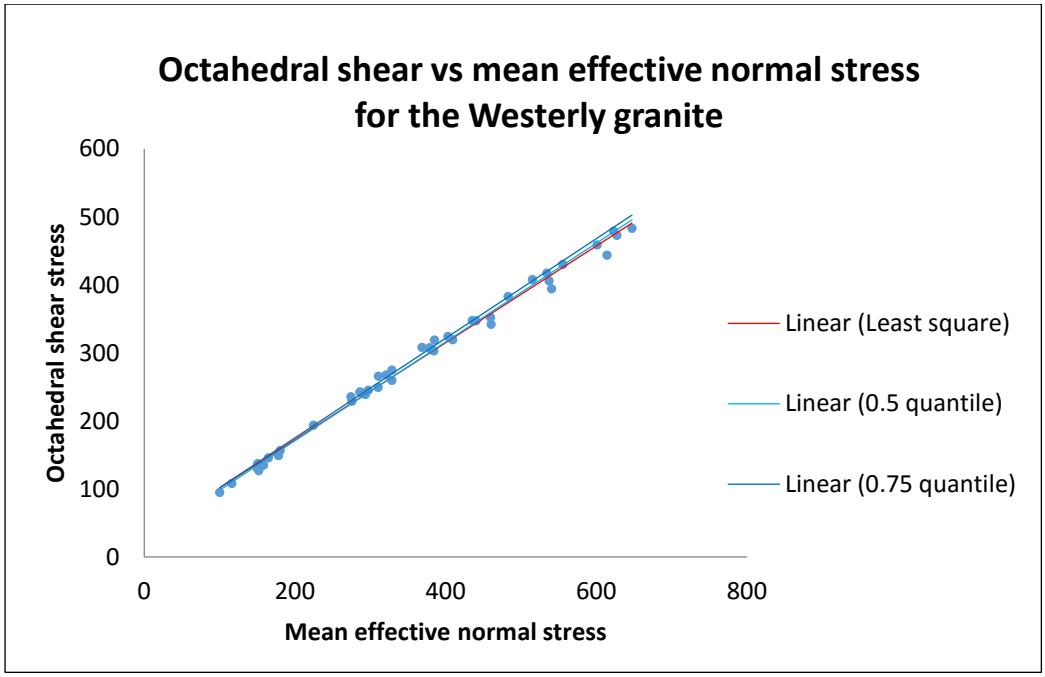

Figure 4.25: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Westerly Granite data.

The R^2 values in Table 4.9 are used to determine the relationship between the observed data and fitted model. The values in this table are all greater than 0.7 suggesting a strong fit. The least square regression has the highest R^2 value while the 0.5 quantile has the lowest coefficient of determination. Based on the R^2 values, it can be concluded that this data can be fitted using linear functions.

Table 4.9: R^2 values evaluating the relationship between the observed Westerly Granite data and the fitted models

Model type	R^2
Least square regression	0.994
0.5 quantile	0.974
0.75 quantile	0.987

Failure prediction

Figure 4.26 shows the number of points stable, indecisive and resulting to failure. The 0.50 quantiles has estimated 5 points to result to failure whilst the 0.75 quantile has estimated 9 points to result to rock failure. Most points (13) are estimated as being stable in the 0.5 quantile while 5 points are predicted to be stable in the 0.75 quantile. The number of points plotting directly on the 0.5 regression line is high i.e. 11 points suggesting that these points

are indecisive. The least square model on the other hand estimated 13 points to be stable and 20 points as failing. A total of 19 and 13 points are indecisive for the quantile regression and least square regression, respectively.

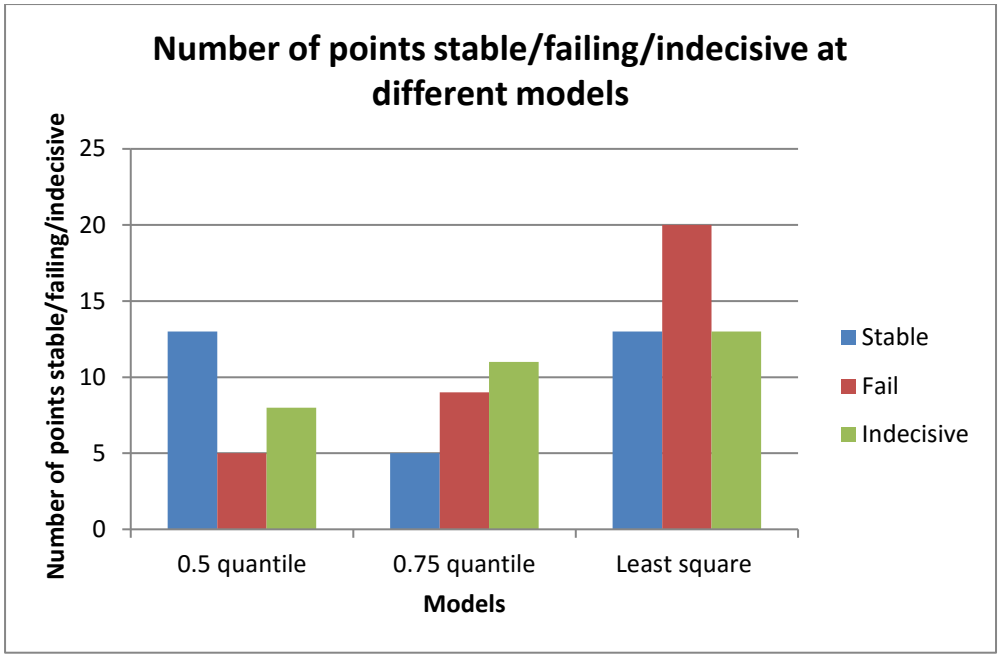

Figure 4.26: The number of points classified as stable/failing/indecisive using both statistical techniques for the Westerly Granite data.

Based on Figure 4.27, it can be noted that the quantile regression technique has estimated 20 points to fail whilst the least square model has predicted 14 points. In Percentages, the least square has estimated a 43% failure while quantile regression predicted that 30% of the points will result in rock failure

.

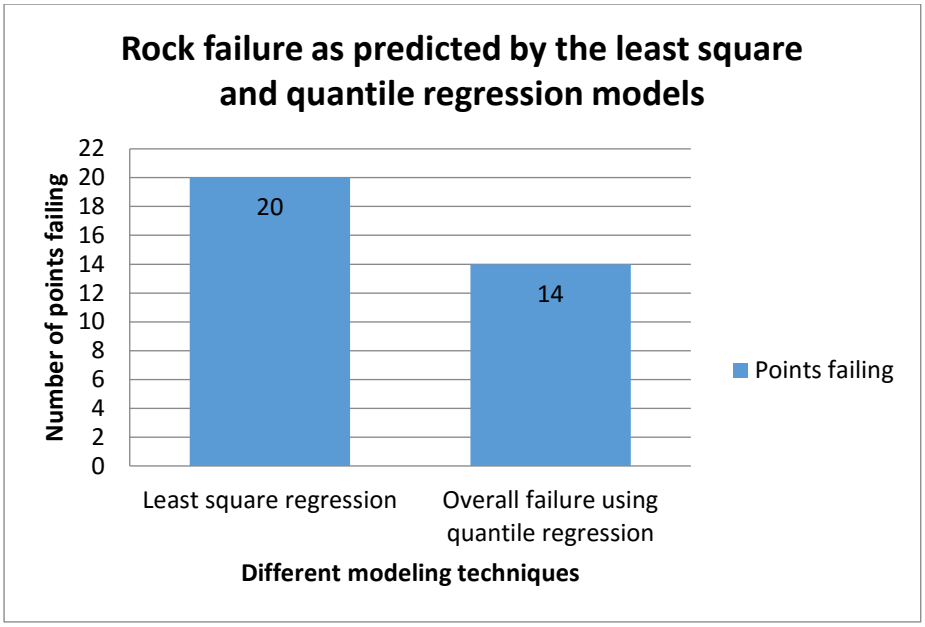

Figure 4.27: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Westerly Granite.

4.1.10. Shirahama Sandstone

The regression lines in Figure 4.28 all show great similarities. The least square and the 0.5 quantiles gradients overlap throughout. The overlapping lines are parallel to the gradient of the 0.75 quantile regression. The gradient of all three models is 0.544.

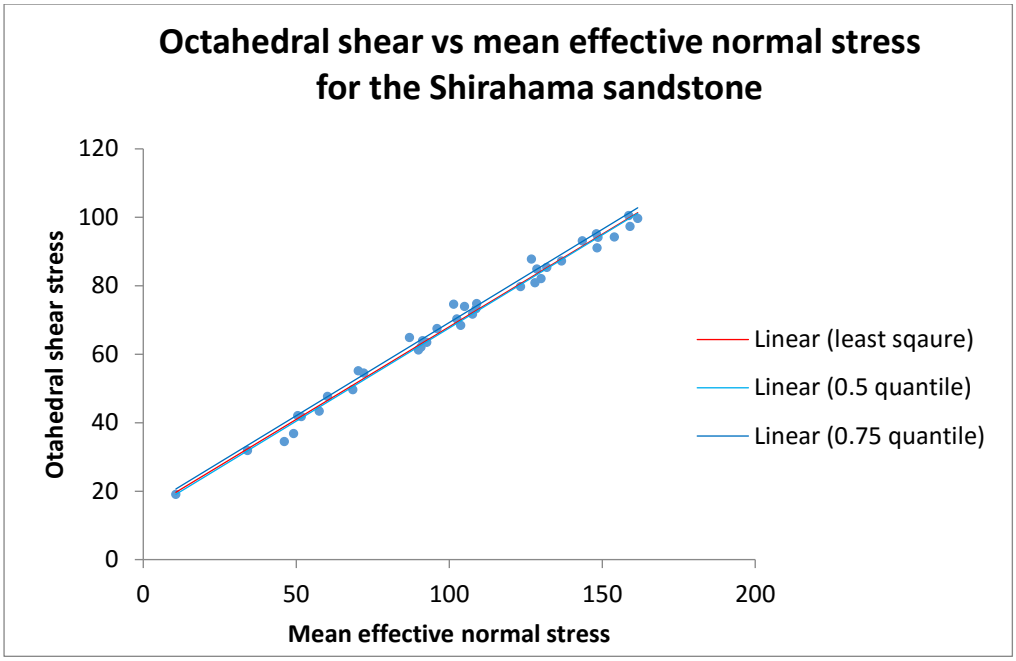

Figure 4.28: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Shirahama Sandstone data.

Table 4.10 shows the various R^2 values for the different models used in the Shirahama sandstone data. The values are all greater than 0.7 suggesting that the models perfectly fit the data. The least square regression model fits the data better as it has the highest coefficient of determination (0.988) while the 0.5 has the lowest value (0.968). Based on R^2, one can conclude that the data behaves in a linear form and can be modelled using linear functions such as the least square and quantile regression.

Table 4.10: R^2 values evaluating the relationship between the observed Shirahama Sandstone data and the fitted models

Model type	R^2
Least square regression	0.988
0.50 quantile	0.968
0.75 quantile	0.982

Failure prediction

According to Figure 4.29, the 0.5 quantile predicts that 12 points are stable; 5 points are resulting to rock failure and indecisive, respectively. On the other hand, the 0.75 quantile model predicts that 10 points are failing, 5 points are stable and 7 are plotting directly on the quantile line. The least square estimates a failure of 15 points, 14 stable and 10 indecisive.

When comparing the indecisive points, quantile regression has more points which cannot be classified (12) compared to the least square regression with 10 points.

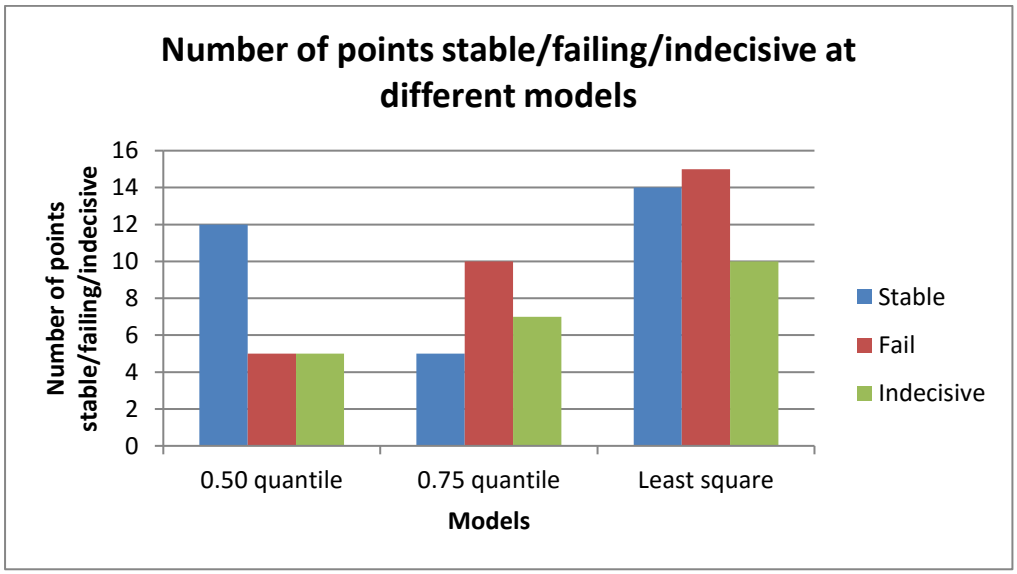

Figure 4.29: The number of points classified as stable/failing/indecisive using both statistical techniques for the Shirahama Sandstone data.

Both the least square regression and quantile regression models have estimated that 15 points are likely to result to rock failure as observed in Figure 4.30.

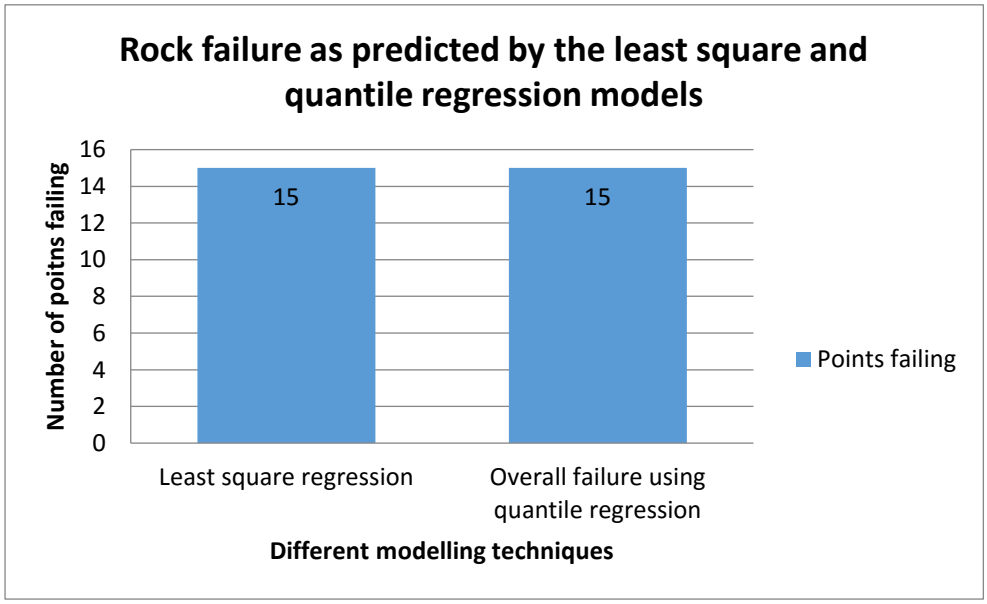

Figure 4.30: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Shirahama Sandstone

4.1.11. Yuubari Shale

Figure 4.31 is a comparison of the regression lines fitted in the Yuubari Shale data. The slopes of all models are the same as they are all parallel. However, the gradient of the least square and 0.5 quantile intersect at a mean effective normal stress of 140 MPa and octahedral shear stress of 175 MPa. The least square line is quite gentle compared to the other regression lines. The regression lines for all three models are quite close to one another.

