Regression Analysis and Neural Network Fitting of Rock Mass Classification Systems

İbrahim Ferid ÖGE

1Muğla Sıtkı Koçman Üniversitesi, Mühendislik Fakültesi, Maden Mühendisliği Bölümü, 48000, Muğla. (ORCID: 0000-0001-6243-8268)

(Alınış / Received: 16.02.2017, Kabul / Accepted: 22.01.2018, Online Yayınlanma / Published Online: 15.05.2018)

Keywords
Rock Mass Classification, Rock Mass Rating, Q-System, Geological Strength Index, Regression, Neural Network

Abstract: Commonly used rock mass classification systems, Rock Mass Rating (RMR), Q-System, and Geological Strength Index (GSI) were used as input for simple regression and Neural-Network fitting. The relationship between the classification systems can be used for the estimation of unknown classification ratings. The necessary data for this study, consisting of 250 sets of rock mass classification ratings, were collected from an excavation of an underground mine opening during a time interval of more than two years. The rock mass data belongs to the Pliocene-aged Deniş formation in Soma region of Manisa/Turkey. The ratings, basic and adjusted RMR, Q, Q', and GSI were chosen for the simple regression. Three of the equations are suggested to be taken into account due to their strong correlation of determination. These equations can be utilized especially if the rating Q is known and the adjusted RMR is intended to be estimated. Additionally, basic RMR rating can be estimated by considering the GSI as an input. Utilization of the Neural Networks resulted in an improved prediction capability with a greater predicted-measured coefficient of determination. Implementing the Neural Network fitting also overcame the scatter observed in the regression analysis.

Kaya Sınıflama Sistemlerinin Regresyon ve Sinir Ağları Tekniği İle İlişkilendirilmesi

Anatlar Kelimeleri
sınıflamasi, Kaya Kütle Puanı, Q- Sistemi, Jeolojik Dayanım İndisi, Regresyon, Sinir Ağı

Özet: Yaygın olarak kullanılan kaya kütle sınıflama sistemlerinden Kaya Kütle Puanlanması (RMR), Q-Sistemi ve Jeolojik Dayanım İndisi (GSI) temel regresyon çalışmasına ve Sinir Ağı en iyi uyumlamasına tabii tutulmuştur. Sınıflama sistemleri arasındaki ilişkiler, bilinmeyen sınıflama puanlamasının kestirilmesinde kullanlabılır. Çalışmada kullanılan ve kaya sınıflama sistemleri puanlarından oluşan 250 veri iki yılda uzun bir sürede bir maden açılışının zamanından toplanmıştır. Kaya kütle verisi, Soma bölgesinde yer alan Pliyosen yaşlı Deniş birimine aittir. Temel ve düzeltmiş RMR, Q, Q’ ve GSI puanlamaları temel regresyon için
1. Introduction

Rock mass classification systems have a wide application area in rock engineering projects, such as tunnelling, slope stability, and foundation engineering. Among these application areas, tunnelling is the main area of interest. The pioneering studies proposed by Terzaghi [1] and Lauffer [2] can be accounted for the modern rock mass classification systems. Deere et al. [3], introduced a practical and still popular Rock Quality Designation (RQD). Rock Mass Rating system (RMR) is mainly applicable to tunnelling and a few modified versions were proposed by several researchers, [4, 5, 6]. Additionally, it is commonly implemented in foundation engineering and slope stability, [7]. A recent update for the RMR provides sets of equations for the empirical design of support and estimation of rock pressure on the support, [8]. The Q-system is proposed only for tunnel design and an updated version of the system introduces reinforced ribs of shotcrete as a support element for weak ground conditions, [9, 10]. A modified version of the Q-system was proposed for hard rock mining environment, [11, 12]. Rock Mass Index (RMI) [13] and Rock Mass Quality Rating (RMQR) [14] are other common rock mass classification systems. However, the RMR and Q-system are the most commonly used systems in civil and mining engineering. Geological Strength Index (GSI) is a different classification system when compared to other systems since it is only used for the representation of the structural quality of the rock mass and is incorporated in the Generalized Hoek Brown failure criterion, [15, 16]. A failure criterion for a rock mass can be estimated by using the GSI parameter and laboratory test results on intact rock. The deformability and strength parameters of the rock masses can be estimated by the utilization of the previously mentioned systems. The original GSI chart can be used by imposing a visual impression of the rock mass and obtain a GSI rating. However, many quantified versions of the GSI charts are present, [17, 18, 19, 20]. Additionally, some researches presented GSI charts of the particular use for flysch and sedimentary rock masses, [21].

