ESTIMATION OF UNIAXIAL COMPRESSIVE STRENGTH BASED ON REGRESSION TREE MODELS

PROCJENA JEDNOOSNE TLAČNE ČVRSTOĆE POMOĆU MODELA BAZIRANIH NA REGRESIJSKIM STABLIMA

ZLATKO BRIŠEVAC[1)], DRAGO ŠPOLJARIĆ[1)], VLATKO GULAM[2)]

[1)] *University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, p.p. 390, 10000 Zagreb*
[2)] *Croatian Institute of Geology, Sachsova 2, p.p. 268, 10 000 Zagreb.*

Ključne riječi: Procjena, regresijsko stablo, slučajna šuma, karbonati, jednoosna tlačna čvrstoća, indeks čvrstoće, Schmidtova tvrdoća, brzina prolaza ultrazvučnog P-vala

Key words: estimation, regression tree, random forest, carbonates, uniaxial compressive strength, point load strength index, Schmidt rebound hardness, P-wave velocity

Sažetak
Ovaj rad bavi se procjenom jednoosne tlačne čvrstoće za karbonate tipa madston-vekston. Potreba procjene javlja se zbog nemogućnosti ispunjavanja propisane visoke kvalitete obrade uzoraka kod direktnog određivanja tog fizikalno-mehaničkog svojstva na nekim vrstama stijena. Za potrebe modeliranja, u ovom radu, ispitivan je intaktni stijenski materijal sa šest mjesta u Hrvatskoj. Ispitane značajke su: gustoća, efektivna poroznost, indeks čvrstoće, Schmidtova tvrdoća, brzina prolaza ultrazvučnog P-vala te jednoosna tlačna čvrstoća koja je bila i ciljana vrijednost procjene uspostavljenih modela. Prikazani modeli su načinjeni na temelju višestruke regresije i regresijskog stabla, a provedena unakrsna validacija, pokazala je kako najuspješniju procjenu jednoosne tlačne čvrstoće daje model slučajnih šuma (engl. random forests).

Abstract
This paper presents the estimation of the uniaxial compressive strength for mudstone and wackestone carbonates. The need for the estimation has occurred due to inability to fulfil the high quality requirements of sample treatment during direct determination of this physical and mechanical property on certain types of rocks. For the needs of modelling intact rock materials, extracted from six locations in Croatia, were tested. The following properties were examined: density, effective porosity, point load strength index, Schmidt rebound hardness, P-wave velocity and uniaxial compressive strength which was the target value of the used statistical models. The statistical models based on multiple linear regression and regression trees were considered and compared using cross validation, which showed that the most efficient estimation of the uniaxial compressive strength is obtained using random forests

1. Introduction

Uniaxial compressive strength (σ_c) is defined as the maximum load i.e. the strain in one direction with a free lateral expansion which can be borne by the material. It represents one of the most essential mechanical characteristics of the rock material used for designing, analysis and modelling. According to Cargill and Shakoor (1990) almost all the engineering tasks use the uniaxial comprehensive strength, such as classification and estimation of strengths and deformabilities of rock masses and various calculations which include the uniaxial compressive strength.

Therefore, the testing procedure of uniaxial compressive strength is one of the most common testing procedures in rock mechanics. Such testing requires a high-quality processing of samples. At first, it was assessed by using simple tests such as the point load strength index - $I_{S(50)}$, density, porosity,

Schmidt rebound hardness - SRH and the velocity of the ultrasonic P-wave.

Beside them, the estimation comprises newly developed tests, in order to indirectly determine the uniaxial compressive strength in the simplest way. The most important tests are National Coal Board - NCB cone indenter and Shore scleroscope hardness tests, which are carried out upon excavations in tunnels and underground mining of mineral raw materials (Tiryaki, 2008), followed by the Block Punch Index test – BPI (Schrier, 1988), the Core Strangle Test - CST (Yilmaz, 2009) and the Nail Penetration Test – NPT, which was invented by Kayabali and Selcuk (2010).

In addition to such indirect tests, the estimation models were made based on the results of laboratory tests. Different estimation models as input parameters use different combinations of simple tests and are based on regression models (Alvarez Grima et al., 1999; Gokceoglu et al., 2004; Karakus et al., 2006; Kahraman et al., 2008; Yilmaz et al., 2008; Dehghan

et al., 2010; Manouchehrian et al., 2012), fuzzy logic (Alvarez Grima et al., 1999; Meulenkamp et al., 1999; Finola et al., 2001; Singh et al. 2001; Gokceoglu 2002; Gokceoglu et al., 2004; Sonmez et al. 2004; Karakus et al., 2006), artificial neural networks (Meulenkamp et al., 1999; Kahraman et al., 2008; Yilmaz et al. 2008; Dehghan et al., 2010; Manouchehrian et al., 2012) and evolutionary programming (Baykasoglu et al., 2008; Ozbek et al., 2013; Beike et al., 2013).

