Reservoir Property Modeling Using Machine Learning Approaches – A case study from the Stubensandstein, Germany

Table of Contents

List of Figures And Tables

Abstract

The economic and physical constraint of acquiring a high number of spatial petrophysical data in a reservoir is a general complication in geosciences and related engineering fields. A considerable path to reduce these limitations is by application of machine learning prediction techniques to forecast petrophysical properties in space. The Support Vector Regression (SVR) technique, which searches for an "optimal" hyperplane to solve learning problems machine problems appears to be a suitable approach to deal with these limitations due to the high performance of the technique on non-linear relationships using kernel modification on the data. The SVR exhibits the structural risk minimization (SRM) principle which has been exhibited to be more efficient in comparison to the conventional empirical risk minimization (ERM) adopted by traditional statistical methods and neural network. The objective of this research was to predict petrophysical properties (permeability and porosity) from spatial data (longitude, latitude and elevation) using a machine learning technique. The spatial petrophysical data employed for training were digitized from porosity and permeability plots of a quarry in Baden Wurttemberg, Germany. The result obtained from this study shows that the SVR technique using the Radial Kernel is a better model option compared to the other kernel types in the prediction of petrophysical data. The overall accuracy of the technique for porosity prediction and permeability prediction is an RMSE value of 0.2 and 0.1, AAE value of 0.1 respectively. Results obtained also show that the SVR technique outperforms the ANN technique in spatial petrophysical property prediction.

1. Introduction

The information derived from the study of the physical and chemical properties of rocks such as lithology, porosity, permeability, fluid saturations and pressure, chemical fluid composition, fractional flow etc. can be used in identifying and evaluating hydrocarbon reservoirs, hydrocarbon sources, geothermal reservoirs, seals and aquifers. However, it is difficult to determine the actual value of these properties spatially with the certainty required to make economic decisions for production and development (1). The economic potential of reservoirs is dependent on the connectivity of the geological compartments it is being composed of. If the reservoir provides a low degree of inter-connectivity among the reservoir bodies, the asset value thereof is negatively impacted as more drillings must be implemented in order to exploit the entire reservoir (2).The cut-offs in petrophysical properties have been identified as a major cause of poor estimation of reserves in place in a reservoir, and a holistic approach is to consider the range of missing data in petrophysical parameters and quantify the uncertainty. The scarceness of adequate petrophysical data on a spatial scale is widely observed in reservoir characterization but if available should be integrated in order to enrich information about the target region.

Reservoir property models produced with the aid of petrophysical properties prediction can be used to bridge the gap of missing data in the target-specific reservoirs, which cannot be fully surveyed by well logs due to economic and time implications (3). Moreover, these models can be used for planning exploitation and injection wells in order to maximize reservoir recovery.

Moreover, the accuracy of the prediction is dependent on the quality of data acquisition, errors observed during this acquisition and field experience of the geologist acquiring the data (4). However, to acquire this quality and a large number of data is uneconomic and time-consuming. Therefore, methods are been developed for efficient modeling of spatial data for prediction of the cut-offs in petrophysical properties. Researchers have employed geostatistical models in porosity prediction using cokriging with multiple secondary datasets (5). In recent times researchers have used machine learning methods such as artificial neural networks (ANN) and support vector regression (SVR) models to predict rock properties in space For example, rock properties were predicted in temperate mountain regions spatially with the aid of SVR (6). Rock properties have been linked to having a relationship with geological formations and the spatial position of the formations. However, these relationships are rarely linear in real-life scenarios. SVR and ANN models have been proven to establish the non-linear relationship that exists between these properties and spatial information. SVR established on statistical leaning theory is an efficient technique for prediction and classification functions It employs margin based loss functions to optimize model complexity independent of the input space dimensionality with the aid of kernel functions which project the estimated challenge to a higher dimensional space ,This allows problems of complex non linearity to be solved on a universal scale (7).

The ANN, on the other hand, is a technique from machine learning that originates from the studies of the human neuron and possess qualities to automatically analyze multi-source inputs by producing results without hypothesis using self-learning techniques (8). The main objective of this research was to predict the petrophysical properties such as porosity and permeability by developing an SVR model and an ANN model using spatial information in the form of longitude, latitude, and elevation information.

The detailed objective of this research entails

(1) Extracting spatial data and petrophysical data from permeability and porosity plots produced in the sedimentary field course which has been held over the last 15 year at the TU Darmstadt

(2) Selection of optimum parameters for the SVR and ANN models with model performance examination.

(3) Measure of error for model predictions and further recommendations.