2. Material and Method

2.1. Interrelation Between Rock Mass Classification Ratings

It should be reminded that the basic RMR (bRMR) rating lies between 0 and 100 consisting of 5 parameters for uniaxial compressive strength, RQD, discontinuity spacing and condition, and groundwater. Final RMR can be calculated by considering the discontinuity orientation adjustment. Additionally, adjusted RMR (aRMR) can be adjusted for blast damage, field stress, and weakness planes, [22].
However, clear guidelines for such adjustments are missing. The Q-system consists of 6 parameters, RQD, joint set number ($J_s$), joint roughness ($J_r$), joint alteration ($J_a$), ground water condition ($J_w$) and stress reduction factor (SRF).

SRF includes parameters related to weakness zones, rock burst, and squeezing conditions. In this study, $bRMR$ represents the basic RMR value according to the version published in 1989, $aRMR$ represents the discontinuity and the field stress condition adjusted RMR rating. The Q-system rating is directly used, however, $Q'$ was also accounted in this study dropping the $J_w$ and SRF parameters. [7].

Various researchers investigated the relationship between the RMR and Q systems by using several cases, (Table 1).

According to Goel et al. [23], based on 115 case histories, including 77 reported by Bieniawski [24], 4 from the Kielder experimental tunnel reported by Hoek and Brown [25], and 34 collected from India, the correlation coefficients of the approaches were stated as not sufficient. It is important to mention that the classification system guidelines were outdated since the latest version of RMR was proposed in 1989 and the Q-system was updated by minor changes in 2013. N and Rock condition rating (RCR) parameters can also be used for the correlation of well-known classification systems in a transformed form. The value $N$ is obtained for any $Q$ value with $SRF=1$. The RCR value can be obtained by taking the RMR parameters into account with the exception of the uniaxial compressive strength rating and the discontinuity adjustment. In other words, the RCR is a $bRMR$ rating with a zero rating for the UCS value. Then, a correlation with higher prediction capability can be obtained if the condition $UCS>5MPa$ is valid. [7]:

$$RCR = 8 \ln N + 30$$  \hspace{1cm} (1) 

Tzamos and Sofianos [30] used the shared parameters of the four commonly used rock mass classification systems in order to construct correlations between the systems RMR, Q, GSI, and RMi. The parameters on the fabric indices of rock mass, such as the rock structure rating and the discontinuity conditions were utilized in the correlations.

Palmström [31] claimed that the Qc and RMR correlation is stronger and had better estimation performance. The parameter $Q_c = Q(UCS/100)$ is used as a strength adjusted Q-system rating in this approach. This application improves the estimation performance especially if the rock mass is classified as weak rock.

Another relation was proposed by Hoek and Brown [15] which suggested to use $GSI = RMR_{89} - 5$ (if $RMR>23$ or $GSI>18$) in order to relate the GSI and the RMR:

Here the $RMR_{89}$ is calculated for a groundwater rating which is equal to 15. Similarly, if $GSI<18$ the following equation can be used:

$$GSI = 9 \ln Q' + 44$$  \hspace{1cm} (2)
where

\[ Q' = \frac{RQD}{J_r} \left( \frac{J_r}{J_{ra}} \right) \]  

(3)

For the GSI correlations, visual estimates or other derived equations are still used. The quantified GSI ratings can be calculated by the discontinuity condition parameters of both RMR and Q-system, as given below, [19]:

\[ GSI = 1.5 \, D_c + \frac{RQD}{2} \]  

(4)

\[ GSI = 52 \left( \frac{J_r}{J_a} \right) \left( 1 + \frac{J_r}{J_a} \right) + \frac{RQD}{2} \]  

(5)

Where \( D_c \) is the discontinuity condition rating of the RMR. Above mentioned equations provide better GSI estimates by utilizing some of the RMR and Q-system parameters.

2.2. Geology of the Area

Tuysuz and Genç [32] studied the geology of the research area. The Pliocene-aged formation (Deniş) is underlain by Miocene aged Soma formation. Deniş Formation contains clastic limnic deposit succession with coal intercalations. The unit was sub-classified into 6 series by Nebert [33]. They were identified as; Sandstone-Siltstone-reddish/greenish coloured Clay level (P1), Pliocene Lignite level (KP1), Clay-Tuff-Marl series (P2ab), Clay-Sandstone-Conglomerate level (P2c), Finely Graveled (siliceous) Calcareous level (P3) and Tuff-agglomerate (P4 or Pltv) levels, (Figure 1).