The method of regression tree was applied by Tiryaki (2008) upon estimating the uniaxial compressive strength of intact rock material with the machine excavation of rocks. The input data comprised density, NCB cone indenter and Shore scleroscope hardness. The model was based on testing of 44 samples that had a wide range of strength, from very soft to very hard rocks. According to the available literature, the models based on regression tree have not yet been applied in the case where the input data include the results of testing of density, effective porosity, point load strength index, Schmidt rebound hardness and velocity of ultrasonic P-wave, nor in the case of modelling for a narrow type of materials, such as certain types of limestone.

It should be pointed out that the testing of uniaxial compressive strength has generated a large variation of results of the tested samples of the same type of materials. Accordingly, such a fact makes it challenging to create an estimation model. However, the comparison of the models based on the regression tree and the multiple linear regression thus becomes even more interesting.

2.   Performed tests

For the purposes of the model creation the test results on carbonates mudstone-wackestone species collected from six locations on the Croatian territory (Figure 1) were used and are presented in this paper. The tests included 30 samples. Since this is the same type of carbonate, which in the structure does not have grainy support (Tišljar, 1994), the differences in the physical and mechanical sense are more stressed.
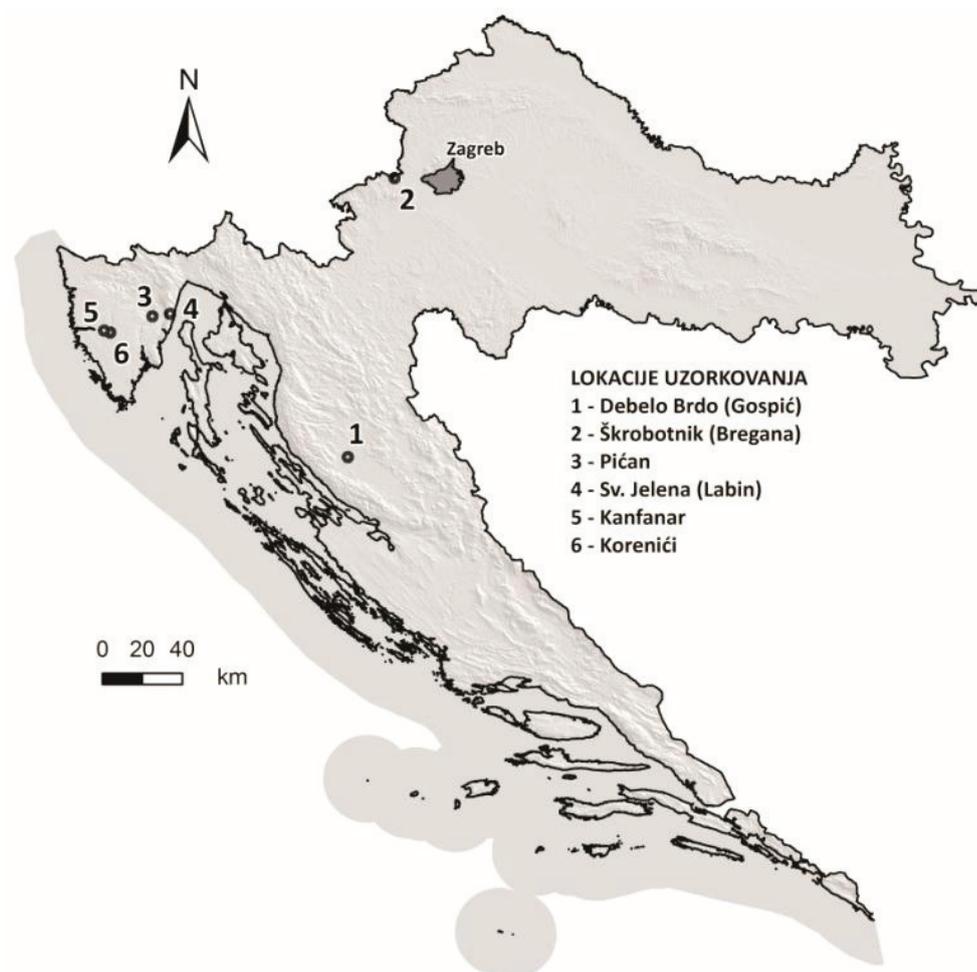

Figure 1: Sampling locations

Slika1: Mjesta uzimanja uzoraka

Rud.-geol.-naft. zb., Vol. 29, 2014.
Z. Briševac, D. Špoljarić i V. Gulam: Estimation of uniaxial compressive strength based on regression tree models

41

According to the recommendations of the International Society for Rock Mechanics the following characteristics were determined on the samples: uniaxial compressive strength (ISRM 1979), density (ISRM 1979b), effective porosity (ISRM 1979), point load strength index (ISRM 1985), Schmidt rebound hardness (ISRM 1978) and the velocity of the ultrasonic P-waves (ISRM 1978b).

The basic descriptive statistics of the results of the test are shown in Table 1. The size of the standard deviation and coefficient of variation for all the methods of testing indicate that this is a data set whose values vary greatly, which is an aggravating fact, but also a challenge to build a model.