2. Materials and Methodology

2.1. Study Area

The quarry is located north of Esslingen about 2 km towards Stetten. It covers a total mining area of approximately 40 hectares with a processing area of few hundred square metres which is been recultivated. The elevation of the quarry ranges from 404 m to 464 m above sea level. The climate related to the quarry area is very cold winter and comfortable summer which is partly cloudy all year round , the temperature typically varies from -2°C to 25°C and is rarely below -9°C or above 31°C, the rainfall typically varies from high of 7.3 cm total accumulation to a low of 3 cm, the average snowfall is 0.3 cm. The quarry exposes clastic deposits of the Stubensandstein. This unit belongs to the Middle Keuper, a lithostratigraphic group of the Germanic Triassic, which is subdivided into the three groups of Buntsandstein, Muschelkalk and Keuper (16). The coloured sandstone deposits present in the study area were deposited by rivers. The Keuper is dominated by lake playa sedimentation and river transported sedimentation (9). Sedimentation of the Stubensandstein in the intracratonic Keuper basin in the south of Germany occurred under semi-arid to semi-humid climatic conditions (10). The Bohemian Massif in the east and the Vindelizische Land in the southeast of the basin served as source rock areas (11). The dominant poorly sorted fine to coarse-grained arkose sandstones deposits are a result of the short transport routes of the river (12). Sandstones deposited at the edge of the basin intercalated with the playa lake sediments deposited at the basin centre. Generally, the whole sequence of Stubensandstein is composed of several fan-shaped sandstone wedges, that developed 200 km into the basin and up to a thickness of ca 120 m in the basin depending on the climatic setting (13). The fluctuation from semi-arid to more humid climatic conditions contributed to the deposition of sandstones, intermediate marl, silt or claystone layers in a periodical sequence (14). The change did not occur abruptly but fluctuated between dry phases and more humid episodes with monsoon-like rainfall (15). In the proximal section of the alluvial fans, river channels were formed during the wet season which were periodically cut by erosion (16). The torrential rainfall resulted in the rivers flowing on the alluvial fan as a result of heavy sediment flow and high water which spread the sand extensively (12). The Stubensandstein sedimentation cycle is assumed to last around 6 Ma, which results in an average sedimentation rate of 17 mm/ka (17) which represents a so-called "low accommodation system" along with the low subsidence (16). The Stubensandstein extends to the Swabian and Franconian Albvorland, the Swabian Forest, the Löwenstein Mountains and the Stromberg area on the surface. The underground extension is economically relevant and larger compared to the surface extension as it is a natural gas and oil deposits storage rock in the Alpine foothills. Therefore it is relevant to know the spatial distribution of reservoir properties such as porosity and permeability in this formation to maximize its oil and gas recovery (16).

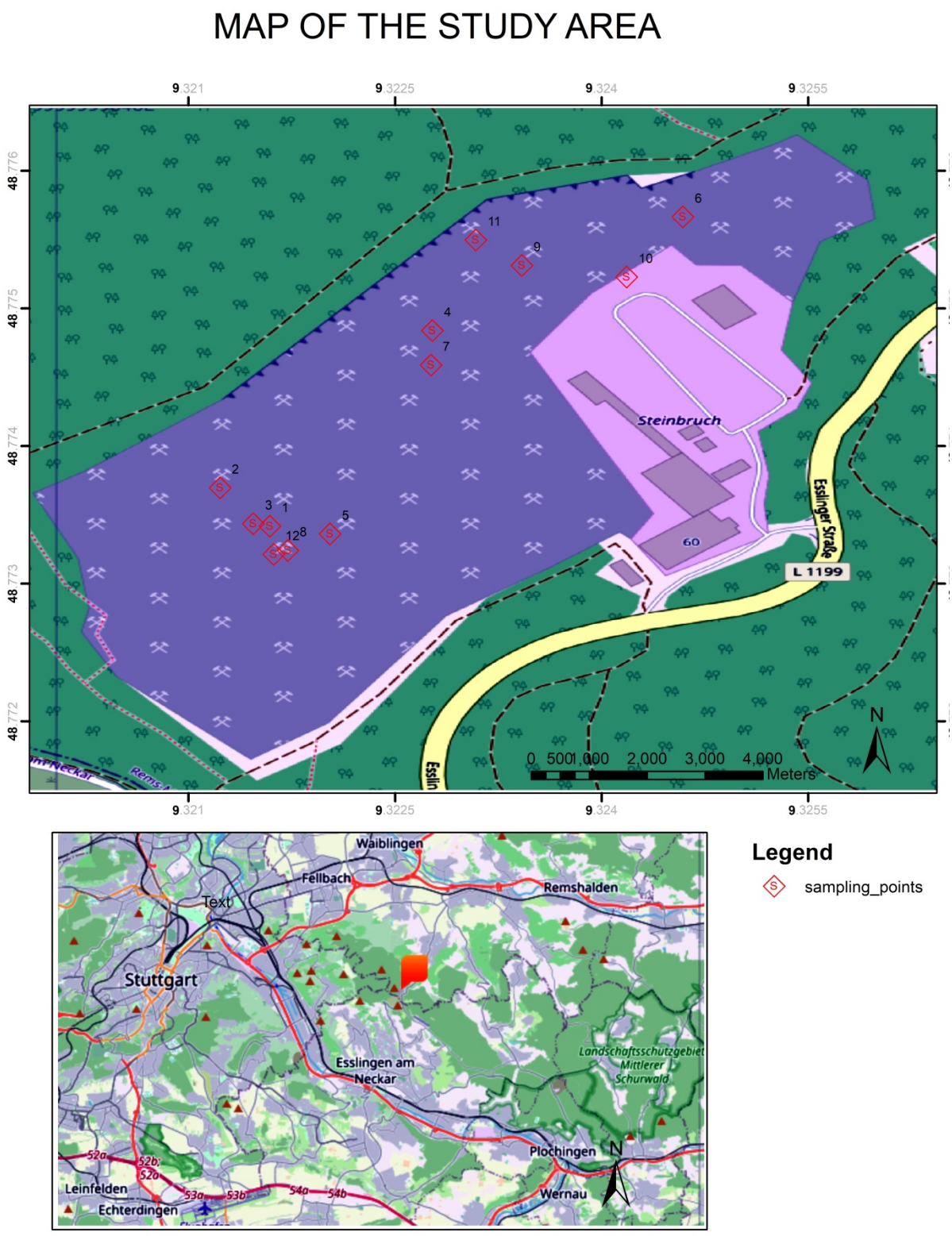

Figure 1: figure showing map of the study area and the sampling points

2.2 Rock Sampling

The investigated quarry is a mining site authorized for mining sandstone. Rock borehole cores were sampled in numerous 1-D sections distributed throughout the entire quarry area to ensure samples are representative for the total population. In the frame of a sedimentary field course, a total of 206 rock samples was collected at 14 1-D sections within the study area with the aid of a handheld drilling machine. These rock samples were further analysed in the laboratory for rock type classification, respective porosity properties using the GeoPyc 1330 pycnometer device and permeability properties using the hassler cell gas permeameter. Two of the borings were reported with coordinates outside of the study area and were omitted for the purpose of this research. The permeability and porosity plots were digitized to obtain the variables required to perform the machine learning techniques.