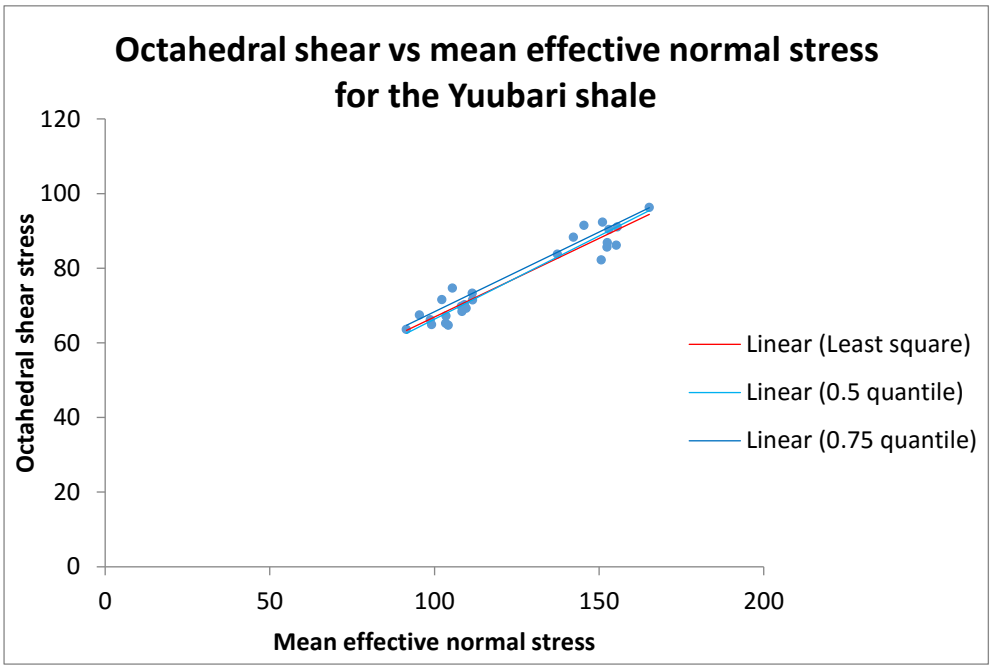

Figure 4.31: The relationship between the octahedral shear and mean effective normal stress modelled using least square and quantile regression for the Yuubari Shale data.

Table 4.11 shows R^2 values for the different quantiles and the least square regression. The R^2 values are all above 0.7. This suggests that the 3 models are a perfect fit for the data and prediction can be done with great confidence.

Table 4.11: R^2 values evaluating the relationship between the observed Yuubari Shale data and the fitted models

Model type	R^2
Least square regression	0.823
0.50 quantile	0.858
0.75 quantile	0.875

Failure prediction

Based on Figure 4.32, 10 points are stable and 4 points are failing in the 0.5 quantile. On the other hand, the 0.75 quantile predicts that 7 points are likely to result to rock failure. The least square regression estimated 13 points to be stable while 10 points will fail. In overall, the majority of the points did not plot directly on the regression line suggesting that this rock has less indecisive points.

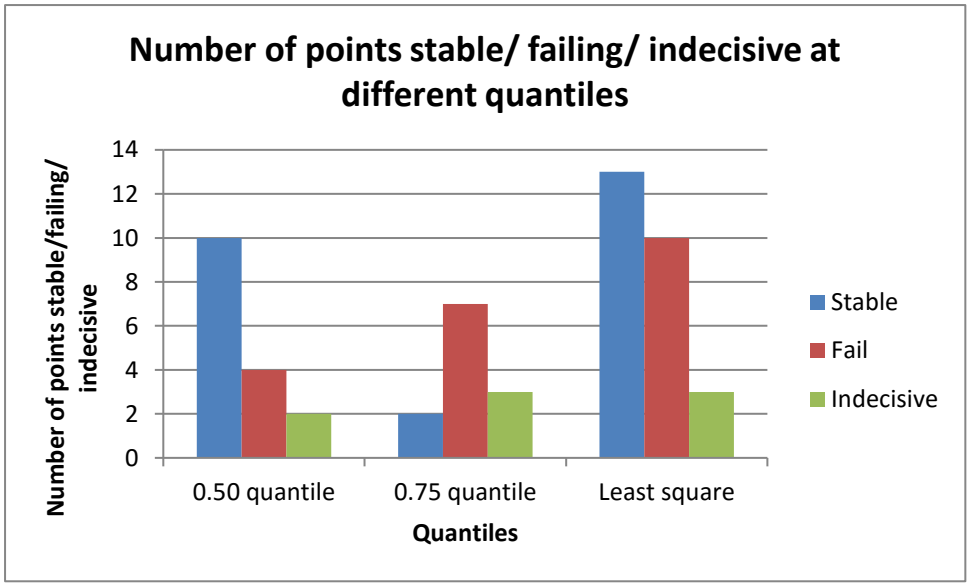

Figure 4.32: The number of points classified as stable/failing/indecisive using both statistical techniques for the Yuubari Shale data.

According to Figure 4.33, the least square regression predicts that 10 points will result to rock failure while the quantile regression estimates 11 points. In terms of percentages, least square predicted 38% failure while the quantile regression predicted 42% failure.

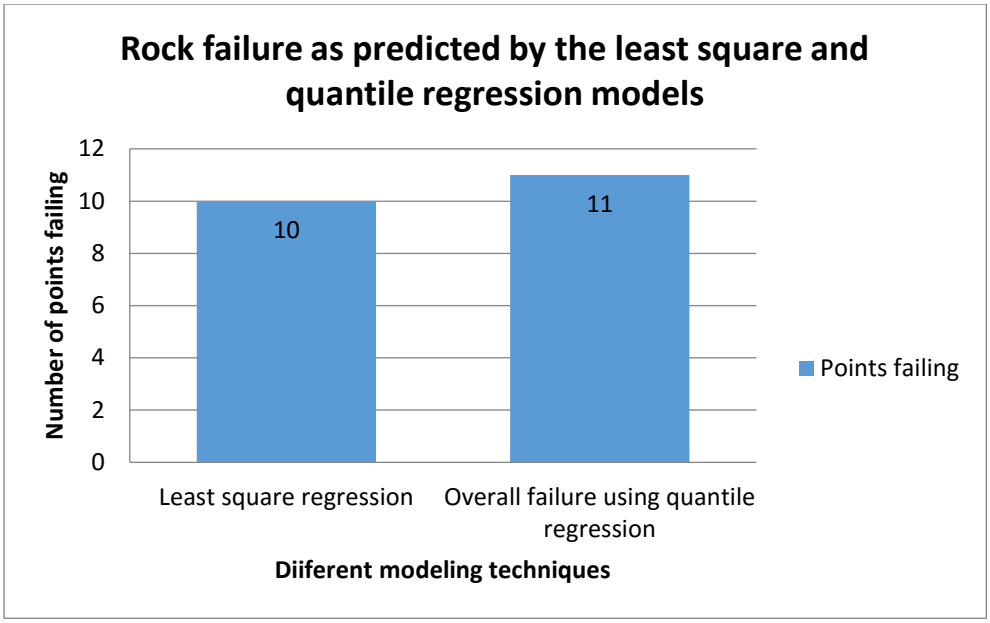

Figure 4.33: A comparison of failure prediction as estimated by the least square regression and quantile regression for the Yuubari Shale.

4.2. Machine learning based techniques

The results from the machine learning based techniques (logistic regression and k-nearest neighbour) are interpreted in this section.

4.2.1. Logistic regression

The logistic regression was conducted at randomly chosen cut-off values to examine its influence on the performance of the model. The cut-off values used were 0.3, 0.5, 0.6 and 0.7. Pertaining to the significance of the variables, the model concluded that σ_2, the mean effective normal stress and octahedral shear stresses were statistically significant hence they were used to predict failure. The results are illustrated in Figure 4.34 and Table 4.12.

The cut-off value has an influence on the prediction of stable or failing points as illustrated in Figure 4.34. From this Figure, it can be observed that a cut-off value of 0.3 results to high number of stable points as compared to the other cut-off values. Similarly, more points are predicted to fail when the 0.7 cut-off value is used. A specific trend exists between the cut-off value and failure prediction i.e. an increase in the cut-off value results to an increase in the failure prediction and a decrease in the estimation of stable points. It should be noted that

in reality, majority of the points were observed to fail suggesting that a high cut-off favours the majority class. It is thus important to specify whether stable or failing points are of interest so that one can decide on the most appropriate cut-off value. From a rock engineering perspective, it is better to predict that more points will cause rock failure rather than estimating fewer points as this can have a dire consequence on the support design.

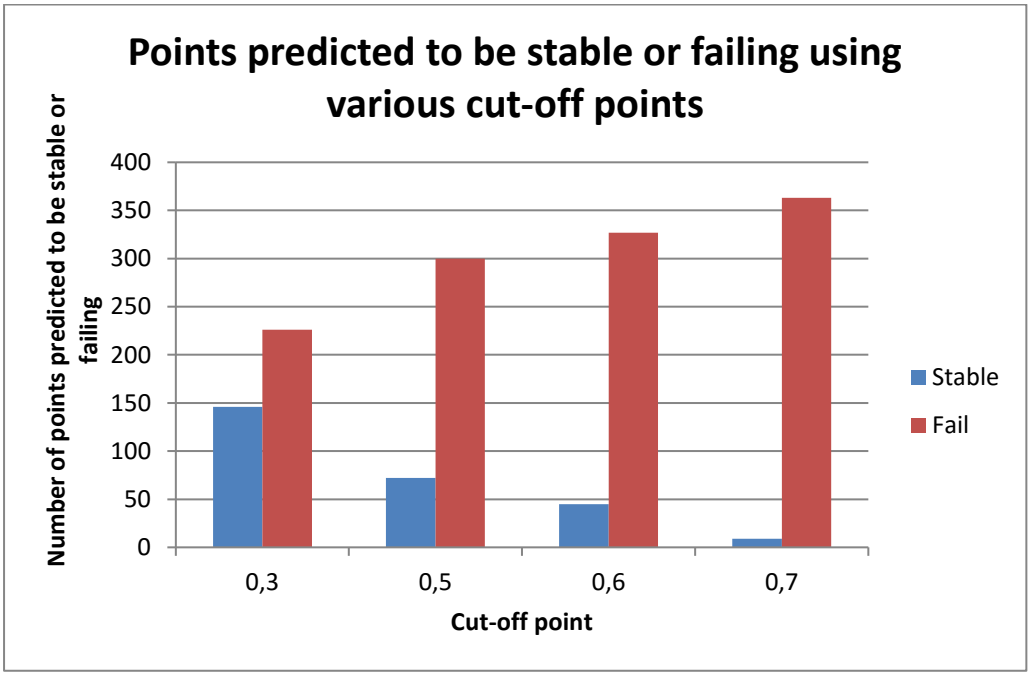

Figure 4.34: Prediction of stable and points resulting to rock failure using logistic regression at different cut-off values.

The results in Table 4.12 are a summary of some of the important parameters in logistic regression. All three principal stresses were modelled however σ_1 and σ_3 had $p>0.05$ (significance level) hence they were considered to be not significant. On the other hand, the p value of σ_2, mean effective normal stress and octahedral shear stress were less than the significance level as observed in Table 4.12 suggesting that there is a statistical significance between the response variable and the three mentioned parameters. The coeffb of the mean effective normal stress is negative while that of the intermediate stress and octahedral shear stress is positive. The influence of this is better explained by the exp (b) parameters which explains the odds of an event occurrence (Peng, et.al. 2002). Odd ratios greater than 1 implies a change in the occurrence of an event as a certain variable changes. According to Table 4.12, the failure points are likely to change when the intermediate stress and octahedral shear stress changes. An increase in these stresses is likely to increase the probability of the points failing. Pertaining to the standard errors, the values are all small implying that if the tests were to be

repeated, the results obtained using these three variables would not show much variation. This suggests that the estimation would be more precise. According to Ranganathan, et.al. (2017) the results from Table 4.12 can be used to formulate an equation providing a model that can be used to estimate the likelihood of an event occurrence given the predictor factors. The equation includes the coeff and given by:

$log_e(odds\ of\ rock\ failure) = 0.932 - 0.042(octahedral\ normal) + 0.042(octahedral\ shear\ ) + 0.007(\sigma_2)$ 19

Table 4.12: Parameters summarising the logistic regression model

	coeff b	*s.e.*	*p-value*	*exp(b)*
Intercept	0.9323	0.2989	0.0018	2.5405
σ_2	0.0071	0.0024	0.0028	1.0071
Mean effective normal stress	-0.0417	0.0088	0.0002	0.9428
Octahedral shear	0.0425	0.0106	0.0064	1.0434

ROC Curve

As stated, for cut-off values were used. The AUC values are similar at all cut-offs, however the prediction, accuracy, sensitivity and specificity are different. The ROC curve (Figure 4.35) is used to evaluate the ability of the model to discriminate between failing and stable points. Since the curve does not form a straight diagonal line, one can have confidence on the model correctly differentiating failing points from the stable ones. From the logistic regression results, the AUC is 0.8202 at all 4 cut-offs. According to Tape (2001), a model with a 0.8 AUC is considered as a good predictive model. This suggests that the logistic regression is good at correctly discriminating between stable and failing points and can be used with confidence in predicting points resulting to rock failure. Figure 4.35 can also be used to determine the best cut-off value (where the errors are minimal or where the sensitivity + specificity are maximal). This is determined using the distance method i.e. the point closer to the left corner of the ROC curve is the optimal cut-off.

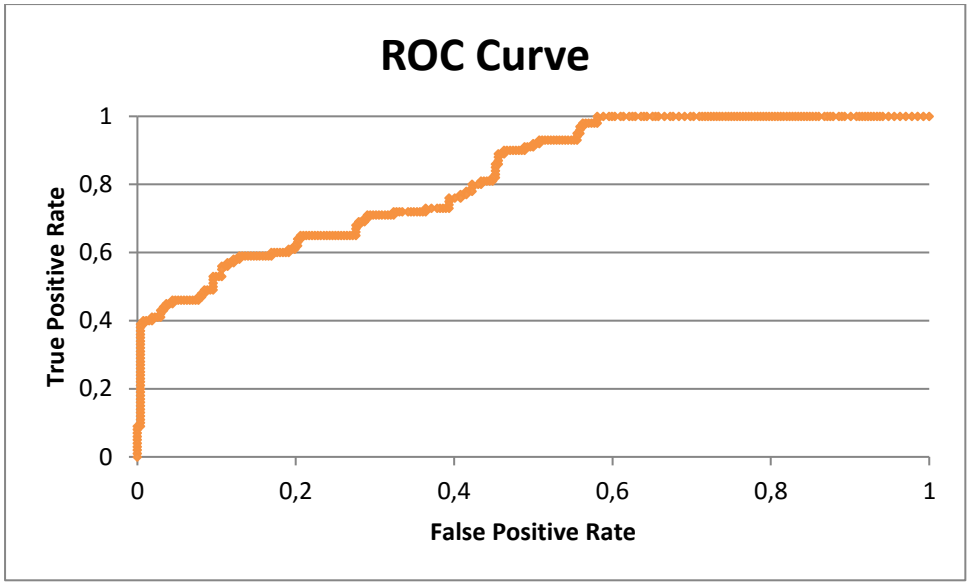

Figure 4.35: AROC curve used to evaluate the logistic regression performance

According to Figure 4.36, the best cut-off point is 0.6. At this cut-off, the combination of the sensitivity and specificity is high. The accuracy is also observed to be at maximum at this cut-off. This suggests that there are minimal classifying errors when running logistic regression at a cut-off of 0.6.