Table 1*: Test results*

Tablica 1: Rezultati provedenih ispitivanja

	σ_c (MPa)	ρ (kg/m³)	n (%)	$I_{S(50)}$ (MPa)	SRH	v_p (m/s)
1	95,73	2695	0,98	4,12	55,8	5718
2	53,94	2698	0,29	3,9	64,9	4769
3	110,52	2699	0,35	3,47	65,1	6346
4	58,93	2696	0,26	2,58	60,5	6216
5	69,88	2694	0,42	2,46	64,4	5947
6	111,3	2674	1,15	5,42	62,4	5958
7	121,17	2673	1,13	5,3	62,4	5955
8	125,64	2698	0,74	5,2	63,8	6145
9	178,19	2686	0,85	5,4	63,8	6096
10	169,45	2684	0,99	4,7	63,8	6089
11	33,1	2164	16,46	2,1	51,4	4424
12	56,97	2195	18,43	2,4	51,4	4506
13	98,17	2678	1,63	3,96	56,7	5777
14	103,94	2692	0,64	4,96	57,1	5159
15	96,47	2778	2,04	3,01	67,8	5802
16	151,26	2683	0,14	5,39	70,7	6338
17	170,6	2685	0,12	2,03	68	6303
18	84,03	2689	0,59	1,94	61,6	5660
19	129,93	2640	3,06	4,4	63	5983
20	126,02	2634	2,43	3,6	63	5932
21	136,1	2638	2,67	3,9	63	5992
22	118,81	2634	2,82	2,1	63	5886
23	109,75	2634	2,29	3,7	63	5880
24	122,46	2652	2,14	4,03	63,8	6057
25	87,49	2635	2,59	4,1	62,7	6054
26	109,68	2675	1,75	4,64	62,4	5884
27	115,26	2677	1,25	4,65	62,4	5909
28	121,27	2676	1,23	3,81	62,4	5896
29	152,37	2684	0,64	5,5	63,8	6149
30	172,93	2688	0,84	4,6	63,8	6136
x_{min}	33,1	2164	0,12	1,94	51,4	4424
x_{max}	178,19	2778	18,43	5,5	70,7	6346
$\overline{x}$	113,05	2644,27	2,36	3,91	62,26	5832,2
s	36,65	129,66	4,2	1,14	4,2	488,02
C_v	32,42	4,9	177,97	29,16	6,75	8,37

σ_c uniaxial compressive strength; ρ density; n effective porosity; $I_{S(50)}$ point load strength index; SRH Schmidt rebound hardness; v_p P-wave velocity; x_{min} minimum value; x_{max} maximum value; $\overline{x}$ mean value; s standard deviation; C_v coefficient of variation (%)

The Figure 2 shows the differences in the range of predictor values. Porosity is specifically stressed, due to the fact that it has a considerably higher variation coefficient than other predictors.
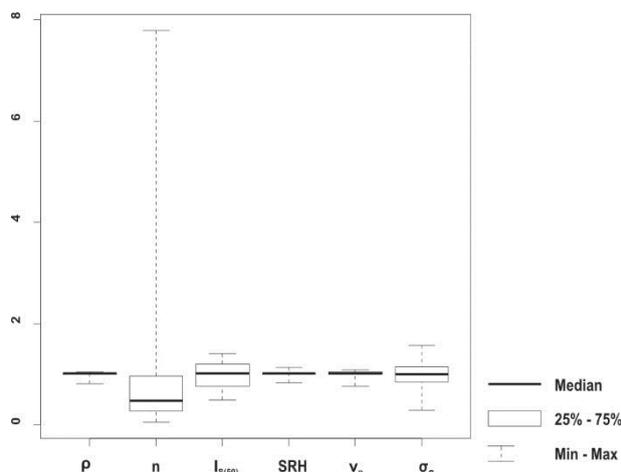
Figure 2: Boxplot for input data divided by mean

Slika 2: B-P dijagram za ulazne podatke normalizirane sa srednjom vrijednošću

3.   Methodology

For the estimation of the uniaxial compressive strength, four statistical models were compared in this paper: multiple linear regression, regression tree and two models based on regression trees known as *bagging* and *random forest*.

Modelling and data analysis were made in the R statistical environment (http://www.r-project.org), which is a standard tool for the statistical analysis of data in the academic community. Besides, it is available as free software. Some of the main advantages are powerful programming language and availability of numerous packages for various statistical analyses often unavailable in commercial software. A constant development in various areas in science and the academic community around R project make this environment useful for statistical data analysis. In this paper, the version 3.1.0 on 32-bit Windows 7 operating system was used.

3.1. Cross-validation and some other measures for model comparison

Several different statistical models will be used for estimation of the uniaxial compressive strength and therefore it is necessary to have a method to choose the best among the used models.

The standard method for estimating the performance of a predictive model is cross-validation which splits the data set in a training and test set of predetermined number of times during the procedure. Since the data set in this paper is relatively small, the *leave-one-out cross-validation (LOOCV)* is used that holds out a single observation from the data set for the test set, while the remaining observations make

up the training set which is then repeated for every observation in a given data set (James et al., 2014).