Another rock group situated together with the volcanic series, where its typical outcrops are commonly observed in the vicinity of Elmadere village, is flowing breccia. These are formed in thick layers of a lithology with bad but distinctive, various sized angular lava gravels as well as lavas in the form of cements. The pyroclastic flow and rubble units are composed of latitic, andesitic, and rarely dacitic fine lava level and intercalated with flowing breccias and lahar levels. The lahar levels contain gravel and blocks in medium-coarse size, generally rounded and spherical andesite and latite, occasionally in dacite composition. Although they are mostly represented by lava flows, their dikes and vein systems intersect with the deposits and the pyroclastics were also observed. Typical examples can be observed at the southern part of the area, in the east and west of Kocadere, and also in the vicinity of the Kızkaya regions.

![Figure 1. Generalized stratigraphy of Soma coal basin (after Aksoy et al. [35])](image-url)
fine-medium grain sized grains cemented with sand and silt. A lignite zone with a thickness between 3.5 m and 30 m overlays this basement detritus defined as M1. Known as KM2 in the regional Neogene nomenclature, this lignite has generally a hard, massive, black, and bright appearance. A bluish-grey coloured marl level overlays KM2 zone.

Brinkman et al. [34] combines this lithology defined as M2 with the upper limestone level (M3). It was determined that both marl and limestone was defined together as “marl-calcareous series”. The marl directly overlays KM2 lignite zone, with a sharp contact. The marls are grey, grey-green coloured, hard, and massive. It is medium-thick layered and abundant of leaf fossils. In these levels, marls are splitted into small plates and in almost cardboard shale appearances.

The study area is located at the south of Bakırçay Graben. This graben is one of the most important graben formations in NW Anatolia. Bakırçay Graben starts with Dikili-Çandarlı at the West, extends to the east, gets narrow eastwardly, and changes direction in the vicinity of Soma. The basin is bounded at the north by an oblique slip active fault which is also a boundary for the Bergama Valley. At the south, small fault segments dominated by vertical slip components limit the graben's boundaries. Dirik et al. [36] claimed that the coal basins located at the western part of Soma developed in Pliocene-Quaternary period and remained over the Çamlıca Rise which was disintegrated by block faults.

Another structural element in the study area are the folds. They are determined by the strike and dips of the deposits, and changing features of the layers through the borehole cores. It is known from previous studies that Soma and Deniş Formations are folded. From both Dirik's [36] data and the borehole data, some folds are syn-sedimentary folds (slump structure). The slump nature of the folds is clearly seen through the boreholes, also from the low-angled slopes or horizontal beddings that overlie or underlie steeply layered zones. Many medium-scaled folds were also determined in the area as well. These are symmetrical and are located along the NE-SW direction. The fold structures are especially important for synclinal structures, exploration of underground waters, and underground water movements. Although the strata dip directions are generally towards SE, SW and NW directions, the strikes of the layers are dominated in the NE-SW direction and are partially NW-SE directed. NE directed layer dip directions are rarely observed.

2.3. Data Characteristics
The study area is a decline heading of an underground coal mine development which is planned to pass through all geological formations in the Soma coal basin located in Soma/Manisa and Kınık/Izmir. A section with 2100 m length is considered in the study. During the decline heading, aRMR, bRMR, Q, Q', and GSI ratings were collected directly from the face mapping and logging, separately. 125 different chainages were considered in the data collection. Due to the geological complexity and strongly varying geology within short advance intervals, two values for each classification parameter were calculated. One of the parameters belongs to the dominant zone and the other one was attained for the minor zone at face. As a result, a total of 250 data points were obtained for the study. The data is obtained from the sections of the decline driven through P2c, P2ab, and volcanic units. The units, P2c and P2ab are generally observed to have poor rock mass character, while volcanic rock
masses in the study area generally exhibit good rock mass conditions.

Frequency histograms of the collected data based on the related rock mass classification systems are provided in Figure 2. When the RMR is considered for the basic and adjusted values, it is observed that the rock mass quality is mostly in fair to very poor rock classes.