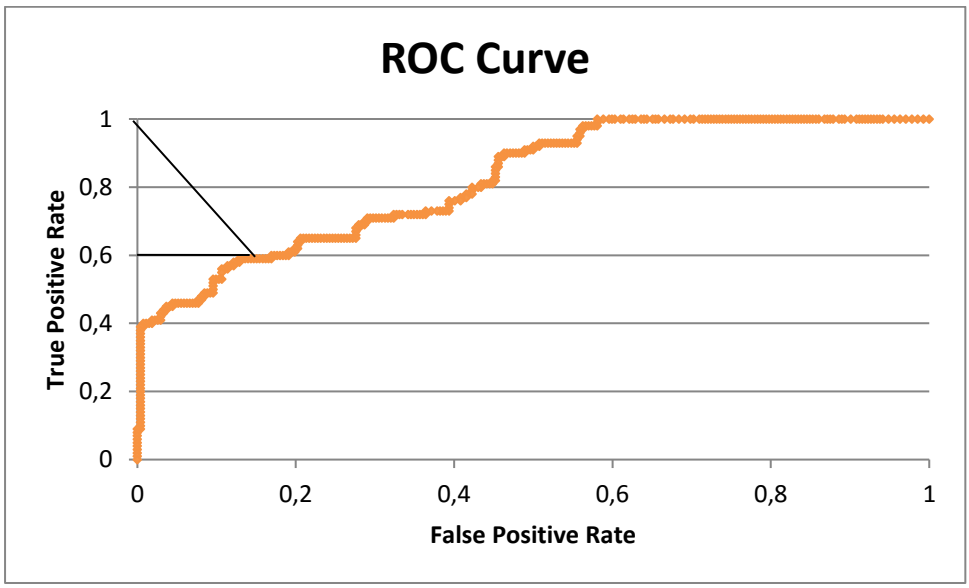

Figure 4.36: ROC curve used to determine the best cut-off value.

4.2.2. K-nearest neighbour

When modelling data using k-nearest neighbour, different parameters can be adjusted to illustrate their influence on the prediction accuracy of the model. The graphs below show the effect of the k value and the split on the performance of the model. Three k values were used (3, 5 and 7) and three split values were evaluated (0.35, 0.67 and 0.85). When the k values were changed, the split was kept constant at 0.67. A split of 0.67 implies that the data is divided into 67% training dataset and 33% test dataset. This simply means that the model was trained using 67% of the data and the performance of the model was tested using 33% of the data. Secondly, the split values were altered while keeping the k value constant as 3.

In Figure 4.37, the split was kept constant as 0.67 while the k values were changed. At a split of 0.67, 248 points were used to train the algorithm while 124 dataset were tested for all three k values. From Figure 4.37, it is observed that the highest points (90) resulting to rock failure were predicted when the k value was 5 and more stable points (38) were predicted when k was 3. The influence of k value is not that visible but does have an effect on the accuracy as explained in the confusion matrix section.

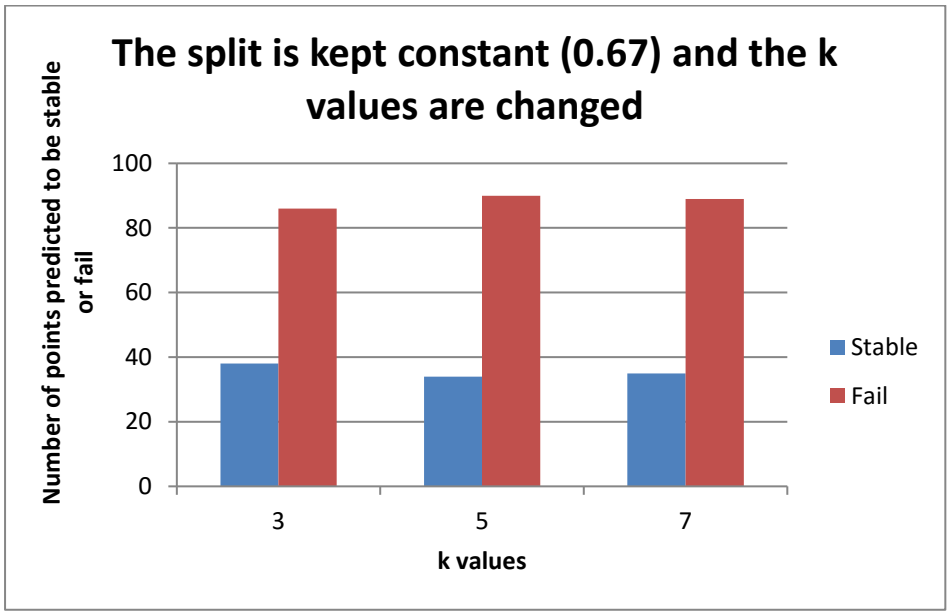

Figure 4.37: The influence of the k-value on k-nearest neighbour modelling

The influence of the split was evaluated by keeping k constant whilst changing the split values. Based on Figure 4.38, it can be deduced that the split determines the amount of data set to be trained and tested by the model. When the split was 0.35, 243 points were tested. Out of the 243, 178 points were predicted to fail while 65 were classified as stable. When the

split was increased to 0.67, 124 points were tested. From the 124 points, 86 were predicted to fail and 38 were stable. With a further increase of the split (0.85), 56 points were tested with 37 points classified as failing and 19 as stable. This has an influence on the accuracy of the model.

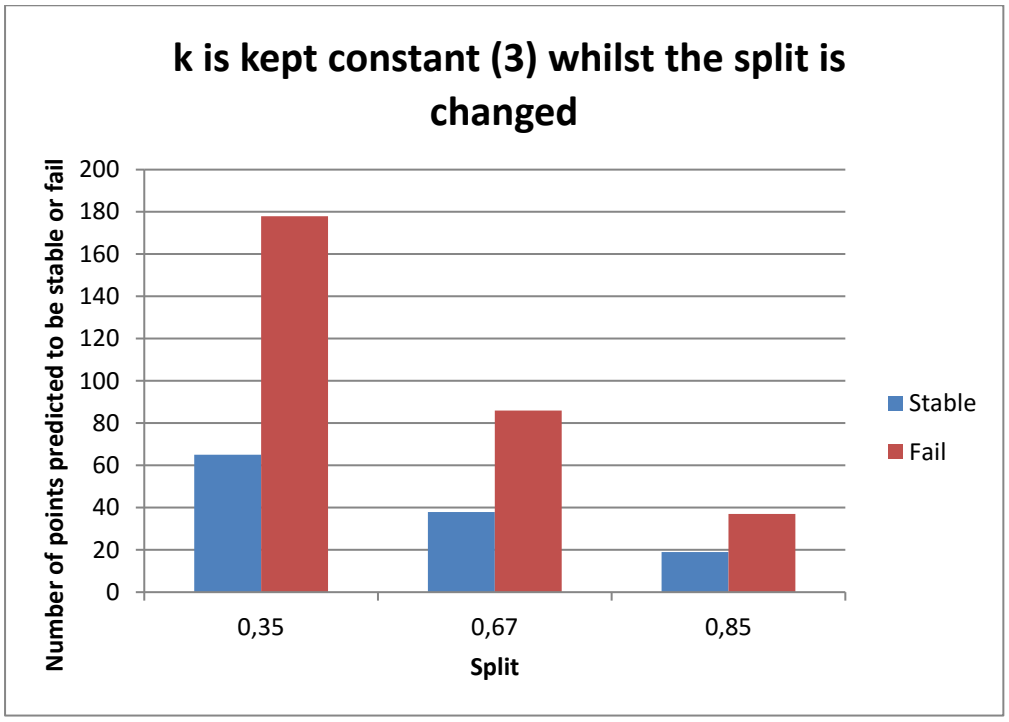

Figure 4.38: The influence of the split on k-nearest neighbour modelling

4.4. Confusion matrix

A confusion matrix is a system used to evaluate the performance of a modelling technique. The modelling techniques were evaluated using this system as illustrated by the following tables. Table 4.13 shows the total number of points tested under true triaxial loading for each rock. These points were then classified either as stable or failing using the definition of rock failure proposed by Mogi i.e. "Failure will occur when the distortional energy reaches a critical value that increases monotonically with the effective mean pressure on the slip planes parallel to σ_2" (Colmenares and Zoback, 2002, 700p). An important parameter is the true triaxial strength which is provided by Kwasniewski, et.al. (2012). The true triaxial strength of the Yuubari Shale is not known; hence it was impossible to model data of this rock using machine learning techniques. This implies that rock failure prediction for the Yuubari Shale could only be conducted using statistical techniques thus the two machine learning techniques

only modelled 372 points instead of 398 points (excluded Yuubari Shale). In a perfect scenario, the models should have predicted as Table 4.13, however that was not the case due to various factors. The following tables are representation of what was predicted by each model.

Table 4.13: A table showing the true classification of points based on Mogi's failure criterion

Rock type	**Stable**	**Fail**	**Total points tested**
Solnhofen Limestone	31	0	31
Dunham Dolomite	24	30	54
Yamaguchi Marble	8	18	26
Mizuho trachyte	2	29	31
Orikabe Monzonite	3	35	38
Manazuru Andesite	1	22	23
Inada Granite	4	40	44
KTB Amphibolite	5	35	40
Westerly Granite	12	34	46
Shirahama Sandstone	11	28	39
Yuubari Shale	N/A	N/A	26

Least square regression

Out of 398 points, the least square regression could classify 259 points as either stable or failing. The remaining 139 points were indecisive. According to Table 4.14, a total of 178 points were observed as failing. Out of the 178 points, 105 were correctly predicted to fail whilst 73 were incorrectly predicted to be stable when they are actually failing. Also 81 points were observed to be stable however only 46 were correctly predicted to be stable whilst 35 were misclassified as failing. In total, 108 points were misclassified whilst 151 were correctly classified. The accuracy of this model in correctly predicting and classifying failing and stable points is 60%.

Table 4.14: Confusion matrix for the least square regression model

	Predicted stable	Predicted fail
Observed stable	46	35
Observed failing	73	105

Accuracy: 60 %

Quantile regression

From Table 4.15, it can be observed that a total of 106 points were observed to be failing. Within the 106 points, 25 were correctly predicted as failing while 81 were predicted to be stable whereas they are actually failing. Also 13 points were incorrectly predicted to fail and 40 were correctly predicted to be stable. The accuracy of the 0.5 quantile regression in correctly predicting failure is 41%.

Table 4.15: Confusion matrix for a 0.50 quantile model

	Predicted stable	Predicted fail
Observed stable	40	13
Observed failing	81	25

Accuracy: 41%

In the upper quantile (0.75), 76 points were observed to result to rock failure whilst 24 were stable. Out of the 76 failing points, 51 were correctly predicted to fail and 25 were incorrectly classified as stable (Table 4.16). Pertaining to stable points, 13 were correctly classified as stable and 11 were incorrectly classified as points resulting to failure. In overall, 36 points were incorrectly classified whilst 64 were correctly classified. The accuracy of the 0.75 quantile to correctly predict failure is 64%.

Table 4.16: Confusion matrix for a 0.75 quantile model

	Predicted stable	Predicted fail
Observed stable	13	11
Observed failing	25	51

Accuracy: 64%

Logistic regression

In reality, 271 points were observed to fail and 101 were stable (excluding the Yuubari Shale data). When modelling at a cutoff value of 0.3 as shown in Table 4.17, the model predicted 226 points to fail whilst 146 were estimated to be stable as observed in Table 4.17. From the 226 failing points, 195 were correctly predicted to fail by the model and 77 were

misclassified as stable whereas they are failing. On the other hand, 69 points were correctly predicted to be stable and 31 were misclassified as failing whereas they were stable. The accuracy of the logistic regression at a 0.3 cutoff value is 71%.

Table 4.17: Confusion matrix for a Logistic regression model at 0.3 cutoff

	Predicted stable	Predicted fail
Observed stable	69	31
Observed failing	77	195

Accuracy: 71 %

When the cut-off point was increased, the accuracy of the model also increased as it is observed that an accuracy of 80% was achieved at a cutoff of 0.5. Out of the 271 points observed to fail, 249 were correctly predicted as failing whilst 23 were incorrectly predicted to be stable. Pertaining to stable points, 49 points were predicted to be stable and 51 were predicted to fail although they were stable. According to Table 4.18, a total of 74 points were misclassified.

Table 4.18: Confusion matrix for a logistic regression model at 0.5 cutoff.

	Predicted stable	Predicted fail
Observed stable	49	51
Observed failing	23	249

Accuracy: 80%

When the cutoff point was further increased to 0.6, maximum accuracy of 83% was achieved. Out of a total of 271 observed failing points, 267 were correctly predicted as failing and 5 points were predicted as stable whereas they are failing. Giving reference to the 101 stable points, 40 points were correctly predicted as stable whilst 60 points were misclassified to be failing whereas they are stable. At a cut-off value of 0.6, a total of 65 points were misclassified as shown in the Table 4.19.

Table 4.19: Confusion matrix for a logistic regression model at cutoff-0.6 "optimum"

	Predicted stable	Predicted fail

Observed stable	40	60
Observed fail	5	267

Accuracy: 83%

With a cutoff value of 0.7, all 272 points were correctly predicted to fail whilst no point was misclassified as stable. On the other hand, 91 stable points were classified as failing whereas they are stable and only 9 were correctly predicted as stable as observed in Table 4.20. The accuracy of the logistic regression at a cutoff value of 0.7 is 75%.

Table 4.20: Confusion matrix for a logistic regression model at 0.7 cut-off

	Predicted stable	Predicted fail
Observed stable	9	91
Observed failing	0	272

Accuracy: 75%

K-nearest neighbour

When the split was 0.67, the number of tested dataset points was 124 whilst the trained dataset points were 248. Out of the 124 points, 96 points were observed to fail whilst 28 were stable. From Table 4.21, 82 were correctly predicted to fail while 14 were incorrectly classified as stable. From the 28 stable points, 4 were incorrectly predicted as stable and 24 were correctly classified. In total, 18 points were incorrectly classified. The accuracy of the k-nearest neighbour to correctly predict failed and stable points is 85%.

Table 4.21: Confusion matrix for a k-nearest neighbour model with k=3, split=0.67

	Predicted stable	Predicted fail
Observed stable	24	4
Observed failing	14	82

Accuracy: 85%

Keeping the split constant, the k value was increased to 5. Giving reference to Table 4.22, it is observed that 97 points were observed to fail and 27 were stable. Out of the 97 points, 83

were correctly classified as failing and 14 were incorrectly predicted to be stable. Also 20 points were correctly predicted to be stable whilst 7 were incorrectly classified as failing. In total 21 points were misclassified thus lowering the accuracy of the model to 83%.

Table 4.22: Confusion matrix for a k-nearest neighbour model at k=5, split=0.67

	Predicted stable	Predicted fail
Observed stable	20	7
Observed failing	14	83

Accuracy: 83%

When the k value was further increased to 7, a decrease was observed on the accuracy of the model. A total of 96 points were observed to fail while 28 were stable. From the 96 failing points, 81 points were correctly predicted as failing whilst 15 points were incorrectly predicted as stable whereas they are failing. Pertaining to the 28 stable points, 20 were correctly predicted as stable and 8 were incorrectly predicted to fail. The accuracy of the k-nearest neighbour when k =7 is 81.5%.

Table 4.23: Confusion matrix for a k-nearest neighbour model at k=7, split=0.67

	Predicted stable	Predicted fail
Observed stable	20	8
Observed failing	15	81

Accuracy: 81.5%

Since an optimal accuracy is achieved when the k value is 3, it was further used to check the influence of the split on accuracy. The split was lowered to 0.35 and the results are observed in Table 4.24. With a split of 0.35, 129 points were used to train the model while 243 points were used as test the datasets. From the 243 points, 205 were observed to fail while 38 were stable. Out of the 205 failing points, the model predicted 168 failing points and misclassified 37 to be stable as observed in Table 4.24. Similarly from the 38 stable points, 28 were correctly classified and 10 where misclassified. In total, 47 points were incorrectly classified. The accuracy of the model decreased from 85% to 80%. This suggests that the higher the number of test data, the lower the prediction accuracy.