2.3 Digitization of Variables

The Graphclick software was employed for the digitization of the permeability and porosity plots. Data obtained from the digitization include geographical coordinates, elevation, permeability and porosity. Permeability and porosity graph images were obtained from the sedimentology field practical reports and were imported into the Graphclick software. The y-axis limit and x-axis limits were declared and the elevation limit was also defined. Finally, the permeability and porosity points were digitized and the respective permeability and porosity variables were obtained. The data was previewed and filtered for null variables, the total number of entries, duplicate entries and further plotted to understand the general data trend. A total of 206 points were digitized from the permeability and porosity plots and attributed to 12 different coordinates in the study area with their respective elevation above sea level.

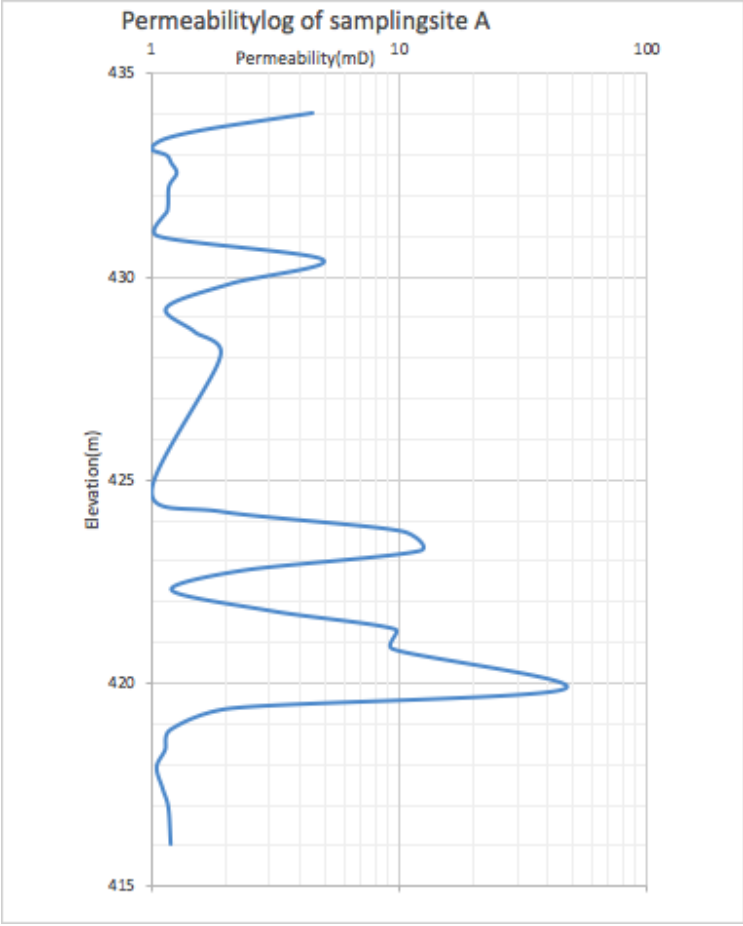

Figure 2 : showing the permeability log at longitude:9.321593, latitude:48.77342 in the study area**.**

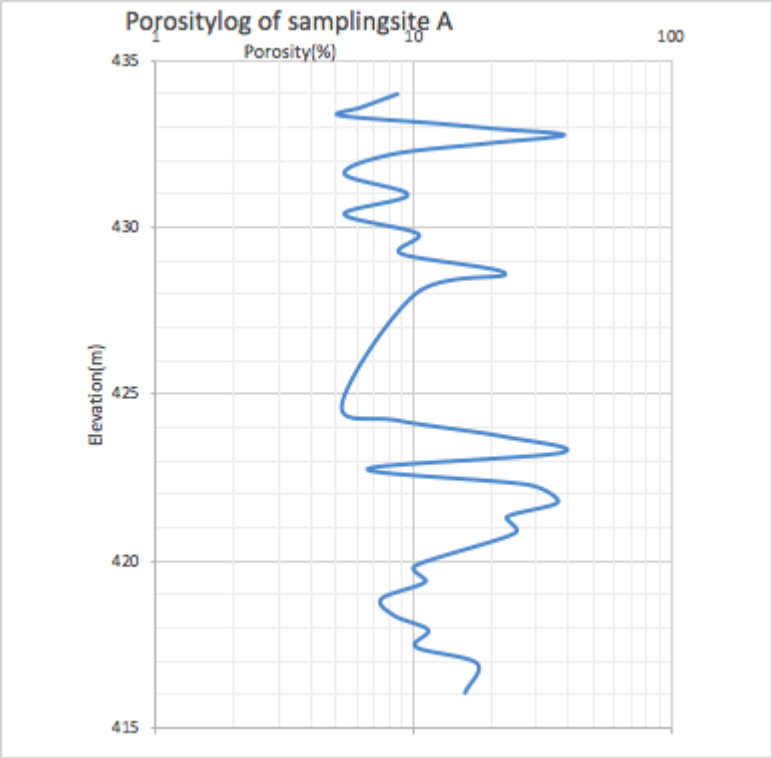

Figure 3: showing the porosity log at longitude:9.321593, latitude:48.77342 in the study area.