Precisely, if $\{(x_1,y_1),...,(x_n,y_n)\}$ denotes observations in a given data set, where x_i denotes the ith input parameter, and y_i i^{th} output observation (in this paper σ_c), then LOOCV estimate for the test error is given by:

$$CV_{(n)} = \frac{1}{n}\sum_{i=1}^{n}MSE_i \qquad (1)$$

$$MSE_i = (y_i - \hat{y}_i)^2 \qquad (2)$$

where:

$CV_{(n)}$ – LOOCV estimate for the test mean squared error

MSE_i – mean squared error for the ith observation

n – number of observations

$\hat{y}_i$ – predicted value for y_i using x_i and statistical learning method fitted on the remaining n-1 observations i.e. on training data set

y_i – observed i^{th} value

This paper reports the square root of $CV_{(n)}$ known as *root mean square error of cross-validation* (RMSECV) given by:

$$RMSECV = \sqrt{CV_{(n)}} \qquad (3)$$

The quantity which is often used as a measure of a fit of the model to the data is the coefficient of determination, denoted by R^2 and given by (4) (James et al., 2014)

$$R^2 = 1 - \frac{RSS}{TSS} \qquad (4)$$

$$TSS = \sum(y_i - \bar{y})^2 \qquad (5)$$

$$RSS = \sum(y_i - \hat{y}_i)^2 \qquad (6)$$

where:

$y_i - \hat{y}_i$ is called i^{th} residual

R^2 – coefficient of determination

RSS – residual sum of squares

TSS – total sum of squares

$\bar{y}$ - sample mean

$\hat{y}_i$ - predicted value for y_i using x_i and statistical learning method

TSS measures the total variance in the response variable, while *RSS* measures the amount of variability that is left unexplained after applying the statistical learning method.The R^2 statistic takes values between 0 and 1, where the values close to 1 indicate that a large proportion of the variability in the response has been explained by the fitted model. Still, it is well known that a fitted model with R^2 very close to 1 can show a very poor prediction

performance on the test data. Therefore, it is advisable to have more measures of a model performance which can serve as the basis for the model selection.

One such measure is *adjusted R^2* that penalises an increase of the parameter numbers in the model, and it is given by:

$$adjusted\ R^2 = 1 - \frac{RSS/(n-d-1)}{TSS/(n-1)} \qquad (7)$$

where:
RSS = residual sum of squares
TSS = total sum of squares
n – number of observations
d – number of parameters in a model.

For more details about the given measures for model comparison and some other alternatives we refer to James et al. (2014).

3.2. Multiple linear regression

Multiple linear regression is a very straight-forward and common approach for predicting a quantitative response Y on the basis of multiple predictor variables $X_1,..., X_k$. It assumes that there is approximately a linear relationship between the response Y and the predictors $X_1,..., X_k$. Therefore, the multiple linear regression model takes the form given in the equation (8)

$$Y = \beta_0 + \beta_1 X_1 + ... + \beta_k X_{k} + \varepsilon \qquad (8)$$

where:
Y – response variable
ε – mean-zero random error term
$X_1,..., X_k$ - predictors
$\beta_0,...,\beta_k$ - regression coefficients.

Regression coefficient β_j quantifies the association between the predictor X_j and the response variable Y. More precisely, β_j is interpreted as the average effect on Y of one unit increase in X_j, holding all other predictors fixed (James et al., 2014).

In this paper, multiple linear regression models for all subsets of five predictors (31 subsets in total) are built. The Figure 3 shows the ranking of those 31 models according to the adjusted R^2 scores (vertical axis). Every row represents one particular model, and the predictors (horizontal axis) included in the model are denoted by the black square, while the white square means that that particular predictor is not included in the model.