Table 4.24: Confusion matrix for a k-nearest neighbour at k=3, split=0.35

	Predicted stable	Predicted fail
Observed stable	28	10
Observed failing	37	168

Accuracy: 80.8%

A split of 0.85 implies that 85% of the data is used to train the model while 15% is used to test the model. Hence 56 points were used to test the capability of the model to predict failing /stable points. From the 56 points, 45 points were observed to fail whilst 11 were stable. Out of the 45 failing points, 37 were correctly predicted to fail and 8 points were misclassified as shown in Table 4.25. In terms of stable points, all 11 points were correctly predicted to be stable suggesting a 100% prediction. The accuracy of the model is 86%. This implies that an increase in the split decreases the amount of data to be tested thus modelling becomes more accurate.

Table 4.25: Confusion matrix for a k-nearest neighbour at k=3, split=0.85

	Predicted stable	Predicted fail
Observed stable	11	0
Observed failing	8	37

Accuracy: 86%

Sensitivity and specificity

As stated in the literature review, sensitivity and specificity are two critical parameters when evaluating the ability of a model to correctly classify points. According to Table 4.26, it is observed that the level of sensitivity and specificity differs per model. Also the variation between the amount of sensitivity and specificity is not the same. Both the sensitivity and specificity for k-nearest neighbour when k=3, split=0.67 is 85, respectively implying that the sensitivity and specificity of the model is the same. For least square regression and logistic regression at 0.3 cut-off, the difference between the sensitivity and specificity is very small whilst a greater variation is observed for logistic regression at cutoff of 0.7 and 0.5. Furthermore, the specificity for the 0.5 quantile and the k-nearest neighbour (k=3, split=0.85) is greater than the sensitivity whilst an opposite trend is observed for the other models as seen

in Table 4.26. Logistic regression (0.7 cutoff) has the maximum sensitivity (100) and k-nearest neighbour (k=3, split=0.85) has the maximum specificity (100). Pertaining to the sum of the two parameters, k-nearest neighbour has the greatest combination (180) and the 0.5 quantile shows the least (99) combination. The sum of the sensitivity and specificity has an influence on the accuracy of the model.

Table 4.26: Sensitivity /specificity for all models

Predictive model	Sensitivity/Specificity	
Least square regression	Sensitivity	59
	Specificity	57
0.50 Quantile	Sensitivity	24
	Specificity	75
0.75 Quantile	Sensitivity	67
	Specificity	54
Logistic regression (0.3 cutoff)	Sensitivity	72
	Specificity	69
Logistic regression (0.5 cutoff)	Sensitivity	92
	Specificity	49
Logistic regression (0.7 cutoff)	Sensitivity	100
	Specificity	9
Logistic regression (Best cutoff 0.6)	Sensitivity	98
	Specificity	40
K-nearest neighbour (k=0.3, split=0.67)	Sensitivity	85
	Specificity	85
K-nearest neighbour (k=5, split=0.67)	Sensitivity	86

	Specificity	74
K-nearest neighbour (k=7, split=0.67)	Sensitivity	84
	Specificity	71
K-nearest neighbour (k=3, split=0.35)	Sensitivity	82
	Specificity	74
K-nearest neighbour(k=3, split=0.85)	Sensitivity	82
	Specificity	100

4.5. Conclusions

Generally, graphical representations are used to analyse statistical techniques. The concept of such representation was applied for least square regression and quantile regression. Since quantile regression required multiple quantiles, three quantiles were randomly chosen i.e. the 0.25, 0.5 and 0.75 quantiles. Based on the nature of the data, modelling of the 0.25 quantile was impossible hence the results were based on the 0.5 and 0.75 quantiles. From the graphical results together with R^2 values, it can be concluded that the two quantiles and the least square regression perfectly fits the data. However, the use of two quantiles results to majority of the points plotting directly on the regression lines thus producing a high number of indecisive points (neither stable nor failing). In overall, it was observed that the least square regression predicted more points to fail in comparison to the combined two quantiles. The two machine learning based techniques i.e. logistic regression and k-nearest neighbour were analysed individually. Logistic regression was modelled at four cutoff values (0.3, 0.5, 0.6 and 0.7) which were randomly chosen. Based on the variation of the results from the four cutoff values, it was concluded that the cutoff value has an influence on the prediction and classification of stable and failing points. Furthermore, it was observed that the amount of variables used at the initial stages of modelling has an influence on the end product. Results from k-nearest neighbour illustrated the influence of the k value and the split on the prediction accuracy. Three random k values were chosen (3, 5 and 7). Also three split values

(0.35, 0.7 and 0.85) were used. The lower split implied that a lower dataset was used to train the algorithm whilst a higher split suggests that a higher dataset was trained and few data points were used to test the algorithm. It was then concluded that this has an influence on the predictive accuracy of the model. The predictive accuracy of all four models was evaluated using a confusion matrix. Based on this system, k-nearest neighbour had the highest predictive accuracy while the 0.5 quantile had the least. The sensitivity and specificity of each model was also analysed based on the confusion matrix theory. The higher the combination of the sensitivity and specificity, the better the predictive performance of the model. K-nearest neighbour had the highest sensitivity and specificity combination whilst quantile regression had the least. This corresponds to the predictive accuracies observed.

CHAPTER 5: DISCUSSION

5.1.1 Statistical Techniques

Least square regression and Quantile regression

The distribution of the data has an influence on the manner in which the data will be model. As it was observed, the true triaxial data for almost all 11 rocks followed a linear path. Hence modelling using the 0.25 quantile was impossible. This implies that the quantile regression is quite sensitive to the distribution of the data. For much better results, scattered data should be used.

Gradient

According to Schneider et.al. (2010), the gradient in a regression model describes the change in the dependent variable per unit change in the independent variable. In the context of this research, it explains the rate of change of the octahedral shear stress with respect to the mean effective normal stress. The models with parallel regression lines have equal gradients suggesting that the rate of change of the octahedral shear stress is the same. On the other hand, the gradients of non-parallel lines are not equal meaning the rate of change varies. In overall, not much variation was observed amongst the gradients of the three models. However, there was a variation in the slopes of some rocks such as the Solnhofen Limestone. The observed variation in the gradients suggests that the data behaves differently throughout. It is then not ideal to model the entire data using as single line of best fit, which is the principle followed by the least square regression.

Coefficient of determination (R^2)

As stated in the literature review, the relationship between the observed data and fitted model is expressed by the coefficient of determination (R^2). In overall, the R^2 values for both least square regression and quantile regression (0.50 and 0.75) were greater than 0.7 for all eleven rocks. According to Moore, et.al. (2013), R^2 values greater than 0.7 suggest a strong perfect fit. With R^2 values of this range, it can be deduced that the data behaves in a linear form and the variation between the observed data and predicted values is small. This is supported by Frost (2013) where he explains that when the difference between the observed values and predicted values is small, that model is considered to fit perfectly

It was however observed that the rocks with the highest range in terms of X axis show an increase in the R^2 compared to those with shorter range. A comparison is made using the Orikabe Monzonite and the Yuubari Shale rocks. The range of the Orikabe Monzonite is 700-78=622 while that of the Yuubari Shale is 152-68=84. When comparing the R^2 values of the two rocks, the one with the highest range had the highest R^2values as compared to the Yuubari Shale with the lowest range. In 1986, Cornell and Berge conducted a study with the aim of identifying the factors affecting the R^2 values. From this study, one of the factors was the increase in the range of the X values. They concluded that R^2 increases with an increase in the range of X values and their conclusion was based on Figure 5.1. This supports the author's observations.

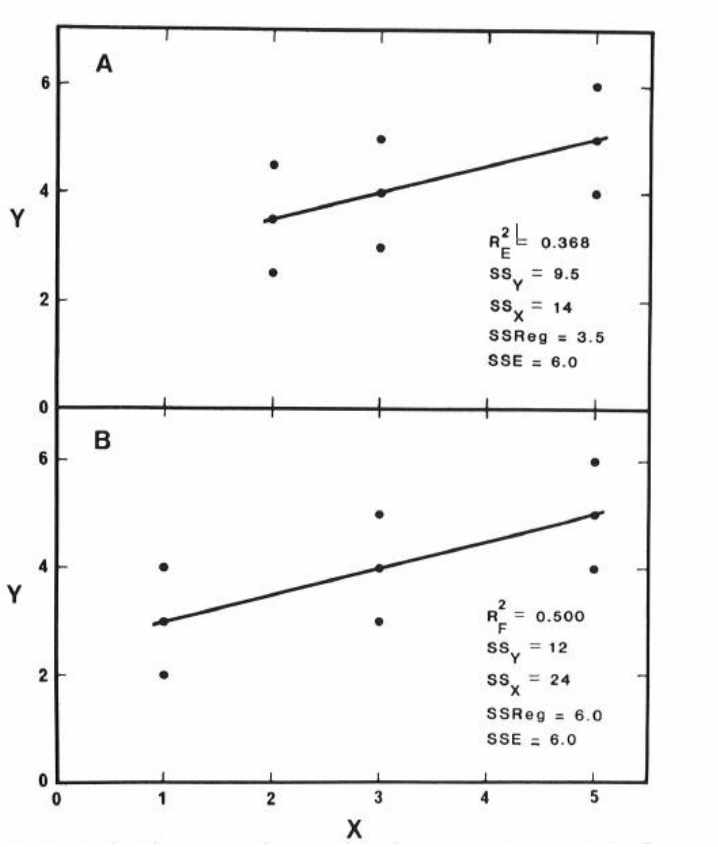

Figure 5.1: Two graphs illustrating an increase in R^2 due to an increase in the range of X values (Cornell and Berger, 1986).

Furthermore, it was noted that the closer the points are to the fitted line of a model, the higher the value of R^2. This implies that R^2 decreases when the points are further from the regression line. In simple terms, R^2 parameter is affected by outliers i.e. points that are outside the central range. The Solnhofen Limestone data shows the influence of outliers on the R^2 value. In this data, two points are identified as outliers i.e. the (43:18 and 70: 74) points as observed

in Figure 5.2a. The least square regression was affected by the two outliers as the regression lines moves towards the two points with the aim of accounting for the points. In the presence of the two outliers (Fig: 5.2a), the R^2 value for the least square regression model is 0.933. When the outliers are emitted as shown in Figure 5.2b, the R^2 value is observed to increase to 0.955. The influence of outliers was identified by several authors including Hao and Naiman (2007). These two authors concluded that conditional mean techniques do not extend to locations that are not within the centre and thus are strongly affected by outliers.

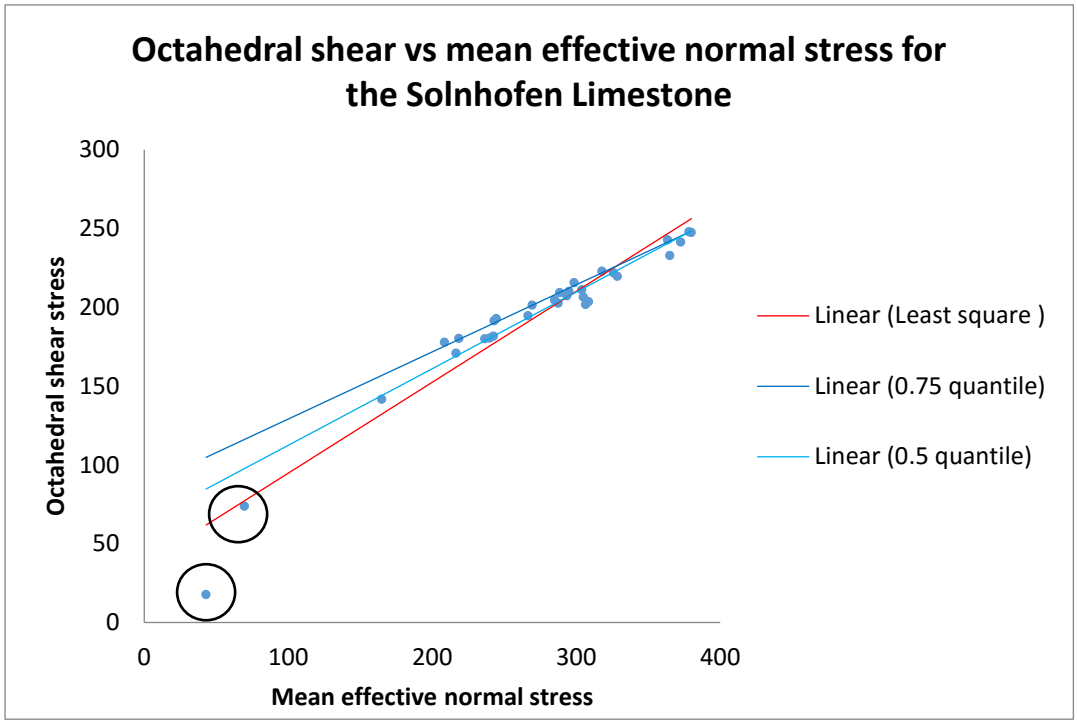

Figure 5.2a: The modelling of the Solnhofen Limestone data with the two outliers included

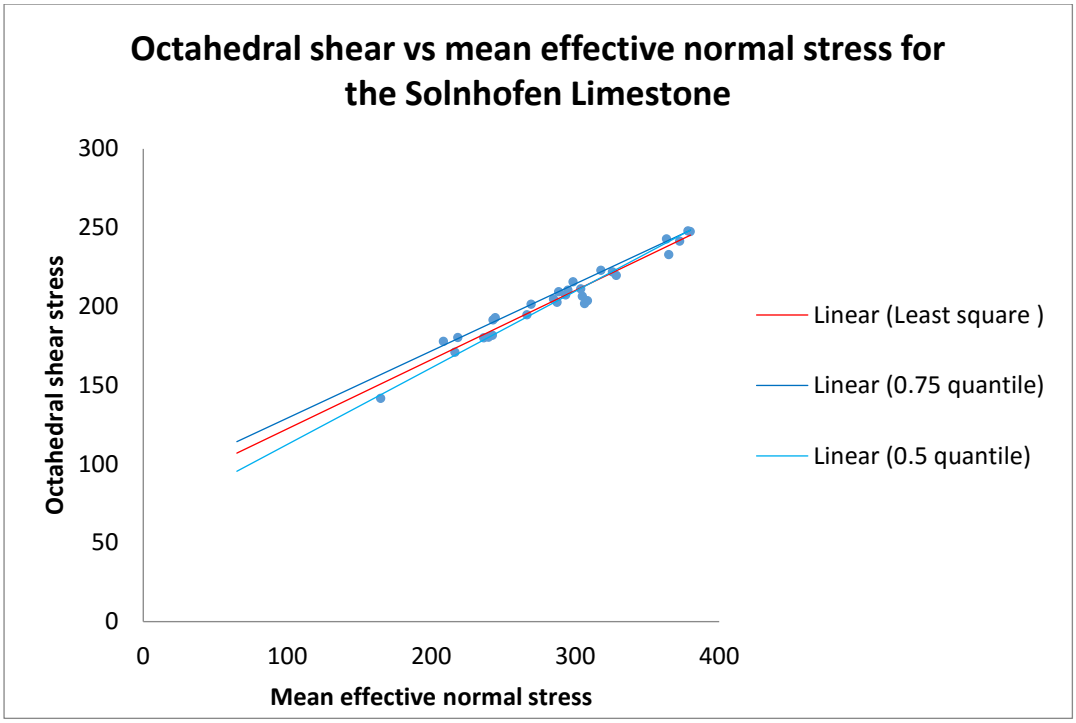

Figure 5.2b: The modelling of the Solnhofen Limestone data without the two outliers

Indecisive points

The outcomes of a rock failure prediction should either be stable points or points resulting to failure. This is true when using logistic regression as it models binary dichotomous data; hence all 372 points were predicted and classified as stable or failing. However, statistics based techniques (least square and quantile regression) resulted to three outcomes i.e. stable, failing and indecisive. Out of 398 points (including the Yuubari Shale), 259 points could be predicted and classified by the least square regression whilst 223 were predicted by the quantile regression. This was expected since the quantile regression models in quantiles thus more regressions lines and more points plotting exactly on the line. This however had an influence on the accuracy of the model.