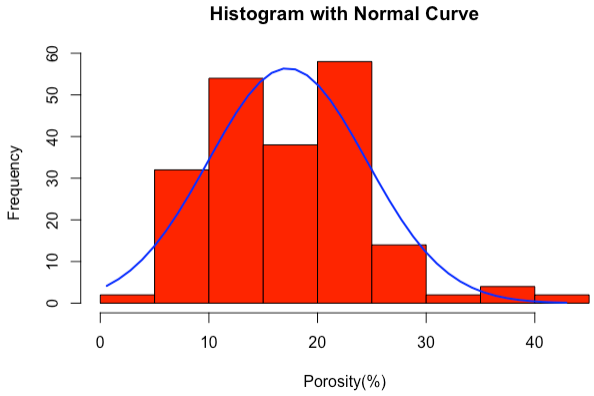

Figure 4: Histogram showing porosity frequency with Normal curve.

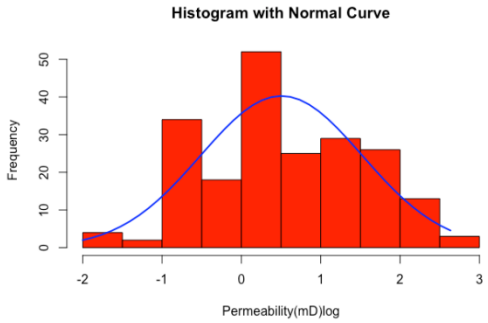

Figure 5: Histogram showing the permeability frequency with Normal curve.

A summary statistic of the data derived from the plot was performed before visualizing the data. The summary statistics in Table 1 show that the elevation data have a very similar median, mode and mean

indicating that it is normally distributed. Visualizing the elevation on a histogram and a normal curve in Figure 2 shows the elevation plot as symmetric with a slight negative skew which was also validated by the summary statistics. The summary statistics in Table1 show that the porosity data have similar mean and median but different mode giving a hint that the data is not normally distributed. Validation with the histogram and normal curve in Figure 3 shows that the porosity data is not normally distributed with a positive skew however logarithmic transformation of the permeability data resulted in a normal distribution. The permeability data show very different median, mean and mode giving an idea that the data is not normally distributed. The permeability histogram with a normal curve in Figure 4 shows the data to be positively skewed

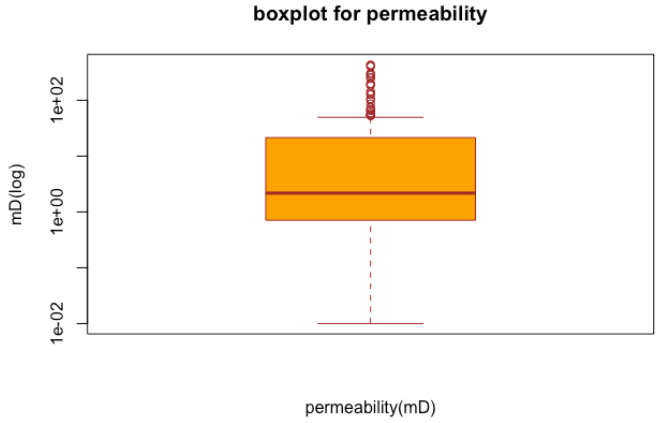

Figure 6: Boxplot showing the upper limit, lower limit, outliers and median of the permeability variable

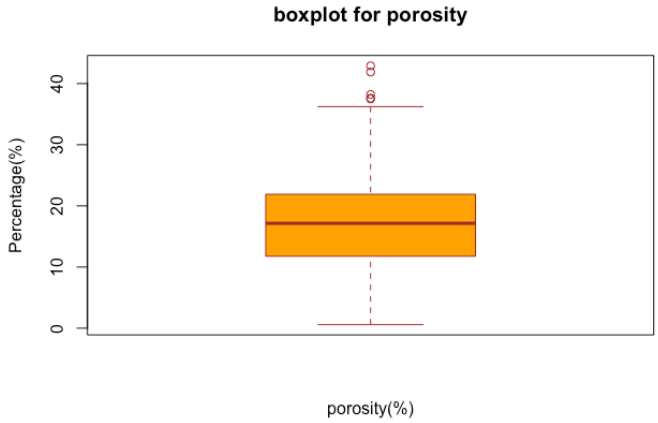

Figure 7: Boxplot showing the upper limit, lower limit, outliers and median of the permeability variable

The boxplot in Figure 6 shows that the 2nd quartile or median of the elevation variable is around 418 m above sea level, 25 percent of the elevation data is below the lower limit (ca. 405 m) and 25% of the elevation data is above the upper limit (ca 427 m) with no outliers. Figure 7 shows the 2nd quartile of the permeability data equal to 2 mD, 25%of the permeability data above 25 mD (75th percentile/upper limit) and 25 percent of the permeability data below 0.8 mD (25th percentile/lower limit), the permeability boxplot also shows numerous outliers occurring above 60 mD. The porosity box plot shows the 2nd quartile of the porosity variable equal to 17%, 25%(lower limit) of the porosity data fall

below 12% and 25% of the porosity data occur above 22% porosity value with 4 outliers occurring beyond 37%. The summary statistics and visualization prove the further need for data processing for the machine learning techniques.

	elevation	permeability	porosity
Number of observations	206	206	206
Mean	416.08	28.8	17.24
Standard deviation	13.16	70.19	7.29
Median	415.97	2.17	17.15
Mean absolute deviation	16.84	3.05	7.6
Minimum value	389.1	0.01	0.6
Maximum value	442.43	431.74	42.9
Range	53.33	431.73	42.3
Skewness	-0.04	3.81	0.55
Kurtosis	-1.05	15.39	0.72

Table 1: Summary statistics of the digitized variables.