![Frequency histograms for rock mass classification data](image)

**Figure 2.** Frequency histograms for rock mass classification data

The distributions of the ratings Q and Q' are similar to the RMR values. A rock mass class could not be assigned for Q' since groundwater and stress reduction factor were missing. The majority of the data points have Q values smaller than 1. There is a significant amount of data between 0.001 and 0.01. Considering the database, it can be remarked that the study is mostly concentrated on very poor to fair rock mass classes. This fact is also verified by the normality test. In the probability functions, p-values are smaller than 0.005 for all classifications. Anderson-Darling coefficients lie between 1.8 and 6.1 which are relatively high. Indicators show that the data is non-homogenously distributed. In none of the data points, extreme water inflow was observed. Mostly, conditions were dry, damp, or wet. This fact limits the utilization of the findings presented in this study. The findings can be used in cases where serious groundwater ingress...
is not present. In the field, sedimentary rock units were generally moist. Due to the excavation induced permeability increase around the excavation, the groundwater was observed to deteriorate the rock material. However, during the data collection the initial state of the rock mass was taken into account.

For the uniaxial compressive strength rating of the RMR system, the lithology is matched with the numerous laboratory test results and ISRM suggested field estimates, [37]. The drill core samples obtained from 14 boreholes were used in the laboratory testing and numerous test results were available for the study. The results and observed lithology were matched but thin beddings exhibiting frequent alternation of the geological units caused difficulties in the representation of the rock mass. The dominant and minor zone approach provided identification of a range for the classification data. The dominant and minor zone approach was employed in the chainages where it was necessary. In the presence of homogenous rock mass conditions, only a single set of rock mass classification data was identified.

2.4. Simple Regression and Neural Network Fitting

Simple regression method was implemented commonly for the correlation of RMR, Q, Q', and GSI as it is presented in section 2.1. Linear regression is preferred for the correlation of RMR and GSI since they are linear variables and have a range between 0 and 100. A logarithmic function is necessary when Q or Q' is taken into consideration for the correlation of RMR or GSI. Natural logarithm of Q or Q' parameters were utilized in this study for the regression work. The coefficient of determination is the main performance indicator for the regressions.

Alvarez Grima and Babuska [38] explained that neural networks (NN) are simplified models of the biological structure found in human brains. Neurons are elementary processing units and there is large amount of interconnections among them. The interconnections provide a capability to learn from the processed data led algorithm to exhibit strong classification and prediction performance. Therefore, neural network fitting can be used to develop complex relationships among variables. Neural-networks are widely used in rock engineering and rock mechanics in order to relate several parameters, such as laboratory or field test results, [39, 40].

Rafai and Jafari [41] used the NN as a failure criteria prediction tool. Prediction of the cutting performance was also handled by using the same method, [42, 43, 44]. Sonmez et al. [45] employed artificial NN for the prediction of intact rock modulus. Yesiloglu-Gultekin et al. [46] utilized the neural networks in order to predict the uniaxial compressive strength for six granite types. A study about the excavation performance was carried out by similar approaches with multiple parameters by Kucuk et al. [47]. Prediction of the drilling performance was investigated by the similar approach, [48].

In this study, Levenberg-Marquardt back propagation method was used for the training of the neural network. This algorithm was chosen because it is one of the fastest methods for training moderate-sized neural networks. MATLAB v.9 [39] provides a neural network fitting toolbox which enables researchers to handle the fitting process with a GUI. The structure was constructed in a simple form. One hidden and one output layer were sufficient for fitting two parameters, (Figure 3).
The neural network was constructed in the form of a feedforward network with sigmoid hidden neurons and linear output neurons. In fact, NN was well suited for the solution of multi-variable and complex problems. In this study, one variable (i.e. Q) was selected as a predictor to estimate only a single target (i.e. RMR). In fact, NN is capable of fitting one predictor and one target with a high $R^2$ since the problem is not complex.

To define a fitting problem for the toolbox, a set of input vectors were arranged in the columns of a matrix. Then, another set of target vectors were arranged correspondingly (the actual output vectors for each of the input vectors) into a second matrix. 70% of the 250 data rows were used in training of the NN, 15% of the data was reserved for testing, and 15% for the validation of the model.
3. Results
Initially, the collected data was subjected to simple regression analysis. As it is shown in Figure 4, majority of the data is available in the range of fair to very poor rock mass. For the highest quality rock mass, less amount of data was available.

For Basic RMR (bRMR), it can be indicated that the representation of very poor rock masses on-site was a problem. If the ground is dry, then a minimum rating of 15 is attained. The bRMR rating gets above 20 easily, even though the remaining ratings are lying on the lower bound. In the bRMR-Q correlation, only a few data points having the RMR value
smaller than 20 are observed. For bRMR and Q correlation, when Q rating has the smallest value (0.001), bRMR is greater than 10 according to the intercept of the fitting equation. When the adjusted RMR value in aRMR-Q correlation is investigated, an improved fitting equation with higher $R^2$ value is obtained. It should be reminded that, aRMR value is adjusted by discontinuity orientation rating in accordance with the RMR guidelines.