Out of the 11 rocks modelled using least square regression, only Orikabe Monzonite and Inada Granite had quite a few indecisive points (34% for Orikabe Monzonite and 45% for Inada Granite). On the other hand, more than 40% of the points for Yamaguchi Marble, Inada Granite, Mizuho Trachyte, KTB Amphibolite, Westerly Granite and Shirahama Sandstone were indecisive when quantile regression was used. Hence out of 398 points (including Yuubari Shale), 223 points were predicted and classified whilst the remaining 175 points were neither stable nor failing. When comparing the failure prediction of the two models, it can be observed that the difference is greater. Such great difference might be accounted for

by the indecisive points. This suggests that an increase in indecisive points increases the error estimation of the model and thus affects the accuracy as it was observed.

5.1.2 Machine learning based methods

In machine learning methods, each technique is influenced by unique parameters. For example, logistic regression is highly affected by the cutoff value while k-nearest neighbour is influenced by the k value, distance type and the ratio of training dataset to testing dataset (split).

Logistic regression

According to Ranganathan, et.al. (2017), the most important aspect in logistic regression is to choose the right variables to use when modelling. Using many variables tends to dilute the association thus resulting to large standard error. For this research, 5 variables were likely to be predictive variables (σ_1, σ_2, σ_3, mean effective normal stress and octahedral shear stress). Table 5.1 shows the outcome of modelling data using all five variables.

Table 5.1: Evaluation of the logistic regression when all 5 variables are used.

	coeff b	*s.e.*	*p-value*	*exp(b)*
Intercept	0	0.2468	1	1
σ_1	0	30269.78	1	1
σ_2	0	0.0023	1	1
σ_3	0	30269.78	1	1
Octahedral normal	0	60539.56	1	1
Octahedral shear	0	0.0155	1	1

Based on Table 5.1, all variables were not statistically significant since the p values were greater than the significance level of 0.05. Also the standard errors were significantly high. This agrees with what was stated by Ranganathan et.al. (2017) that including many variables dilutes the association and increases the standard error. Hence the variables were assessed individually and based on the p value; the σ_1 and σ_3 were eliminated. It should be noted that the p values change when each variable is assessed individually as seen in Table 5.2 below. These are the statistically significant variables according to the logistic regression model. Eliminating σ_1 and σ_3 seems to contradict with observations made by Li et.al. (2002) as he

discovered that the major principal stress at failure is linked to the intermediate principal stress under constant minimum principal stress. This implies that all principal stresses should be statistically significant when predicting rock failure.

Table 5.2: Evaluation of the performance of logistic regression using the statistically significant variables

	coeff b	*s.e.*	*p-value*	*exp(b)*
Intercept	0.9323	0.2989	0.0018	2.5405
σ_2	0.0071	0.0024	0.0028	1.0071
Octahedral normal	-0.0417	0.0088	0.0002	0.9428
Octahedral shear	0.0425	0.0106	0.0064	1.0434

The equation of the odds of rock failure given in Equation 20 is explained as follows: the odds of rock failure or probability of rock failure is likely to increase with an increase in the octahedral shear stress and σ_2 and a decrease in the mean effective normal stress.

$$\log_e(odds\ of\ rock\ failure) = 0.9323 - 0.0417(mean\ effective\ normal\ ) + 0.0425(octahedral\ shear\ ) + 0.0071(\sigma_1) \quad \text{.....20}$$

This equation contradicts with work done by most authors as it was shown that rock failure increases with an increase in the mean effective normal stress.

Another important parameter in logistic regression is the cutoff value. As observed from the results, the accuracy of prediction increases with an increase in the cutoff value until the optimum cutoff is reached. Beyond the optimum cutoff value, the accuracy decreases again. This implies that with no knowledge on the best cutoff, the maximum accuracy will not be obtained. When comparing results obtained from the three cutoff values, it can be observed that the 0.7 cut off value favours the dominating outcome i.e. failing points rather than stable meanwhile the 0.3 cutoff favours the least dominating outcome. This creates biasness in the prediction performance of the model. For example, when the cutoff point was 0.7, the accuracy of predicted failing points was 100% while that of predicting stable points was very low. Also, when the cutoff value was 0.3, the accuracy of estimating stable points was higher compared to that of predicting failing point. This biasness affects the overall accuracy of the model hence it is significant to use a cutoff that gives minimum error and true accuracy. Although the data was modelled at different cutoffs, critical parameters such as coeffb and

standard error remained the same. This suggests that the cutoff only changes the classification of points i.e. predicted accurately stable or failing points. The effect is thoroughly observed in the confusion matrix and affects the sensitivity and specificity of the model.

K-nearest neighbour

The performance of the k-nearest neighbour model is influenced by the k value, type of distance used and data pre-processing methods (split). This research explored the influence of the k value and data pre-processing methods by using different k values and split when testing the data. According to Ma et.al. (2014), there is no clear guideline for determining the k value. It can be any number however; this affects the predictive performance of the model. From the results obtained, it was observed that a smaller k value gave the highest accuracy, while a larger k value lowered the accuracy of the model. The difference in accuracy is explained using the noise concepts. According to Ma, et.al. (2014) a smaller k value is overly sensitive to noise points whilst a larger k includes points from other classes thus misplacing points and lowering the accuracy. Another important parameter in k-nearest neighbour is the split (ratio between trained dataset and tested dataset). Ma, et.al. (2014) explains that in statistics, splitting of dataset into training and testing set is done randomly while machine learning requires an individual to decide on the split ratio. Based on the results, a lower split (0.35) implied that the model was trained using few points hence more errors were observed when the model was tested. A split of 0.85 on the other hand suggested that more points were used to train the model and fewer points were tested hence the accuracy was high. From a study conducted by Ma, et.al. (2014), it was observed that a lower training dataset reduces the accuracy whilst a higher training dataset increases the accuracy. The results thus agree with Ma, et.al. (2014) findings. It was then concluded that a larger split results to the accuracy approaching the optimum value. This however creates biasness. For example, when the split is 0.3, it simply suggests that 30% of the dataset was used for training the model whilst 70% of the data was tested. On the other hand, a split of 0.85 implies that 85% of the dataset was used as training data while 15% of the data was tested. It is fair enough to say the latter scenario will definitely have a higher accuracy compared to the 0.3 split case. A split of 50% should give a true reflection of the performance of the model as observed in Table 5.3. When the split is 0.5, 185 points are used as training dataset whilst 187 are tested dataset. The accuracy of the model is 82% suggesting that the model can accurately discriminate between failing and stable points as shown in Table 5.3.

Table 5.3: Confusion matrix for a k-nearest neighbour model at k=0.3, Split=0.5

	Predicted stable	Predicted failing
Actual stable	20	21
Actual failing	9	137

Accuracy: 82%

Confusion matrix

The confusion matrix explains the accuracy, specificity and sensitivity of the model. Out of all models, the k-nearest neighbour had the highest accuracy despite changes in k value and split as shown in Figure 5.4. The predictive performance accuracy of this model ranged between 80.8% and 86%. On the other hand, quantile regression had the least average accuracy (53%). The predictive accuracies for all the models are shown in the Figure 5.4. When plotting the graph, the maximum accuracy was used for logistic regression and k-nearest neighbour and the average was used for the two quantiles. Pertaining to the sensitivity and specificity of the models, the logistic regression at 0.7 cutoff had the maximum sensitivity and least specificity. According to a medical study conducted by Parikh (2008), it was concluded that a lower sensitivity is likely to overlook the disease. In the context of this research, a lower sensitivity implies that the model is likely to overlook points that can result to rock failure. As stated in the literature review, a high sensitivity lowers the specificity hence the specificity at 0.7 cutoff was minimum. This suggests that the logistic regression at a 0.7 cutoff cannot discriminate stable points from failing points. It is thus important to specify the parameter of interest at the initial stages i.e. points resulting to rock failure or stable points before modelling data. Giving reference to the sensitivity and specificity of k-nearest neighbour when k=0.3, split= 0.67, both parameters are equal (85). This suggests that the model is good at predicting both stable and failing points. When the sensitivity and specificity were combined, k-nearest neighbour had the highest combination whilst quantile regression had the lowest. The highest combination was observed when the accuracy was at maximum (86%), and the lowest combination was at 41% accuracy. It can then be concluded that sensitivity and specificity are a measure of accuracy.

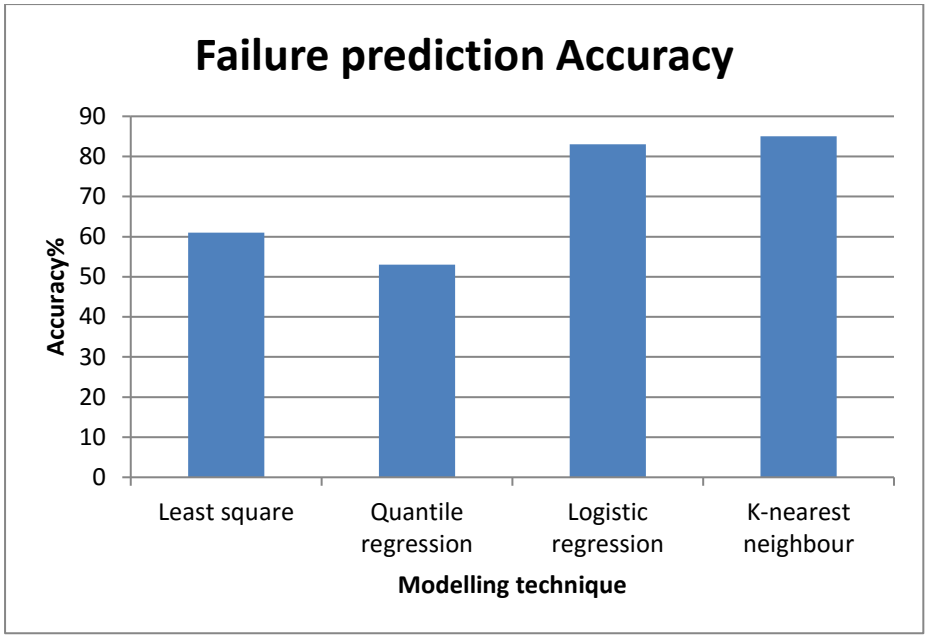

Figure 5.4: predictive performance accuracy of all four models

Robustness

From the four modelling techniques, only least square regression was non-robust. It is not robust to assumptions and outliers as it was observed from the Solnhofen Limestone results. This has an influence on the predictive performance of the model as well as the accuracy.

Bias vs. variance trade-off

As stated in the literature review, quantile regression and k-nearest neighbour are non-parametric techniques whilst least square regression and logistic regression are parametric. According to Brownlee (2016), parametric techniques tend to show a high biasness while non-parametric techniques have a higher variance. A high biasness is caused by an increase in the assumptions made by the model whilst a high variance is as a result of high specifics made by the trained dataset (Brownlee, 2016). The least square regression and logistic regression are examples of high bias-low variance techniques. On the other hand, the k-nearest neighbour is an example of a high variance-low bias method. Not much work has been done on the biasness and variance of quantile regression. For modelling techniques with high variance, the procedure in which prediction is carried out changes when the training dataset changes while the performance of high biasness techniques changes when assumptions made by the model change. High bias-low variance techniques tend to under fit data creating high training error and test error while low bias-high variance methods normally over fit the data by including irrelevant things like noise (Brownlee, 2016). The aim of a

model is to balance the biasness and variance. According to Parry, et.al. (2010), k-nearest neighbour balances the high variance by increasing the k value thus increasing the biasness of the model. Also by increasing the training dataset, the variance is reduced thus balancing the biasness. These are the errors influencing the performances of the models.

The implication of the study on numerical modelling

Four components are fundamental in solving geo-engineering problems i.e. the ground profile (genesis, site investigation), material behaviour (laboratory and field testing), appropriate model (conceptual, idealization) and empiricism (Burland, 2000). This is also applicable to rock engineering problems. A triangular model explaining the link amongst the components is shown in figure. 5.5. This diagram is known as the Burland triangle. According to Steenfelt (2000), it is critical to keep the four components balanced. It is thus essential to properly conduct site investigations and model laboratory data as this data has an influence on the selection of an appropriate numerical model. Poor modelling of data can indeed cause dire consequences on the support design.

During the development of majority of the failure criterion, constants relating to the strength of the material behaviour were introduced. An example being the Hoek and Brown criterion were two dimensionless strength parameters, m and s as well as σ_c were introduced (Colmenares and Zoback, 2002). These parameters are determined from modelling of laboratory results. Moreover, these constants are generally used as rock mass properties when using certain numerical modelling packages such as Examine2D.

Based on the results from the four models, it was concluded that the machine learning based techniques have a higher accuracy as compared to the statistical techniques in terms of predicting rock failure. It should however be noted that statistical techniques are also classified as explanatory techniques. This implies that they have the ability to measure specific parameters which are likely to have an effect on the dependent variable. Hence they are normally used in determining constant variables (m and s) for failure criterions which are then used in numerical modelling as stated. Machine learning techniques on the other hand are pure predictive techniques i.e. they predict the likelihood of an event occurring in the future based on a specific criterion. In the context of this research, they were used to predict rock failure based on the Mogi (1971) failure criterion. They cannot measure unknown variables as compared to statistical techniques. Therefore, the statistical techniques can be used mainly for numerical modelling whilst the machine learning based techniques can be

used to compare results obtained from numerical modelling as both methods will predict rock failure.