2.4 Support Vector Regression

Support Vector Machines (SVM) have originated recently in the framework of statistical learning theory (SLT) otherwise known as Vapnik–Chervonenkis (VC) theory (18) (19), and have been successfully employed in a wide range of fields, from time series prediction (20), to facial recognition (21), to biological data processing for medical diagnosis (22), or the spatial prediction of rock properties (6). The success of these experimental projects and theoretical studies encouraged further research on SVM. SLT gives fundamental and ample conditions for the rapid rate of uniform convergence and distribution independent consistency of empirical risk minimization (ERM) (23). The concept of VC dimension was introduced to regularize the ERM consistency. For instance, in a binary classification learning problem, the function set has a VC dimension **h** if a maximum number of training points is present upon which all combinations of separation can be derived from the function set members accurately (24). According to results obtained from the SLT, the empirical risk was concluded as a safe measure of true risk when the sample size is large. However, large deviations which cannot be ignored occur between the true risk and empirical risk when the sample size is small, these large deviations result in the need of a process to control the learning by assigning complexity of approximating functions to the available finite samples. The process is accomplished with the aid of SVM inductive principle where a structure is defined on set of approximating functions and the VC dimension with low empirical error is chosen (24).

Statistical learning also considers the training examples as identically generated and also independent in respect to a known distribution. The input response sample data are considered a learning tool for the anonymous functional correlation between the output and input variables (23). The SVM approach is established on SLT and it solves constrained quadratic problems with target functions of minimizing the combination of a margin based loss function with a optimization term (25).

2.5 SVM Methodology

With the introduction of the Vapnik ε insensitive loss function, SVR has been implemented in solving nonlinear regression estimation problems (19). A linear regression function can be expressed by (26) :

$$\boldsymbol{w} = \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) X_k \quad (1)$$

$$f(x) = w^T x + b \qquad (2)$$

With input values of $x_k \epsilon R^{Nxn}$ and output values of $y_k \epsilon R^N$ for N given training points. The empirical risk is given by :

$$R\text{-} = \frac{1}{N}\sum_{k=1}^{N} \left| y_k - w^T x - b \right|_\varepsilon \qquad (3)$$

with the Vapnik's ε insensitive loss function is defined as $\left| y - f(x) \right|_\varepsilon =$

$$\begin{cases} 0 & if \ \left| y - f(x) \right| \leq \varepsilon \\ y - f(x) \left| - \varepsilon \right. & \text{otherwise} \end{cases} \qquad (4)$$

the problem can be redefined in dual space by

$$max J_D(\alpha, \alpha^*) = \frac{1}{2}\sum_{k,i=1}^{N} (\alpha_k - \alpha_k^*)(\alpha_i - \alpha_i^*) X_k^* X_i \qquad (5)$$

$$-\varepsilon \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) + \sum_{k-1}^{N} y(\alpha_k - \alpha_k^*) \qquad (6)$$

Subject to $\begin{cases} \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) \\ \alpha_k, \alpha_k^* \epsilon [0, c] \end{cases}$

Likewise the expansion in feature space is

$$f(x) = \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) X_k^* X + b \qquad (7)$$

And W has defined in equation 1

The training points with values of α_k that are non-zero conform to the support vectors that are free, this allows the bias term b to be calculated. It is a safe practice to derive the bias term is by solving the regression function of the bias term and then balancing it out with the free support vectors (27). This solution justifies the SVM over the conventional statistical learning algorithms. The SVM is both unique and universal. The dual problem formulation is independent of the input space dimension and depends solely on the amount of training patterns (26). Nonlinear regression estimation can be derived by applying kernel functions. The common kernel functions applied are sigmoid functions, polynomial, and Gaussian radial basis (RBF) kernel. The sigmoid functions, polynomial, and Gaussian radial basis (RBF) kernels were implemented for modeling porosity and permeability in this study (23) . The nonlinear hyperplane regression

$$fx = \sum_{K=1}^{N} \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) K(X_k^* X) + b \qquad (8)$$

where α_k and α_k^* are the solutions to the quadratic problem above and b is the bias term which is derived as an average value over the free support vectors. This solution is also universal and unique as long as the kernel function is positive definite (23).

2.5.1 Kernel function and Kernel Examples

Kernel functions can be implemented to exploit nonlinear features of data over more expressive hypothesis spaces, the kernel functions project the data into high dimensional feature space to increase the ability of the linear learning machine to represent the nonlinear relationship that occurs in the original input space (23). Hence, the SVR solves the nonlinear regression challenge by mapping the input training

vectors x into ϕ(x) of a higher dimensional feature space where ϕ aids in transforming the input data space to the feature space. The kernel function k($\mathbf{x_i}$,$\mathbf{x}$) can be re pre-processed with the outcome being stored in a kernel matrix k= $(k(\mathbf{x_i},\mathbf{x}))^{l}_{ij=1}$. The kernel matrix must be positive definite in to assure a unique solution (23). The kernel functions implemented in this study as displayed in Table 2 all result in positive definite kernel matrices (28). Hence, the use of the kernel matrix result in nonlinear function approximations while preserving the simplicity and computational capability of linear SVR approximations (23).