Additionally, due to the high overburden and weak rock mass conditions, the field stress corrections were applied for the data points of the aRMR, where necessary. In Q' rating, SRF and $J_w$ parameters are disregarded hence it only represents the facts of the structural quality of the rock mass. Since GSI represents the structural quality only, it is expected to be correlated with Q' with high $R^2$, which is not the case in this study. Unexpectedly, the parameter Q' is strongly correlated with the ratings aRMR and bRMR represented by a higher value of $R^2$ when compared to the rest of the parameters used in the correlation study. Among the correlations with the GSI, the bRMR has the greatest correlation coefficient. In the correlation of aRMR and GSI, the scatter is greater than the others. For the parameters Q and Q', although satisfactory $R^2$ values are obtained, considerable scatter is clearly visible in the correlations. Several correlations are presented in this section to discuss and evaluate the issue and present the performance of correlations. However, only the suggested equations are given in Table 2.

In Table 2, calculation of the Q or Q' is suggested first then the RMR ratings should be calculated. The ambiguity in the field stress adjustment of the RMR, especially for squeezing conditions, led to the abovementioned suggestion. Another concern is the performance of RMR for the representation of very poor rock masses. While Q lies between 0.001 and 0.01, Basic RMR or, discontinuity adjusted RMR does not fall in the range between 0 and 20 for the same rock mass conditions. This study reveals that, in very poor rock masses, it is suggested to acquire Q rating first and then aRMR rating should be correlated. The correlation of bRMR with GSI had the highest $R^2$ among the other GSI relationships. Still, the scatter of data is significant when GSI is considered.

| Table 2. Suggested correlations for the rock mass classification systems |
|---------------------------------|--------|
| Correlation                     | $R^2$  |
| aRMR = 6.645 lnQ + 45.50        | 0.87   |
| aRMR = 6.696 lnQ' + 32.10       | 0.84   |
| bRMR = 0.698 GSI + 11.97        | 0.83   |

Following the regression work, three NN fitting models were constructed for aRMR-Q, aRMR-Q', and bRMR-GSI. All of the three prediction models performed better than regressions with small error distribution and high $R^2$ values, (Figure 5). The predicted (target)-estimated (output) plots are given for the complete data, not individually for the sets of training and validation.

When the error histograms are inspected (Figure 5.a-c), very small errors (<0.1) are observed. The error values are insignificant since predicted RMR or GSI ratings do not have such small sensitivity.
That is to say, predicted-estimated $R^2$ values are close to the possible maximum. Generally, other statistical parameters are also used to measure the performance of NN models, such as Root Mean Square Error or Variation Accounted For.

In this study, it is concluded that employing such effort is not necessary due to the high $R^2$ values and insignificant error, (Figure 5.d-f). However, the NN fitting tool utilizes a very complex algorithm and background thus, an equation cannot be produced. Instead a MATLAB compatible file can be constructed.

4. Discussion and Conclusion
The ratings, basic and adjusted RMR, Q, Q’ and GSI were correlated and all derived equations resulted in having higher $R^2$ values than the previous studies in the same field. Although all of the correlations converged to $R^2$ values greater than 0.73, only three of them are suggested due to having greater $R^2$ values and relatively low scatter. The correlations in the form of simple regressions seem to be satisfactory but the scatter seems to be considerable and in particular cases the difference between estimated and predicted values can be crucial. In fact, the best practice is the collection of data on site for all classification systems and the computation of each rating separately. In such cases, the correlations can be used for verification purposes. However, the equations given in this study can be expected to best suit the Pliocene-aged formations of Soma coal basin due to the obtained data. In the available data, poor end ratings are present but it is difficult to represent the same class by the rating bRMR especially if the rock mass is very weak and the depth is great. In such cases, Q-system fits the problem better. Alternatively, the field stress adjustment is necessary for RMR. This observation
can be verified by having greater $R^2$ values for the correlations of aRMR and Q. The Neural Network application is found to be superior to the regressions for this research study as the error values of the NN is smaller and practically negligible when compared to regression.

Acknowledgement
The author thanks to Polyak Eynez Energy Mining A.Ş., Fina Energy, and its management group for granting permission to use necessary data. The author presents his gratitude to the Geological Engineers of Polyak Eynez, Feridun Emre Yaşgılı, Ali Türkoğlu, and Mehmet Kılıç for providing extensive data on the geology of the area, and their additional help during the data collection.

References


[29] Abad, J., Celada, B., Chacon, E., Gutierrez, V. and Hidalgo, E., 1983,