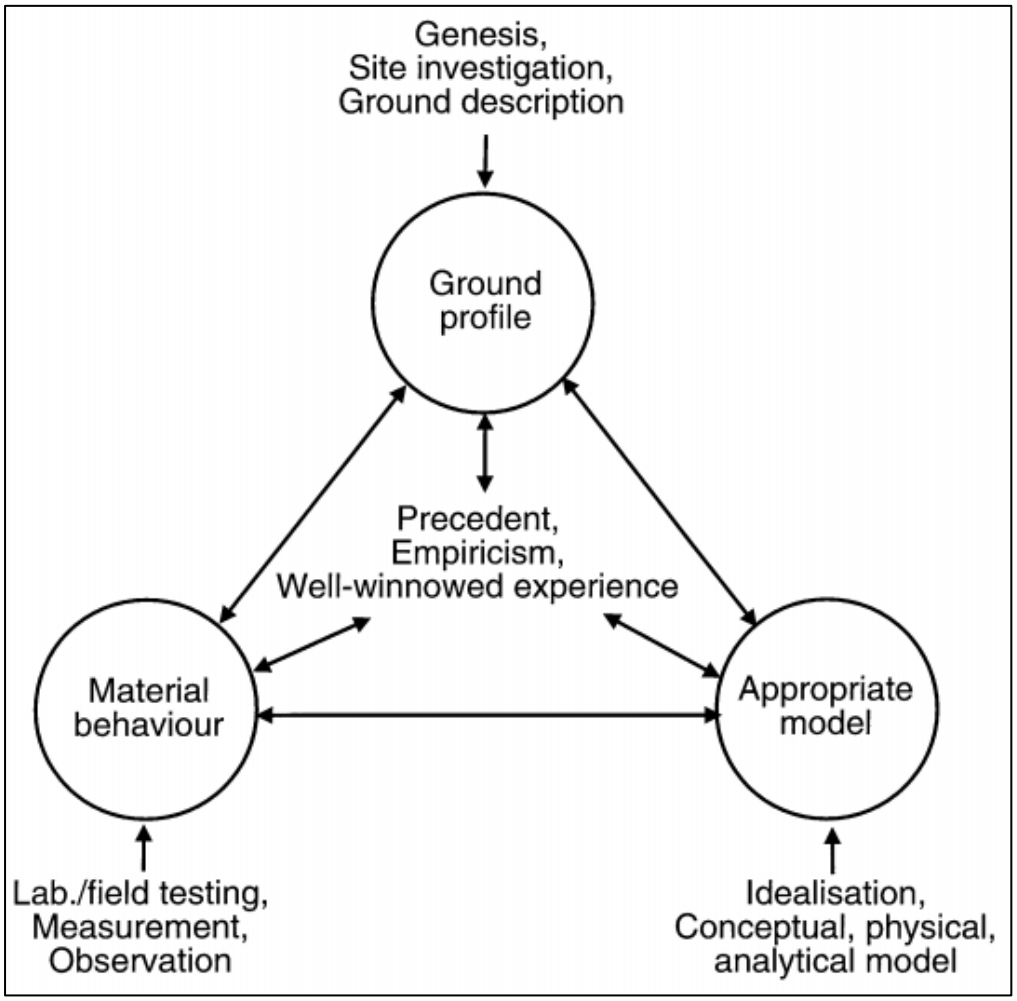

Figure 5.5: The Burland triangle (Burland, 2000)

Applicability of the study on rock masses

Mogi's failure criterion can be applied to both laboratory samples and rock masses. Giving reference to the Hoek-Brown failure criterion, it should be noted that the original criterion was derived empirically from the results of laboratory triaxial tests however the main aim of the criterion was to estimate the strength of jointed rock masses (Hoek, et.al., 1992). This implies that any failure criterions including Mogi (1971) can be used on either laboratory samples or rock masses despite their origin unless stated otherwise.

A clear relationship between the lab/field testing and site investigation exists as shown in figure 5.5 above. This implies that the two components work hand in hand in order to produce meaningful results. Not only are the four modeling techniques (least square, quantile

regression, logistic regression and k-nearest neighbour) applied in laboratory test data, they can also be applied to rock masses. With the use of a strain gauge, the various strains of a rock mass can be deduced and converted into principal stresses. These stresses can then be modelled using the four modelling techniques to determine the state of the rock mass i.e. whether it will fail or not. Logistic regression can further provide the probabilities of failure occurring within that rock mass. The stresses and failure rate determined from the laboratory testing can be compared with that from the strain gauge through modeling.

Practical application of the methodology

Triaxial data (both from true or conventional triaxial) should be analyzed after testing. This is normally conducted to get the perfect fit therefore determining the right failure criterion and the strength parameters correlating to that failure criterion. The two statistical tests i.e. the least square and quantile regression can be used to determine the failure criterion and parameters such as m and s for Hoek and Brown as they are explanatory techniques. On the other hand, the two machine learning based techniques are mainly used to predict the probability of failure occurrences using a specific failure criterion. For example, the k-nearest neighbour can be used to determine the failure probability of loose rock wedges in a hanging wall depending on the state of the rocks surrounding the loose wedges. The procedure would be to collect the strain readings using a strain gauge and convert them into stresses. The important factor would then be to select the appropriate failure criterion based on the results from the statistical techniques, create a code using the chosen failure criterion then model the stresses. If the rocks surrounding the wedges do not show any signs of failure, then the machine learning technique (k-nearest neighbour) will predict that the wedge is stable. But if there are signs of failure around the vicinity of the wedges, then it will be predicted that the loose rock wedges will failure. The machine learning techniques are mainly dependent on the defined failure criterion and the neighbouring rocks. For better results, it would be recommended for both statistical and machine learning techniques to be used simultaneously.

5.2. Conclusions

In statistics, the relationship between the observed data and fitted model is given by the coefficient of determination ($R^{2)}$. It was noted that this parameter is influenced by the range of x values as well as the presence of outliers. A conclusion was then drawn that R^2 can be biased and would not give a true reflection of how the model fits the observed data. Furthermore, it was observed that individual modelling techniques are influenced by various

factors. These factors have an influence on the predictive accuracy of that individual technique. It was noted that the least square regression was strongly affected by outliers. In the presence of outliers, the regression line shifted towards those outliers. This had an influence on the coefficient of determination and the predictive performance of the technique. It was then concluded that least square regression is not suitable for widely distributed data with outliers. The quantile regression on the other hand was robust however; this technique had a high number of indecisive points as two regression lines were used. Hence a rock state prediction of 223 out of 398 points was observed. The remaining 175 points were neither stable nor failing. This had a direct influence on the predictive accuracy of the model and it was concluded that an increase in indecisive points increases the error estimation of the model. The last two techniques were also affected by specific parameters, respectively. It was noted that the logistic regression was influenced by the number of variables used when modelling. For this research, five variables (σ_1, σ_2, σ_3, octahedral shear stress and mean effective normal stress) were initially used as modelling variables. From the five variables, only three were statistically significant (σ_2, octahedral shear stress and mean effective normal stress). This implied that rock failure is affected by σ_2, octahedral shear stress and mean effective normal stress based on the logistic regression model. The latter statement contradicted Mogi's conclusion that rock failure is influenced by all three stresses. It can be concluded that it is fundamental for the user to correctly select the right variables when modelling using logistic regression. Furthermore, this model was strongly influenced by the cutoff value. This value has the ability to change failure predictions. It was observed that an increase in cutoff value increased the model prediction accuracy until an optimum cutoff was reached. It is thus important to select the correct cutoff value in order to determine the maximum accuracy. The k-nearest neighbour was influenced by the k value and the split value. Based on work done by several authors, it was concluded that there is no guideline for choosing k-values. This is done randomly and influences the predictive accuracy of the model. The split value also affects the accuracy. A higher split will produce a higher accuracy since the dataset used for training the algorithm is large and only a limited dataset is actually tested on the model. A lower split on the other hand implies that the algorithm is trained by fewer dataset and tested on a large dataset. In the latter scenario, there is a high probability of misclassification since a larger dataset is tested. It can be concluded that the specifics applied when training and testing the dataset results to a high variance. Since more assumptions are made when modelling using least square regression and logistic regression, these two

techniques show a high biasness and low variance. K-nearest neighbour on the other hand shows high variance and low biasness as a result of high specifics made on the trained dataset. Not much work has been done on the biasness and variance of quantile regression. It should be noted also that the study can be applicable to rock masses and plays a crucial role in numerical modelling. Some rock mass parameters can only be determined through laboratory testing hence the balance between numerical modelling and lab testing should exist.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

The relationship between the strength of rocks and the stresses subjected to rocks is critical in rock engineering. In order to understand this significant relationship, rocks are exposed to laboratory testing. These testing as stated already can either be conventional or true triaxial. In general, the results from the tests are expressed as principal stresses and analysed with the aim of predicting rock failure. One method of analysing the results is to apply a modelling technique that can interpret the relationship between the rock strength and stresses and therefore predict rock failure. It is tradition to use least square regression as a predictive modelling technique. This research was aimed at exploring various modelling techniques focusing on their abilities and limitations.

The background principles behind each modelling technique are fundamental as they assist in understanding the limitation and abilities of each technique. The four modelling techniques of interest were categorized into two groups i.e. statistical techniques (least square regression and Quantile regression) and machine learning based techniques (Logistic regression and K-nearest neighbour). From the statistical techniques, the least square regression was easy to implement, however it was strongly influenced by outliers and assumptions made. On the other hand, the quantile regression was a robust technique as it was not affected by outliers. Both statistical techniques use regression lines as failure envelops. This implies that points plotting directly on the lines are indecisive. Since quantile regression uses regression lines as quantile divisions, more points tend to plot directly in those lines thus becoming indecisive. The same was observed for the least square regression but the effect was not severe since this method uses a single line as a regression line. In total, 398 points were recovered from 11 rocks. The least square regression could only classify 259 points as either failing or stable. The remaining 139 points were indecisive. With the quantile regression model, 223 out of 398 points were classified. The remaining 175 points were indecisive. This had an influence on the predictive performance of the models. Furthermore, the performances of the two statistical techniques were evaluated using the coefficient of determination (R^2). It was observed that R^2 is influenced by the range of the x-axis as well as the outliers. This suggests that R^2 might not give a true reflection in terms of the goodness of fit of the model. It should be used with other evaluating measures and not as an individual measure.

The machine learning techniques were mainly dependent on the availability of data. It was impossible to predict and classify the points of the Yuubari Shale using the machine learning techniques since the true triaxial rock strength of this rock was not available. Hence out of 398 points, 372 were classified using machine learning techniques. Both logistic regression and k-nearest neighbour were influenced by the pre-processing methods of the data. Logistic regression was mainly affected by the cutoff value as well as the chosen variables. On the other hand, k-nearest neighbour was affected by the split and the k value. These parameters had an influence on the sensitivity and specificity thus affecting the predictive performance accuracy of the modelling techniques.

Least square regression and logistic regression are parametric techniques whilst quantile regression and k-nearest neighbour are non-parametric in nature. Parametric techniques are highly biased implying that errors are mainly due to assumptions made on the pre-processing stages. Non-parametric techniques have a high variance i.e. errors resulting from the trained data. If different data is used in the algorithm using non-parametric techniques, the results will show great variation. Four components are critical in solving rock engineering problems i.e. site investigation, lab/field testing, numerical modelling and well-winnowed experienced personnel. It is critical to balance these components hence poor modelling of lab data will have a direct influence on the numerical modelling therefore affecting the final design. It is then crucial to thoroughly model laboratory data.

In overall, machine learning based techniques were better predictors than statistical techniques. Logistic regression and k-nearest neighbour were more sensitivity to failure than the least square and quantile regression. Hence k-nearest neighbour had the highest accuracy in predicting rock failure followed by logistic regression then least square regression and lastly quantile regression. It was also observed that statistical techniques were easy to implement and interpret as compared to machine learning techniques. This agrees with the statement made by Plate in 1999 that modelling techniques with high accuracy are difficult to interpret while those with poor accuracy are quite easy to interpret.

The following recommendations rose from the study:

More background is needed on the predictive modelling techniques.

For future research:

- More modelling techniques such as support vector machines, decision trees, naïve Bayes and neural networks should be evaluated. These mentioned techniques are known to be very accurate in predicting outcomes.
- Construction of a failure criterion using algorithms should be explored. This algorithm should be in such a way that information from laboratory is directly loaded and the results are produced i.e. failing or stable without having to run any techniques since all modelling techniques have limitations. This will be only possible if rock failure is thoroughly understood from a mathematical perspective.

REFERENCES

Abedallah, L., & Shimshoni, I. (2012). K-nearest neighbour using ensemble clustering. *International conference ondata warehousing and knowledge discovery*, (pp. 265-278).

Adel, M., Ajmi, A. L., & Zimmerman, R. W. (2007). The Mogi-Coulomb true triaxial failure criterion and some implication for rock engineering. *International Society for Rock Mechanics*, 6.

Akbligic, O., & Akinci, E.D. (2009). A novel regression-approach: Least square ratio. *Communication in Statistics -theory and methods*.

Alkasassbeh, M., Hassanat, , A., & Altarawnel, , G. A. (2015). On enhancing the performance of nearest neighbour classifiers using hassanat distance metric . *Indian Journal of Pure and Applied Mathematics*.

Austin, P., & Steyerberg, E. (2012). Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. *BMC Medical Research Methodology*.

Bieniawski, Z. T. (1974). Estimating the strength of rock materials. *Journal of the South African Institute of Mining and Metallurgy*, 312-320.

Brady, B., & Brown, E.T. (2004). *Rock mechanics for underground mining.* Kluwer Academic Publishers.

Brownlee, J. (2016). *Gentle introduction to the bias-variance trade-off in machine learning* . Machine Learning Algorithm.

Buhai, S. (2004). Quantile regression: overview and selected applications Ad-Astra. *The Young Romanian Scientists Journal, 5*, 1-17.

Burland, J. (2000). The soil mechanics triangle revisited. *Keynote, 1st International conference on Geotechnical Engineering Education and Training.* Sinaia.

Chen, Y. (2016). Spatial autocorrelation approaches to testing residuals from least square regression. *PLOS One*.

Colmenares, L. B., & Zoback, M. D. (2002). A statistical evaluation of intact rock failure criteria constrained by polyaxial test data for five different rocks. *International Journal of Rock Mechanics and Mining Sciences, 39*, 695-729.

Cornell, J., & Berger, R. D. (1986). Factors that influence the value of the coefficient of determination in simple linear and non-linear regression models. *The American Phytopathological Society, 77*.

Eskandarinia, A., Nazarpour, H., Teimouri, M., & Ahmadi, M. Z. (2010). Comparison of neural network and k-nearest neigbor method in daily flow forcasting. *Journal of Applied Sciences, 10*, 1006-1010.

Fleiss, J. L., Williams, J. B., & Dubro, A. F. (1986). The logistic regression analysis of psychiatric data. *Journal of Psychiatric Research, 20*(3), 195-206.

Florkowski, C. (2008). Sensitivity, specificity, receiver-operating characteristics (ROC) curves and Likelihood ratios: communicationg the performance of diagnostic tests. *The Clinical Biochemist Reviews, 29*, 83-89.

Fox, J., & Weisberg, S. (2013). Robust regression. *Thousand Odds*.

Hagquist, C., & Stenbeck, M. (1998). Goodness of fit in the R-squared and G-squared considered. *32*, 229-245.

Haimson, B. (2006). True triaxial stresses and the brittle fracture of rock. *Pure Applied Geophysics, 163*(5), 110-1130.

Haimson, B. & Chang, C. (2000). A new true triaxial cell for testing mechanical properties of rock and its use to determine rock strength and deformability of westerly granite. *International Journal of Rock Mechanics and Mining Sciences*, 285-296.

Haimson, B. & Chang, C. (2002). True triaxial strength of the KTB amphibolite under borehole wall conditions and its use to estimate the maximum horizontal in situ stress. *Journal of Geophysical Research.*

Hao, L., & Naiman, D. Q. (2007). *Quantile regression.* Califonia: Sage publication.

Hoek, E. (1968). *Rock mechanics in engineering practice.* London: Willey 99 to 124.

Hoek, E. W. (1992). A modified Hoek-Brown Criterion for Jointed Rock Masses. *Rock Characterization: ISRM Symposium*, 209-213.

Hoek, E., & Brown, E. T. (1980). Empirical strength criteria for rock mass. *Journal of the Geotechnical Engineering Division, GT9*, 1013-1035.

Hu, L., Huang, M., Ke, S., & Tsai, C. (2016). The distance function effect on k-nearest neigbor classification for medical datasets. *Springer Plus*.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*, 33-50.

Kwasniewski, M., Li, X., & Takahashi, M. (2012). *True triaxial testing of rocks* (1st ed.). Geomechanics Research Series.

Lalkhen, A., & McCluskey, A. (2008). Clinical tests: sensitvity and specificity, " continuing education in anaesthesic critical care and pain". *Scientific Research, 8*, 221-223.

Lee, H., & Haimson, B.C. (2011). True triaxial strength, deformability and brittle failure of granodiorite from the San Andreas fault observatory at depth. *International Journal of Rock Mechanics and Mining Sciences, 48*, 1199-1207.

Li, Z., Shi, J, & Tanga, A. (2015). An investigation of failure modes and failure criteria of rock in complex stress state. *The Southern African Institute of Mining and Metallurgy, 117*.