Kernel Functions	**Mathematical Expressions**
Polynomial	$k(x_i, x) = (\alpha x_i^T x + c)^d$
Gaussian radial basis (RBF)	$K(x_i, x) = e^{-(\frac{\|x_i - x\|^2}{2\sigma^2})}$
Sigmoid	$K(x_i, x) = \tanh(k\langle xi, x\rangle + \boldsymbol{\vartheta})$

Table 2: showing kernel functions with their corresponding mathematical expressions.

2.6 Accuracy Assessment

The Root Mean Square Error (RMSE), Average Absolute error (AAE) were considered for assessing the model's accuracy i.e. how good the model performs in predicting porosity and permeability. The various accuracy measure mathematical expressions are shown in Table 3 below.

Accuracy Measure	**Mathematical Expressions**
RMSE	$\sqrt{\frac{1}{l}\sum_{i=1}^{l}(y_i - \hat{y}_i)^2}$
MAE	$max\left\|y_i - \hat{y}_i\right\|, i = 1, \ldots, l$
AAE	$\frac{1}{l}\sum_{i=1}^{l}\left\|y_i - \hat{y}_i\right\|$

Table 3: Accuracy measures with corresponding mathematical expressions.

.

3. Results and Analysis

3.1 Data Transformation

For the purpose of this research two data transformation techniques were considered namely: Logarithmic transformation and Standardization. These transformation technique were selected because of their simplicity and good performance on spatial data.

Logarithmic Transformation: The observed data were initially transformed to base 10 using the ***Log 10*** function in the R scripting language. After the logarithmic transformation has been performed it was observed that data points with values below 1 resulted in negative values, this was corrected by adding a value of 2 to all data points to avoid biasness and ensure all data points are positive. The transformed data was modeled using the different SVM kernel techniques and the error of the prediction was measured, a RMSE value range of 0.2 to 3 was derived and AAE value range of 0.1 to 2, this prediction accuracy derived shows improvement compared to the original data without transformation but also shows room for further improvement and further transformation was carried out.

Standardization: As a result of the prediction accuracy derived from the logarithmic transformation a different transformation technique was employed. The standardization technique was employed by declaring a normalize function:

$$z_o = x_o - (\min(x) * \max(x)) - \min(x) \qquad (9)$$

In the R scripting language, where z_o is the standardized value, x_o original value , $\min(x)$- minimum column value and $\max(x)$- maximum column value. This function which scales the data from 0 to 1 produced the smallest error of RMSE value range of 0.1 to 0.6, MAE value range of 0.5 to 2 and AAE value range of 0.05 to 0.4 compared to the other data types when calculated.

3.2 Kernel Selection and Optimization

Selection of the right kernels and tuning its parameter to obtain optimum results is often a problem in SVR techniques. Nevertheless, a set of rules can be employed in easing this optimization process and obtaining better results (25). Each kernel selection determines the parameters to tune for obtaining the desired result. For instance, the Gaussian RBF kernel requires little parameter tuning as it requires few parameters and performs well with data with no prior knowledge. For this research, three kernels were chosen, namely: polynomial, sigmoid and Gaussian radial and the results were all compared in terms of accuracy and performance. The cost of the constraint violation value (C) was set at a value of 1000 and 1 respectively while other parameters such as degree, epsilon and gamma were tuned to obtain optimum results (26). The high-value C (1000) resulted in a higher range of RMSE values for the three kernels and the low-value C (0.1) resulted in the lower range of RMSE values for the three kernels and was chosen for the model fitting of all kernels. The observed data of 206 data points was split into random training and testing data with a ratio of 80% and 20% respectively with the aid of caTools functions in code line 13 to 17 in figure 15 in the appendix. A seed of 101 was set to randomly select the exact testing and training data for consistency purpose, afterwards the e1071 package function ***svm*** was used to train the model for various kernels and setting permeability and porosity as response variable, furthermore the model derived was used to predict the test data with the function ***predict.*** The prediction correlation was observed by plotting the observed test data and predicted test data, Finally the accuracy of the prediction was determined using the RMSE, AAE and MAE functions.The R code used for fitting the model, predicting the data and examining model performance is shown in the Figures 15 respectively.

1) POLYNOMIAL KERNEL: After carefully tuning the parameters with a range of C values, the value of the parameters used for creating the polynomial kernel regression model in predicting both permeability and porosity in code line number 29 and 30 respectively are shown in Table 2 below:

Kernel	C	Degree	Coef	Epsilon
Polynomial	1	3	0	0.1
Radial	1	-	-	-0.1
Sigmoid	1	-	0.1	0.1

Table 4:showing the tuned parameters for the different kernel types.

Any further deviation from the value assigned to these parameters results in a negative effect on the RMSE value and model performance.

2) GAUSSIAN RADIAL KERNEL: The tuned parameters used for developing the Gaussian radial kernel regression model in predicting both permeability and porosity in code line number 31 and 32 respectively are displayed in Table 2 above. We recall that the Gaussian radial kernel requires little parameters for good performance, keeping the C value constant at 1000 and adapting the epsilon (ε) value results in an increase of the RMSE value as shown in Figure 8. This helps in adjusting the insensitive tube wideness until the optimal value is reached for ε.