Further, all the constructed models are then compared using the leave-one-out cross-validation, and ranked according to the RMSECV. The LOOCV ranking is shown in Figure 3 in the LOOCV RANK column.
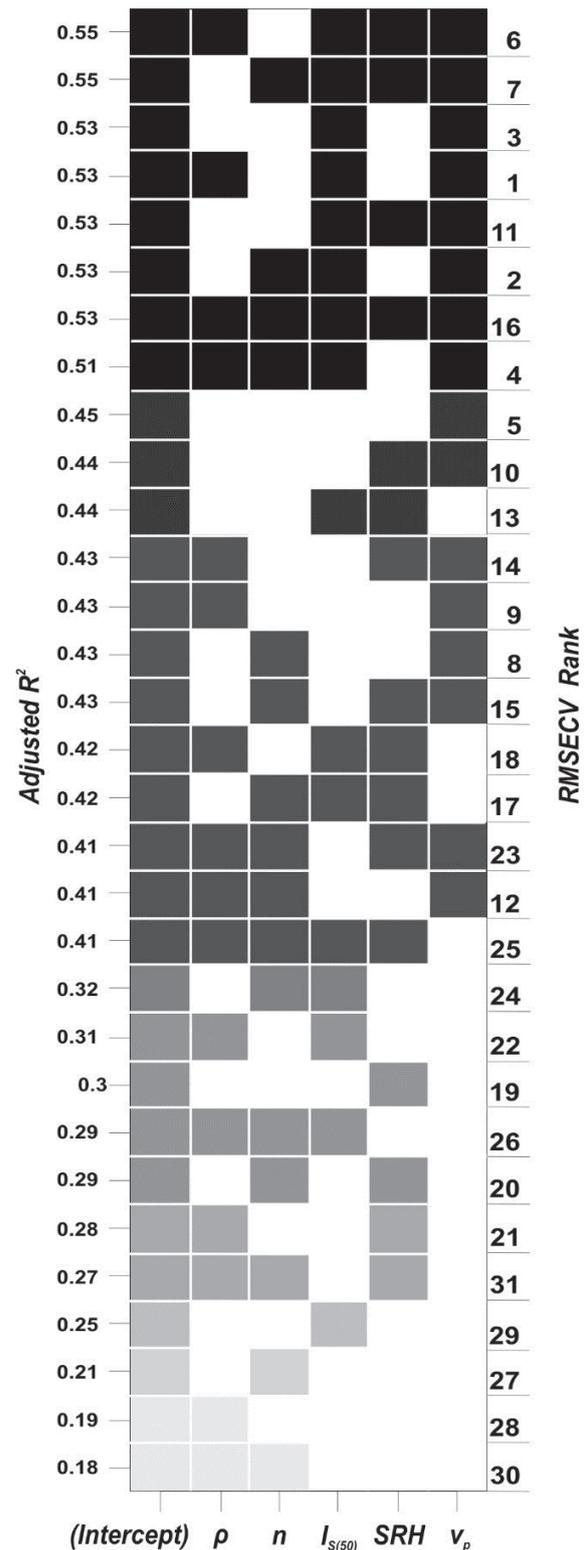

Figure 3 Rank of multiple linear regression models according to the adjusted R2 and RMSECV

Slika 3: Višetruki linearni regresijski modeli sortirani prema vrijednosti prilagođenog R^2 i RMSECV

Note that the first two models according to RMSECV are the 4th and the 6th models according to the adjusted R^2, respectively. However, notice that

the model with all the predictors included, takes up the 7th place according to the adjusted R^2 (and the first one according to R^2) while it is on the 16th position according to RMSECV. Therefore, the right measure for model accuracy should be carefully chosen.

The equations for the first two of the best models according to RMSECV are given in (9) and (10),

$$\sigma_c = -106{,}2093 - 0{,}04868 \, \rho + 11{,}5110 \, I_{S(50)} + 0{,}052 \, v_p \tag{9}$$

$$\sigma_c = -240{,}0109 + 1{,}5087 \, n + 11{,}5916 \, I_{S(50)} + 0{,}0522 \, v_p \tag{10}$$

where:

σ_c – uniaxial compressive strength (MPa)
ρ – density (kg/m^3)
n – effective porosity (%)
$I_{S(50)}$ – point load strength index (MPa)
v_p – P-wave velocity (m/s).

3.3. Regression trees

Regression tree is a model which is simple and useful for interpretation, but typically is not competitive, in terms of prediction and accuracy, with the best supervised learning approaches, including sometimes even linear regression (James et al., 2014).

Nevertheless, in this paper the analysis is conducted also using regression trees, since they are simple for interpretation and form a basis for other modern techniques, like bagging and random forests.

The process of building a regression tree first divides the predictor space into the M distinct and non-overlapping regions $R_1,..., R_M$, and then, for every observation that belongs to the region R_j, the response variable (σ_c) is predicted by the mean of the response values for the testing observations in R_j (James et al., 2014). The question is: „How are regions $R_1,..., R_M$ chosen?" The goal is to find high-dimensional rectangles $R_1,..., R_M$ that minimize RSS given by

$$\sum_{j=1}^{M}\sum_{i \in R_j}(y_i - \hat{y}_{R_j})^2 \tag{11}$$

where $\hat{y}_{R_j}$ is the mean response for the training observations within the rectangle R_j. Since, it is computationally very demanding to consider every possible partition of the predictor space into *M* rectangles, the standard approach, to that minimization problem, is a greedy algorithm known as *recursive binary splitting*. As the name of the algorithm suggests, at each step of the tree-building process, the predictor X_j and the cutpoint *t* are chosen such that splitting the predictor space into the regions $\{X|X_j < t\}$ and $\{X|X_j \geq t\}$ leads to the greatest possible reduction in RSS. Here the notation $\{X|X_j \geq t\}$ denotes the region of predictor space in which X_j is at least *t*.

The building process stops when a stopping criterion is reached – usually, until no region contains more than five observations. For more details about regression trees we refer to James et al. (2014).