Ma, C., Yang, W., & Cheng, B. (2014). How the parameters of k-nerarest neghbour algorithm impact the best classification accuracy: in case of parkinston dataset. *Journal of Applied Sciences, 14*, 171-176.

Mogi, K. (1971). Fracture and flow of rocks under high triaxial compression. *Journal of. Geophysical. Research*, 1255-1269.

Montgometry , D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). New Jersey: John Wiley & Sons Inc.

Moore, D., Notz, W.I., & Flinger, M.A. (2013). *The basic practice of statistics* . New York: Freeman and Company.

MOSHAB. (1997). Geotechnical consideration in underground mines. *Department of Industry and Resources*, 6-66.

Oneata, D. (2011). *Fast low-rank metric learning.* Univesity of Edinburgh.

O'Neill, C. R. (1999). *Least square line fitting.* Oklahomo State University , Engineering. Stillwater: School of Mechanical and Aerospace Engineering.

Parikh, R., Mathai, A., Parikh, S., Sekha, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictivie values. *Indian Journal of Ophthalmology, 56*, 45-50.

Parry, R., Jone, W., Strokes, T. H., Phan, J. H., Moffitt, R. A., Fang, H., et al. (2010). K-nearest neighbour models for microarray gene expression analysis and clinical outcome prediction . *The Pharmacogenomics Journal*, 292-309.

Peng, C., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting . *The Journal of Edicational Research*.

Plate, T. A. (1999). Accuracy versus interpretability in flexible modeling: Implementing a tradeoff using Gaussian process models. *Behaviormetrika, 26*(1), 29-50.

Ranganathan, P., Pramesh, C. S., & Aggarwal, R. (2017). Common pitfalls in statistical analysis : Logistic regression. *Perspectives in Clinical Research, 8*.

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis: part 14 of a series on evaluation of scientific publication. *Deutshes Arzteblatt International*.

Shmueli, G. (2010). To explain or to predict ? *Institute of Mathematical Statistics, 25*.

Soureshjani, M., & Kimiagari, , A. L. (2012). Calculating the best cut-off point using logistic regression and neural network on credit scoring problem- A case study of a commercial bank. *African Journal of Business Management* .

Steenfelt, J. (2000). Teaching for the millennium or for the students. *Research Gate*.

Takahashi, M. & Koide, H. (1989). Effect of the intermediate principal stress on strength and deformation behaviour of sedimentary rocks at the depth shallower than 2000m. *In: Maury V, Fourmaintraux D, editors. Rock at great depth.*, 19-26.

Tape, T. (2001). Interpretation of diagnostic tests. *Annals of Internal Medicine*.

Visa, S., Ramsay, B, Ralesce, A, & VanDerKnaap, E. (2011). Confusion matrix based feature selection. *Proceedings of the 22nd Midwest Artificial Intelligence and Cognitive Science Conference*, (pp. 120-127).

Yao, Z., & Ruzzo, W. L. (2004). A regression -based k-nearest neighbor algorithm for gene function prediction for heterogeneous data. *BMC Bioinformatics, 7*.

York, D. (1966). *Least-square fitting of a straight line.* University of Toronto, Physics. Toronto: Geophysics Division.

You, M. (2010). Three independent parameters to describe conventional triaxial compressive strength of intact rock. *Journal of Rock Mechanics and Geotechnical Engineering, 4*, 350-356.

Zelle, M. (2004). *Python programming: an introduction to computer science.* Franklin, Beedle & Associates Inc.

APPENDICES

Appendix A

Table A1: True triaxial data for the eleven rocks used in the research

Solnhofen limestone (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
310	0	20	310
397	20	20	310
417	51	20	310
413	92	20	310
453	165	20	310
460	206	20	310
465	233	20	310
449	40	40	310
446	40	40	310
46	80	40	310
499	113	40	310
530	193	40	310
547	274	40	310
535	315	40	310
473	60	60	310
517	87	60	310
537	102	60	310
530	113	60	310
576	164	60	310
550	197	60	310
553	275	60	310
557	345	60	310
528	80	80	310

572	126	80	310
577	150	80	310
647	208	80	310
59	225	80	310
677	283	80	310
665	298	80	310
650	378	80	310
680	454	80	310

Dunham Dolomite (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
258	0	0	261.5
265	0	0	261.5
400	25	25	261.5
475	66	25	261.5
487	45	45	261.5
495	96	25	261.5
540	60	65	261.5
545	272	25	261.5
560	129	25	261.5
568	65	65	261.5
570	97	45	261.5
571	174	25	261.5
576	126	45	261.5
586	229	25	261.5
606	160	45	261.5
620	85	85	261.5

622	294	45	261.5
638	117	65	261.5
639	183	45	261.5
644	153	65	261.5
670	240	45	261.5
670	266	45	261.5
682	105	105	261.5
684	128	85	261.5
685	262	65	261.5
687	208	65	261.5
701	393	65	261.5
719	153	85	261.5
725	125	125	261.5
744	233	85	261.5
746	318	65	261.5
773	306	85	261.5
778	167	105	261.5
786	205	105	261.5
798	445	105	261.5
805	268	105	261.5
818	376	85	261.5
822	415	105	261.5
824	187	125	261.5
824	334	105	261.5
840	356	105	261.5
860	239	125	261.5

863	270	105	261.5
863	293	125	261.5
883	253	145	261.5
886	516	125	261.5
897	362	125	261.5
918	463	125	261.5
922	349	145	261.5
923	324	145	261.5
927	296	145	261.5
941	414	125	261.5
1002	410	145	261.5
1015	392	145	261.5

Yamaguchi Marble (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
82	0	0	82
118	6	6	82
140	12.5	12.5	82
179	26	12.5	82
177	28	12.5	82
196	45	12.5	82
213	67	12.5	82
225	90	12.5	82
228	105	12.5	82
200	115	12.5	82
189	25	25	82

209	39	25	82
240	58	25	82
252	78	25	82
275	107	25	82
268	132	25	82
268	157	25	82
250	168	25	82
243	40	40	82
290	64	40	82
288	88	40	82
309	88	40	82
319	112	40	82
307	143	40	82
336	160	40	82
321	177	40	82
341	208	40	82

Orikabe Monzonite (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
234	0	0	234
339	5	50	234
504	20	20	234
584,7	40	40	234
636	59	40	234
698	80	40	234
673	101	40	234

775	102	40	234
739	121	40	234
747	143	40	234
777	168	40	234
748	187	40	234
751	80	80	234
834	95	80	234
810	108	80	234
836	117	80	234
854	135	80	234
893	147	80	234
889	182	80	234
930	183	80	234
906	216	80	234
973	218	80	234
926	281	80	234
956	284	80	234
966	311	80	234
962	140	140	234
1098	205	140	234
1144	259	140	234
1161	331	140	234
1168	424	140	234
1107	200	200	234
1168	235	200	234
1244	251	200	234

1305	298	200	234
1352	343	200	234
1329	401	200	234
1358	473	200	234
1364	537	200	234

Mizuho Trachyte (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
100	0	0	100
196	15	15	100
259	30	30	100
302	45	45	100
314	58	45	100
327	67	45	100
341	90	45	100
350	138	45	100
359	204	45	100
368	281	45	100
353	323	45	100
341	60	60	100
353	83	60	100
386	133	60	100
401	186	60	100
403	212	60	100
401	254	60	100

381	306	60	100
368	75	75	100
405	108	75	100
415	147	75	100
438	210	75	100
440	279	75	100
430	318	75	100
452	363	75	100
437	100	100	100
463	126	100	100
493	171	100	100
497	256	100	100
522	354	100	100
510	384	100	100

Manazuru Andesite (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
140	0	0	140
349	16	16	140
364	20	20	140
381	20	20	140
470	67	20	140
516	124	20	140
538	186	20	140
552	40	40	140

577	75	40	140
632	112	40	140
669	126	40	140
653	206	40	140
626	278	40	140
671	70	70	140
735	101	70	140
735	152	70	140
808	193	70	140
812	275	70	140
801	313	70	140
833	375	70	140
806	100	100	140
875	110	110	140
881	130	130	140

Inada granite (Mogi, 1971)			
σ_1	σ_2	σ_3	Rock Strength
236	0	0	229
232	0	0	229
339	5	5	229
504	20	20	229
583	40	40	229
571	40	40	229
600	40	40	229
636	59	40	229

698	80	40	229
673	101	40	229
775	102	40	229
739	121	40	229
747	143	40	229
777	168	40	229
748	187	40	229
718	80	80	229
742	80	80	229
794	80	80	229
834	95	80	229
810	108	80	229
836	117	80	229
854	135	80	229
893	147	80	229
889	182	80	229
930	183	80	229
906	216	80	229
973	218	80	229
926	281	80	229
956	284	80	229
966	311	80	229
943	140	140	229
981	140	140	229
1098	205	140	229
1144	259	140	229

1161	331	140	229
1168	424	140	229
1107	200	200	229
1168	235	200	229
1244	251	200	229
1305	298	200	229
1352	343	200	229
1329	401	200	229
1358	473	200	229
1364	537	200	229

KTB Amphibolite (Haimson and Chang, 2002)			
σ_1	σ_2	σ_3	Rock Strength
165	0	0	164.7
346	79	0	164.7
291	149	0	164.7
347	197	0	164.7
267	229	0	164.7
410	30	30	164.7
479	60	30	164.7
599	100	30	164.7
652	200	30	164.7
571	249	30	164.7
637	298	30	164.7
702	60	60	164.7

750	88	60	164.7
766	103	60	164.7
745	155	60	164.7
816	199	60	164.7
888	249	60	164.7
828	299	60	164.7
887	347	60	164.7
954	399	60	164.7
815	449	60	164.7
868	100	100	164.7
959	164	100	164.7
1001	199	100	164.7
945	248	100	164.7
892	269	100	164.7
1048	300	100	164.7
1058	349	100	164.7
1155	442	100	164.7
1118	597	100	164.7
1147	150	150	164.7
1065	198	150	164.7
1112	199	150	164.7
1176	249	150	164.7
1431	298	150	164.7
1326	348	150	164.7
1169	399	150	164.7
1284	448	150	164.7

1265	498	150	164.7
1262	642	150	164.7

Westerly Granite (Haimson and Chang, 2002)			
σ_1	σ_2	σ_3	Rock Strength
201	0	0	207.8
306	40	0	207.8
301	60	0	207.8
317	80	0	207.8
304	100	0	207.8
231	2	2	207.8
300	18	2	207.8
328	40	2	207.8
359	60	2	207.8
353	80	2	207.8
355	100	2	207.8
430	20	20	207.8
529	40	20	207.8
602	60	20	207.8
554	62	20	207.8
553	61	20	207.8
532	79	20	207.8
575	100	20	207.8
567	114	20	207.8
601	150	20	207.8

638	202	20	207.8
605	38	38	207.8
620	38	38	207.8
700	57	38	207.8
733	78	38	207.8
720	103	38	207.8
723	119	38	207.8
731	157	38	207.8
781	198	38	207.8
747	60	60	207.8
811	90	60	207.8
821	114	60	207.8
860	180	60	207.8
861	249	60	207.8
889	77	77	207.8
954	102	77	207.8
992	142	77	207.8
998	214	77	207.8
1005	310	77	207.8
1012	100	100	207.8
1103	165	100	207.8
1147	167	100	207.8
1155	216	100	207.8
1195	259	100	207.8
1129	312	100	207.8

Shirahama Sandstone (Takahashi and Koide, 1989)			
σ_1	σ_2	σ_3	Rock Strength
68	2	0	60
96	9	5	60
98	16	5	60
87	30	5	60
110	46	5	60
93	65	5	60
112	15	8	60
133	25	8	60
136	40	8	60
13	51	8	60
129	76	8	60
159	29	15	60
168	62	15	60
167	82	15	60
165	87	15	60
170	98	15	60
188	104	20	60
195	100	20	60
198	85	20	60
198	72	20	60
185	60	20	60
190	51	20	60
172	40	20	60

183	30	20	60
224	46	30	60
227	72	30	60
234	92	30	60
230	112	30	60
244	132	30	60
226	152	30	60
217	173	30	60
247	61	40	60
256	73	40	60
257	82	40	60
257	102	40	60
277	102	40	60
268	124	40	60
283	143	40	60
278	162	40	60

Yuubari Shale (Takahashi and Koide, 1989)			
σ_1	σ_2	σ_3	Rock Strength
158	21	25	N/A
166	20	25	
180	30	25	
186	30	25	
173	40	25	
173	50	25	

182	61	25	
198	70	25	
192	75	25	
193	79	25	
198	89	25	
192	95	25	
183	110	25	
194	119	25	
182	130	25	
225	44	50	
234	44	50	
241	44	50	
252	63	50	
256	83	50	
261	93	50	
255	103	50	
255	114	50	
281	122	50	
261	143	50	
251	153	50	

Appendix B

Table B1: Results from the logistic regression modelling

σ_1	σ_2	σ_3	Normal	Shear	Strength	Stable (1)_Fail(0)
310	0	20	165	141.6569	310	1
397	20	20	208.5	177.7195	310	1

417	51	20	218.5	180.2856	310	1
413	92	20	216.5	170.8391	310	1
453	165	20	236.5	179.9562	310	1
460	206	20	240	180.3429	310	1
465	233	20	242.5	181.7257	310	1
449	40	40	244.5	192.8044	310	1
446	40	40	243	191.3902	310	1
46	80	40	43	17.61313	310	1
499	113	40	269.5	201.3858	310	1
530	193	40	285	204.6889	310	1
547	274	40	293.5	207.1859	310	1
535	315	40	287.5	202.4983	310	1
473	60	60	266.5	194.6901	310	1
517	87	60	288.5	209.3583	310	1
537	102	60	298.5	215.6432	310	1
530	113	60	295	210.1846	310	1
576	164	60	318	222.8143	310	1
550	197	60	305	206.4187	310	1
553	275	60	306.5	201.8134	310	1
557	345	60	308.5	203.6277	310	1
528	80	80	304	211.1892	310	1
572	126	80	326	221.8849	310	1
577	150	80	328.5	219.6558	310	1
647	208	80	363.5	242.8063	310	1
59	225	80	69.5	73.80304	310	1
677	283	80	378.5	247.8472	310	1
665	298	80	372.5	241.3936	310	1
650	378	80	365	232.7822	310	1
680	454	80	380	247.4205	310	1
258	0	0	129	121.6224	261.5	1
265	0	0	132.5	124.9222	261.5	1
400	25	25	212.5	176.7767	261.5	1
475	66	25	250	203.1589	261.5	1
487	45	45	266	208.3608	261.5	0
495	96	25	260	206.866	261.5	1
540	60	65	302.5	225.1049	261.5	1
545	272	25	285	212.3775	261.5	1
560	129	25	292.5	231.6132	261.5	1
568	65	65	316.5	237.1165	261.5	1
570	97	45	307.5	236.1868	261.5	1
571	174	25	298	230.4406	261.5	1
576	126	45	310.5	233.5765	261.5	1
586	229	25	305.5	231.8491	261.5	1
606	160	45	325.5	241.9509	261.5	1