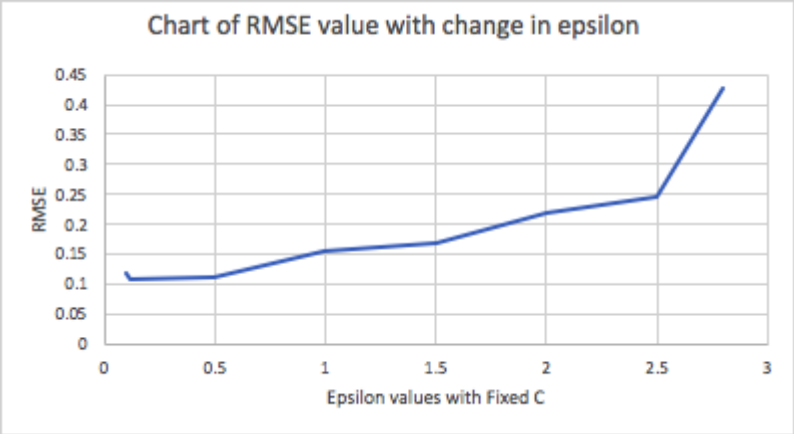

Figure 8: chart of RMSE values with change in epsilon value at constant C value of 1000.

SIGMOID KERNEL: The final kernel used for prediction of the petrophysical parameters is the sigmoid kernel regression model, the tuned parameter used in code line 33 and 34 respectively are displayed in Table 2 above. These parameters produce the optimum result using the sigmoid kernel.

3.3 Model Performance

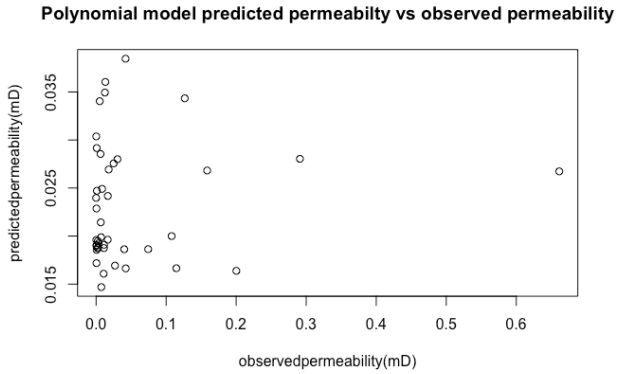

Figure 9: plot showing predicted permeability against permeability in test data using polynomial kernel of standardized dataset.

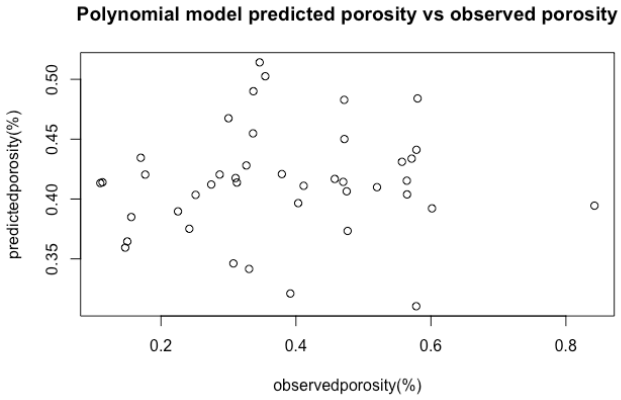

Figure 10: plot showing predicted porosity against porosity variable in test data using polynomial kernel of standardized dataset.

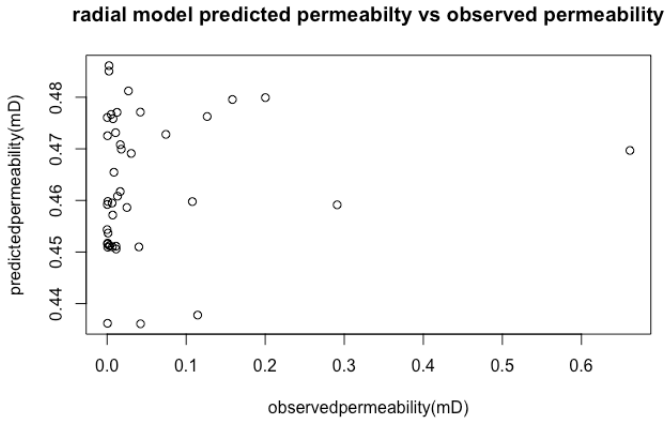

Figure 11: plot showing predicted permeability against permeability in test data using radial kernel of standardized dataset.

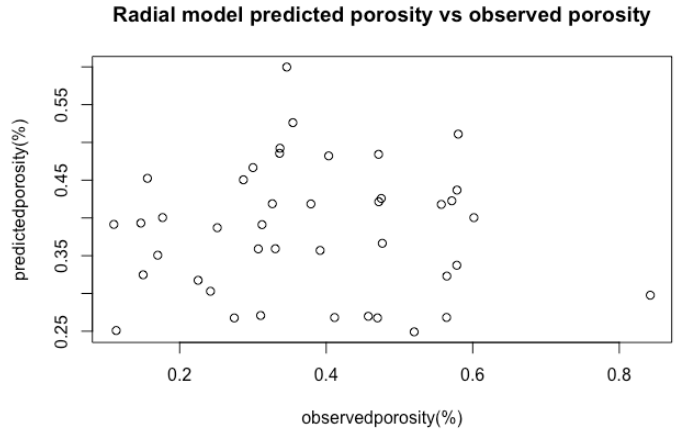

Figure 12: plot showing predicted porosity against porosity variable in test data using radial kernel of standardized dataset.