A tree produced by the procedure described above is likely to overfit the data, leading to the poor test set performance (James et al., 2014). The reason is the fact that the resulting tree might be too complex. Therefore, there is an algorithm, called *tree pruning*, which chooses a subtree (of a tree built by the above procedure) which will have a lower test error rate than the full tree. Again, for more details about tree pruning we refer to James et al. (2014).

If the predictor values are denoted by x and the value of response variable by *f*(x), then regression trees assume a model of the form:

$$f(x) = \sum_{i=1}^{M} c_i \cdot 1_{\{x \in R_i\}} \tag{12}$$

where:

$1_{\{x \in R_i\}}$ - indicator function which takes value 1 if x is in R_i or value 0 otherwise
c_i – the value of response variable for the region R_i.

In other words, regression tree can be interpreted as a sequence of YES/NO questions where the answer YES leads to the left branch in the tree, while the answer NO leads to the right branch in the tree. If the process reaches a leaf, then a fitted response value is used as prediction of the response variable. Otherwise, the process proceeds to another question until it reaches a leaf and a fitted response value.

The Figure 4 shows the regression tree constructed in R environment, using the package „tree"*,* version 1.0.35. with default values for all parameters, while the value of the parameter „best", in function *prune.tree*, is set to 4.
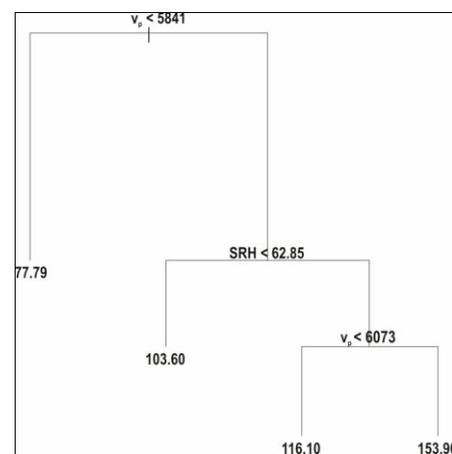

Figure 4: *Regression tree model for estimation of the uniaxial compressive strength*

Slika 4: Regresijsko stablo za procjenu jednoosne tlačne čvrstoće

Rud.-geol.-naft. zb., Vol. 29, 2014.
Z. Briševac, D. Špoljarić i V. Gulam: Estimation of uniaxial compressive strength based on regression tree models

45

As shown in Figure 4, the algorithm for building a regression tree, has chosen the predictors v_p and SRH, among all predictors, as the two most important predictors upon determination of uniaxial compressive strength.

3.4. Bagging

Bagging (the abbreviation of *bootstrap aggregation*) is a general-purpose procedure which uses regression trees as building elements during the creation of prediction model (James et al., 2014). One of the major problems with regression trees is that the resulting fitted tree strongly depends on training data. Namely, if the training data are split into two parts randomly, and a regression tree is fitted to both halves, the results could be quite different.

Bagging is a method which tries to reduce those variability by taking a large number of samples from the training data (using sampling with replacement), building a prediction model on every sample and averaging the resulting predictions.

Suppose that, using the described approach, B different bootstrapped training data sets are generated. Further, on every training set a regression tree is fitted and $\hat{y}_1,\ldots,\hat{y}_B$ resulting predictions are obtained, which are then averaged in order to obtain a single prediction of response (denoted by $\hat{y}_{bag}$) variable with bagging procedure. Precisely

$$\hat{y}_{bag} = \frac{1}{B}\sum_{i=1}^{B}\hat{y}_i \qquad (13)$$

where:
$\hat{y}_{bag}$ – prediction of response variable with *bagging* procedure
B – the number of generated samples using sampling with replacement
$\hat{y}_i$ – predicted response variable using i^{th} regression tree

3.5. Random forests

Although bagging could improve the prediction accuracy of regression trees substantially, the problem is that constructed trees could be highly correlated, in a sense that they will look quite similar to each other (for instance, if there is one very strong predictor in the data set, while other predictors are moderately strong, then most of the trees will use this strong predictor in the top split – in particular, in this paper, the predictor v_p is such a strong predictor).

Random forests try to avoid the above mentioned problem in the way that each time a split in a tree is considered, a random sample of m predictors is chosen as the split candidates from the full set of p predictors (James et al., 2014). Typically, the chosen m is approximately equal to $\sqrt{p}$ (in the analysis in this paper $p = 5$, and the best results are obtained for

$m=2$). Using a small value of m in building a random forest could be helpful if p is large and if there is a large number of correlated predictors.

In contrast to regression trees which can be easily interpreted, a collection of bagged trees and random forest are much more difficult to interpret and it is no longer possible to represent the result of the learning procedure as a single tree. Despite that, some information about the importance of each predictor can be obtained using the *RSS*, where the predictor is more important if the total amount of the *RSS* that is decreased due to splits over a given predictor, averaged over all trees, is larger (see James et al. 2014. for more details).