620	85	85	352.5	252.2014	261.5	1
622	294	45	333.5	236.2941	261.5	1
638	117	65	351.5	258.7307	261.5	1
639	183	45	342	253.8188	261.5	1
644	153	65	354.5	254.7474	261.5	1
670	240	45	357.5	261.0981	261.5	1
670	266	45	357.5	258.7753	261.5	1
682	105	105	393.5	272.0004	261.5	0
684	128	85	384.5	272.8015	261.5	0
685	262	65	375	258.6585	261.5	1
687	208	65	376	265.9937	261.5	0
701	393	65	383	259.6887	261.5	1
719	153	85	402	284.2018	261.5	0
725	125	125	425	282.8427	261.5	0
744	233	85	414.5	282.3131	261.5	0
746	318	65	405.5	281.0603	261.5	0
773	306	85	429	286.7973	261.5	0
778	167	105	441.5	303.6983	261.5	0
786	205	105	445.5	300.2447	261.5	0
798	445	105	451.5	282.9327	261.5	0
805	268	105	455	299.0611	261.5	0
818	376	85	451.5	301.3551	261.5	0
822	415	105	463.5	293.6056	261.5	0
824	187	125	474.5	315.9138	261.5	0
824	334	105	464.5	299.9078	261.5	0
840	356	105	472.5	305.0468	261.5	0
860	239	125	492.5	322.983	261.5	0
863	270	105	484	325.4805	261.5	0
863	293	125	494	315.8354	261.5	0
883	253	145	514	325.4412	261.5	0
886	516	125	505.5	310.7164	261.5	0
897	362	125	511	322.8997	261.5	0
918	463	125	521.5	324.9133	261.5	0
922	349	145	533.5	328.9164	261.5	0
923	324	145	534	332.687	261.5	0
927	296	145	536	338.7044	261.5	0
941	414	125	533	337.8208	261.5	0
1002	410	145	573.5	358.2578	261.5	0
1015	392	145	580	366.0659	261.5	0
82	0	0	41	38.65517	82	1
118	6	6	62	52.79731	82	1
140	12.5	12.5	76.25	60.10408	82	1
179	26	12.5	95.75	75.50828	82	1
177	28	12.5	94.75	74.16311	82	1

196	45	12.5	104.25	79.95103	82	1
213	67	12.5	112.75	84.64731	82	0
225	90	12.5	118.75	87.80502	82	0
228	105	12.5	120.25	88.27073	82	0
200	115	12.5	106.25	76.65761	82	1
189	25	25	107	77.31034	82	1
209	39	25	117	83.63412	82	0
240	58	25	132.5	94.53865	82	0
252	78	25	138.5	96.96162	82	0
275	107	25	150	104.0555	82	0
268	132	25	146.5	99.43954	82	0
268	157	25	146.5	99.32774	82	0
250	168	25	137.5	92.97431	82	0
243	40	40	141.5	95.69512	82	0
290	64	40	165	112.6213	82	0
288	88	40	164	107.3975	82	0
309	88	40	174.5	117.1447	82	0
319	112	40	179.5	118.2624	82	0
307	143	40	173.5	109.9465	82	0
336	160	40	188	121.5602	82	0
321	177	40	180.5	114.7296	82	0
341	208	40	190.5	123.1593	82	0
100	0	0	50	47.14045	100	1
196	15	15	105.5	85.32422	100	1
259	30	30	144.5	107.9516	100	0
302	45	45	173.5	121.151	100	0
314	58	45	179.5	123.8574	100	0
327	67	45	186	128.066	100	0
341	90	45	193	130.2314	100	0
350	138	45	197.5	127.6358	100	0
359	204	45	202	128.1934	100	0
368	281	45	206.5	136.4608	100	0
353	323	45	199	138.6635	100	0
341	60	60	200.5	132.4647	100	0
353	83	60	206.5	133.0322	100	0
386	133	60	223	139.6877	100	0
401	186	60	230.5	140.7843	100	0
403	212	60	231.5	140.3306	100	0
401	254	60	230.5	139.6527	100	0
381	306	60	220.5	137.1058	100	0
368	75	75	221.5	138.1215	100	0
405	108	75	240	148.3981	100	0
415	147	75	245	146.2904	100	0
438	210	75	256.5	149.8065	100	0

440	279	75	257.5	149.3549	100	0
430	318	75	252.5	148.1808	100	0
452	363	75	263.5	160.8982	100	0
437	100	100	268.5	158.8633	100	0
463	126	100	281.5	165.3327	100	0
493	171	100	296.5	171.0016	100	0
497	256	100	298.5	163.3082	100	0
522	354	100	311	173.4692	100	0
510	384	100	305	171.4746	100	0
140	0	0	70	65.99663	140	1
349	16	16	182.5	156.9777	140	0
364	20	20	192	162.1632	140	0
381	20	20	200.5	170.177	140	0
470	67	20	245	201.9675	140	0
516	124	20	268	213.5665	140	0
538	186	20	279	215.9691	140	0
552	40	40	296	241.3591	140	0
577	75	40	308.5	245.3111	140	0
632	112	40	336	263.744	140	0
669	126	40	354.5	278.4652	140	0
653	206	40	346.5	258.8723	140	0
626	278	40	333	240.6343	140	0
671	70	70	370.5	283.3141	140	0
735	101	70	402.5	306.4387	140	0
735	152	70	402.5	296.0552	140	0
808	193	70	439	322.8343	140	0
812	275	70	441	312.8645	140	0
801	313	70	435.5	303.9653	140	0
833	375	70	451.5	313.574	140	0
806	100	100	453	332.8116	140	0
875	110	110	492.5	360.6245	140	0
881	130	130	505.5	354.0248	140	0
236	0	0	118	111.2515	229	1
232	0	0	116	109.3658	229	1
339	5	5	172	157.4491	229	1
504	20	20	262	228.1598	229	1
583	40	40	311.5	255.9727	229	0
571	40	40	305.5	250.3158	229	0
600	40	40	320	263.9865	229	0
636	59	40	338	276.5875	229	0
698	80	40	369	301.1991	229	0
673	101	40	356.5	285.1109	229	0
775	102	40	407.5	332.8326	229	0
739	121	40	389.5	312.1762	229	0

747	143	40	393.5	311.8536	229	0
777	168	40	408.5	321.53	229	0
748	187	40	394	305.0672	229	0
718	80	80	399	300.7561	229	0
742	80	80	411	312.0698	229	0
794	80	80	437	336.5828	229	0
834	95	80	457	351.9568	229	0
810	108	80	445	337.7191	229	0
836	117	80	458	347.9888	229	0
854	135	80	467	352.6191	229	0
893	147	80	486.5	368.4764	229	0
889	182	80	484.5	359.7428	229	0
930	183	80	505	378.7579	229	0
906	216	80	493	361.6124	229	0
973	218	80	526.5	392.5017	229	0
926	281	80	503	360.885	229	0
956	284	80	518	374.2513	229	0
966	311	80	523	375.2602	229	0
943	140	140	541.5	378.5378	229	0
981	140	140	560.5	396.4512	229	0
1098	205	140	619	437.0911	229	0
1144	259	140	642	447.8842	229	0
1161	331	140	650.5	443.1982	229	0
1168	424	140	654	433.4584	229	0
1107	200	200	653.5	427.5639	229	0
1168	235	200	684	448.2978	229	0
1244	251	200	722	480.5767	229	0
1305	298	200	752.5	499.4083	229	0
1352	343	200	776	512.6873	229	0
1329	401	200	764.5	491.7346	229	0
1358	473	200	779	494.2692	229	0
1364	537	200	782	489.0346	229	0
165	0	0	82.5	77.78175	164.7	1
346	79	0	173	148.0413	164.7	1
291	149	0	145.5	118.8117	164.7	1
347	197	0	173.5	142.0946	164.7	1
267	229	0	133.5	117.9331	164.7	1
410	30	30	220	179.1337	164.7	0
479	60	30	254.5	204.9558	164.7	0
599	100	30	314.5	253.3469	164.7	0
652	200	30	341	262.4856	164.7	0
571	249	30	300.5	222.1926	164.7	0
637	298	30	333.5	248.3711	164.7	0
702	60	60	381	302.6417	164.7	0

750	88	60	405	318.8744	164.7	0
766	103	60	413	323.1536	164.7	0
745	155	60	402.5	303.0127	164.7	0
816	199	60	438	328.5568	164.7	0
888	249	60	474	354.2796	164.7	0
828	299	60	444	320.8991	164.7	0
887	347	60	473.5	342.8472	164.7	0
954	399	60	507	368.5078	164.7	0
815	449	60	437.5	308.2751	164.7	0
868	100	100	484	362.0387	164.7	0
959	164	100	529.5	390.7261	164.7	0
1001	199	100	550.5	403.4306	164.7	0
945	248	100	522.5	368.4409	164.7	0
892	269	100	496	340.5802	164.7	0
1048	300	100	574	408.0044	164.7	0
1058	349	100	579	405.8525	164.7	0
1155	442	100	627.5	439.4894	164.7	0
1118	597	100	609	415.6353	164.7	0
1147	150	150	648.5	469.9903	164.7	0
1065	198	150	607.5	420.4783	164.7	0
1112	199	150	631	442.3942	164.7	0
1176	249	150	663	462.0974	164.7	0
1431	298	150	790.5	572.1843	164.7	0
1326	348	150	738	514.0973	164.7	0
1169	399	150	659.5	433.7513	164.7	0
1284	448	150	717	480.0065	164.7	0
1265	498	150	707.5	465.787	164.7	0
1262	642	150	706	454.9735	164.7	0
201	0	0	100.5	94.75231	207.8	1
306	40	0	153	135.8071	207.8	1
301	60	0	150.5	130.0778	207.8	1
317	80	0	158.5	134.6015	207.8	1
304	100	0	152	126.5052	207.8	1
231	2	2	116.5	107.9516	207.8	1
300	18	2	151	136.8633	207.8	1
328	40	2	165	145.5503	207.8	1
359	60	2	180.5	156.4232	207.8	1
353	80	2	177.5	150.4859	207.8	1
355	100	2	178.5	148.7869	207.8	1
430	20	20	225	193.2759	207.8	1
529	40	20	274.5	235.3725	207.8	0
602	60	20	311	265.4321	207.8	0
554	62	20	287	242.4376	207.8	0
553	61	20	286.5	242.174	207.8	0

532	79	20	276	228.7245	207.8	0
575	100	20	297.5	244.9603	207.8	0
567	114	20	293.5	238.8058	207.8	0
601	150	20	310.5	248.9672	207.8	0
638	202	20	329	259.3033	207.8	0
605	38	38	321.5	267.2864	207.8	0
620	38	38	329	274.3574	207.8	0
700	57	38	369	307.6892	207.8	0
733	78	38	385.5	318.6168	207.8	0
720	103	38	379	307.325	207.8	0
723	119	38	380.5	305.6145	207.8	0
731	157	38	384.5	302.5606	207.8	0
781	198	38	409.5	319.294	207.8	0
747	60	60	403.5	323.8549	207.8	0
811	90	60	435.5	347.1698	207.8	0
821	114	60	440.5	346.7125	207.8	0
860	180	60	460	352.2625	207.8	0
861	249	60	460.5	341.8684	207.8	0
889	77	77	483	382.7805	207.8	0
954	102	77	515.5	407.657	207.8	0
992	142	77	534.5	416.86	207.8	0
998	214	77	537.5	405.7457	207.8	0
1005	310	77	541	394.1937	207.8	0
1012	100	100	556	429.9209	207.8	0
1103	165	100	601.5	458.267	207.8	0
1147	167	100	623.5	478.5508	207.8	0
1155	216	100	627.5	472.3701	207.8	0
1195	259	100	647.5	483.0921	207.8	0
1129	312	100	614.5	443.6307	207.8	0
68.3	1.5	0	34.15	31.84926	60	1
96.16	9.36	5	50.58	41.98332	60	1
98.32	15.64	5	51.66	41.7104	60	1
87.39	30.05	5	46.195	34.48589	60	1
110.16	45.57	5	57.58	43.30309	60	1
93.25	65.15	5	49.125	36.81137	60	1
112.45	15.27	8	60.225	47.61723	60	1
132.52	24.88	8	70.26	55.15285	60	1
136.27	40.39	8	72.135	54.4623	60	1
13.32	50.74	8	10.66	19.01831	60	1
129.07	76.23	8	68.535	49.55955	60	1
159.14	29.31	15	87.07	64.83907	60	0
167.71	61.82	15	91.355	63.87936	60	0
166.55	81.77	15	90.775	62.01549	60	0
164.9	87.32	15	89.95	61.20898	60	0

170.4	98.03	15	92.7	63.49152	60	0
187.67	103.57	20	103.835	68.4511	60	0
195.29	100	20	107.645	71.65253	60	0
197.52	85.47	20	108.76	73.29914	60	0
198.11	71.8	20	109.055	74.80391	60	0
185.11	60.34	20	102.555	70.28211	60	0
190.04	51.11	20	105.02	73.92412	60	0
172.14	40.39	20	96.07	67.4293	60	0
183.05	30.05	20	101.525	74.6066	60	0
223.74	46.26	30	126.87	87.74884	60	0
227.46	72.17	30	128.73	84.90763	60	0
233.91	91.75	30	131.955	85.37617	60	0
230.03	112.44	30	130.015	82.0811	60	0
243.55	132.02	30	136.775	87.21023	60	0
226.08	152.34	30	128.04	80.86479	60	0
216.77	172.66	30	123.385	79.70822	60	0
247.07	60.71	40	143.535	93.11697	60	0
256.26	72.9	40	148.13	95.14415	60	0
257.32	81.77	40	148.66	94.15738	60	0
256.7	102.46	40	148.35	91.07391	60	0
277.35	102.09	40	158.675	100.5022	60	0
268.03	123.52	40	154.015	94.19626	60	0
283.18	143.1	40	161.59	99.65971	60	0
278.22	161.58	40	159.11	97.25988	60	0

Appendix C

Figure C1: The K-nearest neighbour algorithm used for rock failure

File Edit Format Run Options Window Help

```
import csv
import random
import math
import operator

def loadDataset(filename, split, trainingSet=[] , testSet=[]):
    with open(filename, 'r') as csvfile:
        lines = csv.reader(csvfile)
        dataset = list(lines)
        for x in range(len(dataset)):
            if (x > 0):

                row = [];
                for y in range(len(dataset[x])):
                    percentageReadSoFar = (x+1)*100/len(dataset);
                    percentageSplit = split * 100;
                    row.append(float(dataset[x][y]))
                if (percentageReadSoFar<percentageSplit):
                    ## trainigSet
                    trainingSet.append(row)
                else:
                    ## testSet
                    testSet.append(row)

def euclideanDistance(instance1, instance2, length):
    #calculating the distance between test and training
    distance = 0
    for x in range(length):
        distance += pow(instance1[x]-instance2[x], 2)
    return math.sqrt(distance)

def getNeighbors(trainingSet, testInstance, k):
    distances = []
    length = len(testInstance)-1
    for x in range(len(trainingSet)):
        dist = euclideanDistance(testInstance, trainingSet[x], length)
##        print(testInstance)
##        print (trainingSet[x],dist)

        distances.append((trainingSet[x], dist))
        distances.sort(key=operator.itemgetter(1))
##        print(distances)
```

```
##          print(distances)
        neighbors=[]
    for x in range(k):
        neighbors.append(distances[x][0])
##    print(neighbors)
    return neighbors

def getResponse(neighbors):
    classVotes = {}
    for x in range(len(neighbors)):
        response = neighbors[x][-1]
        if response in classVotes:
            classVotes[response] +=1
        else:
            classVotes[response] = 1
    sortedVotes = sorted(classVotes.iteritems(), key=operator.itemgetter(1), reverse=True)
    return sortedVotes[0][0]

def getAccuracy(testSet, predictions):
    correct = 0
    for x in range(len(testSet)):
        if testSet[x][-1] == predictions[x]:
            correct += 1
    return (correct/float(len(testSet)))*100.0

def main():
    # prepare data
    trainingSet=[]
    testSet=[]
    split = 0.5

    loadDataset('Data from rocks.csv', split, trainingSet, testSet)
##    print(trainingSet)
    print('Train set: ' + repr(len(trainingSet)))
    print('Test set: ' + repr(len(testSet)))
##    #generate predictions
    predictions=[]
    k = 3
    for x in range(len(testSet)):
        neighbors = getNeighbors(trainingSet, testSet[x], k)
        results = getResponse(neighbors)
```