The Figure 5 displays the importance of predictors in random forest constructed during the analysis in this paper. It can be observed that the most important predictor, in a sense described above, is the P-wave velocity (v_p), while effective porosity is the least important.
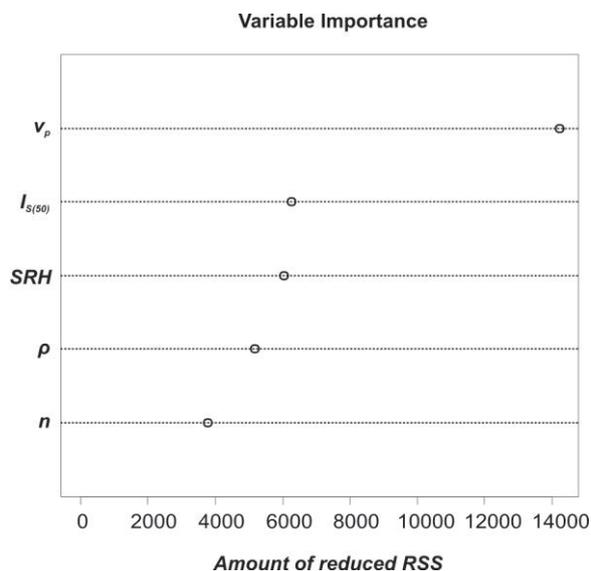

Variable Importance

Figure 5: Importance of variables in a random forest due to the amount of reduced RSS

Slika 5: Važnost varijabli u slučajnoj šumi s obzirom na iznos smanjenog *RSS*-a

The analysis with random forests and bagging is conducted in R environment using the packages „randomForest", version 4.6.10, with default parameters. The best result with random forests is obtained with parameter *mtry* (in function *randomForest*) set to 2, while in *bagging* that parameter is set to 5.

4. Discussion

This paper presents the estimates of the uniaxial compressive strength using multiple linear regression (as the most commonly used model), including the comparison of the models based on regression trees.

In case of multiple linear regression, the models for all possible combinations of 5 predictors are constructed (a total of 31 models). They are compared and ranked using a leave-one-out cross-validation (LOOCV). The worst LOOCV result is obtained by the model which uses density, effective porosity and Schmidt rebound hardness as predictors, while the model with all the predictors included is ranked as the 16th. On the other hand, the best LOOCV result is achieved by the models (9) and (10). It should be pointed out that the model with the best R^2 (for instance, the model with all predictors), has a rather poor LOOCV result compared to other models. This indicates that the measure which serves as benchmark for model accuracy should be carefully selected.

On the other side, regression tree has a larger R^2 than every linear regression model and a considerably larger adjusted R^2, but it has a significantly smaller result on the cross-validation test (Table 2). However, the methods based on regression trees show substantially better results regarding all three measures (R^2, adjusted R^2 and RMSECV). The results in Table 2 show that, under compared models, the best fit to the training data and the most accurate prediction of the uniaxial compressive strength are obtained using random forests.

Table 2*: Results of cross-validation for models based on regression trees with corresponding values of R^2 and adjusted R^2*

Tablica 2: Rezultati unakrsne validacije modela baziranih na regresijskom stablu te vrijednosti mjera R^2 i prilagođeni R^2

Models	**RMSECV**	R^2	adjusted R^2
Regression tree	35,34	0,6152	0,5867
Bagging	26,55	0,9055	0,8858
Random Forest	23,99	0,9081	0,8934

RMSECV root mean square error of cross-validation; R^2 coefficient of determination

The application of the models presented in this paper should take into account the range of the uniaxial compressive strength and predictor values (see Table 1).

5.   Conclusions

Based on the test results of density, effective porosity, point load strength index, Schmidt rebound hardness and P-wave velocity, some statistical models are compared, in order to estimate the uniaxial compressive strength.

One of the goals is to compare the most commonly used model – multiple linear regression – with more modern statistical models, which have an interesting and simple interpretation (such as regression tree), or show a better estimation accuracy, but are much more difficult to interpret.

Despite a high dispersion of data in the training data set, it has been determined that linear regression models show good results (for instance, models (9) i (10)), but less accurate than the models based on regression trees – bagging and random forests.

Since the data set was relatively small, the standard technique known as leave-one-out cross-validation was used as the measure of estimation accuracy. It repeatedly splits the data set into training and test set, which are used for fitting and testing statistical models.

Among the compared models, the model of random forest shows the best estimation accuracy, as well as the best fit to the training data – meaning that it has the highest proportion of variability in the response variable, explained by the model.

6.   Acknowledgment

7.   References

Alvarez Grimaa, M. and Babuška, R., (1999): Fuzzy model for the prediction of unconfined compressive strength of rock samples, Int. J. Rock Mech. Min. Sci., 36, 339-349.

Baykasoglu, A., Gullu H., Canakcı H. and Ozbakır L. (2008): Prediction of compressive and tensile strength of limestone via genetic programming, Expert Systems with Applications, 35, 111–123.

Beiki M., Majdi, A. and Givshad, A.D. (2013): Application of Genetic Programming to predict the uniaxial compressive strength and elastic modulus of carbonate rocks, Int. J. Rock Mech. Min. Sci., 63, 159–169.

Cargill, J. S. and Shakoor, A. (1990): Evaluation of Empirical methods for measuring the uniaxial compressive strength of rock. Int. J. Rock Mech. Min. Sci. Geomech. Abst.,Vol. 27, 495-503.

Dehghan, S., Sattari, G., Chehreh C. S., and Aliabadi, M.A. (2010): Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural Networks. Mining Sci. Technol., 20: 41-46.

Finol, J., Guo, Y.K., Jing, X.D., 2001. A rule based fuzzy model for the prediction of petrophysical rock parameters. J. Pet. Sci. Eng., 29, 97–113.

Gokceoglu, C. (2002): A fuzzy triangular chart to predict the uniaxial compressive strength of the Ankara agglomerates from their petrographic composition. Eng. Geol., 66, 39–51.

Gokceoglu, C., Zorlu, K. (2004): A fuzzy model to predict the uniaxial compressive strength and the modulus of elasticity of a problematic rock. Engineering Applications of Artificial Intelligence, 17, 61–72.

ISRM (1978a): Suggested methods for determining hardness and abrasiveness of rock; Part 3: Suggested method for determination ot the Schmidt rebound hardness. Int. J. Rock Mech. Min. Sci. Geomech. Abst., Vol. 15, 89-98.

Rud.-geol.-naft. zb., Vol. 29, 2014.
Z. Briševac, D. Špoljarić i V. Gulam: Estimation of uniaxial compressive strength based on regression tree models

47

ISRM (1978b): Suggested method for determining sound velocity, International Journal of Rock Mechanics and Mining Sciences & Geomechanical Abstracts., 15 (2), 53–58

ISRM (1979a): Suggested methods for determining the uniaxial compressive strength and deformability of rock materials. Int. J. Rock Mech. Min. Sci. Geomech. Abst., Vol. 16, 135-140.

ISRM (1979b): Suggested methods for determining water content, porosity, density, absorption and related properties and swelling and slake-durability index properties, Part 1: suggested method for determination of the water content of a rock sample; suggested method for porosity/density determination using saturation and caliper techniques. Int. J. Rock Mech. Min. Sci. Geomech. Abst., 16, 141-156.

ISRM (1985): Suggested method for determining point load strength. Int. J. Rock Mech. Min. Sci. Geomech. Abst., Vol. 22, 51-60.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2014) An Intorduction to Statistical Learning with Applications in R, Springer, 426 str., New York

Kahraman, S., Gunaydin O., Alber M. and Fener M. (2008): Evaluating the strength and deformability properties of Misis fault breccia using artificial neural networks, Expert Systems with Applications, (36) 3, 6874-6878.

Karakus M., and Tutmez B. (2006): Fuzzy and Multiple Regression Modelling for Evaluation of Intact Rock Strength Based on Point Load, Schmidt Hammer and Sonic Velocity, Rock Mech. Rock Engng. 39 (1), 45–57.

Kayabali, K., and Selcuk, L. (2010) Nail penetration test for determining the uniaxial compressive strength of rock, Int. J. Rock Mech. Min. Sci., 47, 265–271

Manouchehrian, A., Sharifzadeh M. and Moghadam R.H. (2012): Application of artificial neural networks and multivariate statistics to estimate UCS using textural characteristics., International Journal of Mining Science and Technology, 22, 229–236.

Meulenkamp, F., Alvarez Grima M., (1999): Application of neural networks for the prediction of the unconfned compressive strength (UCS) from Equotip hardness. Int J Rock Mech Min Sci Geomech, 36(1), 29-39.

Ozbeka, A., Unsal, M. and Dikec A. (2013): Estimating uniaxial compressive strength of rocks using genetic expression programming, Journal of Rock Mechanics and Geotechnical Engineering, 5, 325–329.

Schrier van der, J.S., (1988): The block punch index test. Bull. Int. Assoc. Eng. Geol. 38, 121–126.

Singh, V.K., Singh, D.and Singh, T.N., (2001): Prediction of strength properties of some schistose rocks. Int. J. Rock Mech. Min. Sci., 38(2), 269–284

Sonmez, H., Tuncay, E. and Gokceoglu, C. (2004): Models to predict the uniaxial compressive strength and the modulus of elasticity for Ankara Agglomerate. Int. J. Rock Mech. Min. Sci., 41, 717–729.

Tiryaki, B. (2008): Predicting intact rock strength for mechanical excavation using multivariate statistics, artificial neural networks and regression trees. Eng. Geol., 99, 51–60.

Tišljar, J. (1994): Sedimentne stijene. Školska Knjiga, 399 str., Zagreb

Yilmaz I. and Yuksek G. (2008): Prediction of the strength and elasticity modulus of gypsum using multiple regression, ANN, and ANFIS models.Int. J. Rock Mech. Min. Sci., 46 (4), 803–810.

Yilmaz, I. (2009): A new testing method for indirect determination of unconfined compressive strenght of rocks. Int. J. Rock Mech. Min. Sci., Vol. 46, 1349-1357.