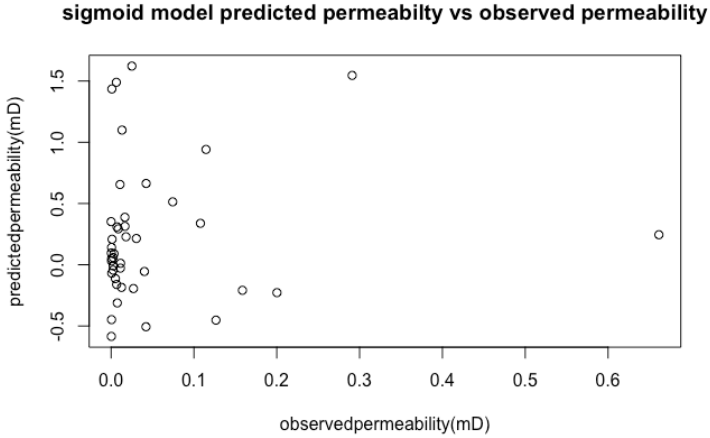

Figure 13: plot showing predicted permeability against permeability in test data using sigmoid kernel of standardized dataset**.**

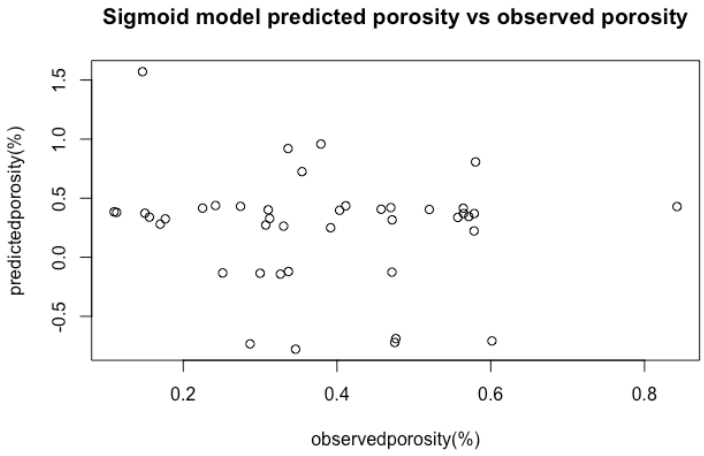

Figure 14: plot showing predicted porosity against porosity variable in test data using sigmoid kernel of the standardized dataset**.**

RMSE	AAE	R^2	Prediction	DataType
47	21	0.01	Permeability Prediction With Polynomial Kernel	Data Without Transformation
50	22	0.1	Permeability Prediction With Radial Kernel	Data Without Transformation
248	174	0.01	Permeability Prediction With Sigmoid Kernel	Data Without Transformation
9	7	0.06	Porosity Prediction With Polynomial Kernel	Data Without Transformation
9	7	0.05	Porosity Prediction With Radial Kernel	Data without Transformation
20	16	1.3E-05	Porosity Prediction With Sigmoid Kernel	Data Without Transformation
1	0.8	0.01	Permeability Prediction With Polynomial Kernel	logarithmic Transformation
0.8	0.6	0.3	Permeability Prediction With Radial Kernel	logarithmic Transformation
3	2	0.02	Permeability Prediction With Sigmoid Kernel	logarithmic Transformation
0.2	0.1	0.004	Porosity Prediction With Polynomial Kernel	logarithmic Transformation
0.2	0.2	0.02	Porosity Prediction With Radial Kernel	logarithmic Transformation
0.7	0.5	0.04	Porosity Prediction With Sigmoid Kernel	logarithmic Transformation
0.1	0.05	0.01	Permeability Prediction With Polynomial Kernel	standardized
0.1	0.4	0.2	Permeability Prediction With Radial Kernel	standardized
0.6	0.4	0.01	Permeability Prediction With Sigmoid Kernel	standardized

0.3	0.1	0.01	Porosity Prediction With Polynomial Kernel	standardized
0.2	0.2	0.001	Porosity Prediction With Radial Kernel	standardized
0.5	0.4	0.07	Porosity Prediction With Sigmoid Kernel	standardized

Table 5: The different kernel and data types with the respective RMSE, AAAE and coefficient of determination values.

Table 2 summarizes the measure of precision of the different models employed for the purpose of this research. Initial analysis performed on the original data resulted in a very low coefficient of determination results **(**R^2) for all of the kernels, AAE values of range 7 to 174 and very high RMSE values of range 9 to 248 which shows a high deviation of the value predicted from the original permeability and porosity data, from the three kernels used the sigmoid kernel produced the lowest coefficient of determination value and the highest RMSE value for both permeability and porosity data prediction of [0.01, 248], [1.3E-05,20] respectively. The radial kernel model produced the highest R^2 value of 0.1for the permeability prediction with the polynomial kernel model producing the lowest RMSE value of 47 and lowest AAE value of 21 for the permeability prediction and for the porosity prediction the polynomial kernel model developed the lowest RMSE value of 9, the lowest AAE value of 7 and highest coefficient of determination of 0.06. The low coefficient of determination and very high RMSE value produced from the original data resulted in further analysis to understand the data better.

The original data was transformed by a logarithmic function and alternatively declaring a standardized function. The standardized function which has been proven to give better SVR results according to *Crone* et al (27) was used to normalize the variables from 0 to 1 where the minimum value is 0 and the maximum value is 1. The analysis performed on the standardized set of data with the line of codes displayed in Figure 15 and16 produced the best results out of the three data types used. The radial kernel produced the lowest RMSE value of 0.1 for the predicted permeability data which proves the radial kernel model accurately predicted the response value (permeability data), The radial kernel also produced the highest coefficient of determination of 0.15, however this low value of the coefficient of determination produced explains the inability of the radial model to explain the variability of permeability generally with input data. The radial kernel model likewise produced the lowest RMSE value of 0.2 for the predicted porosity data from the standardized data type showing that the radial kernel model predicts the porosity data for the study area accurately. The highest value of 0.07 for coefficient of determination obtained was from the sigmoid model kernel which shows that the variability of the porosity data not generally explained by the model. Finally, the radial kernel model using the Logarithmic transformation data produced the highest explanation of the general variability of permeability with an R^2 value of 0.32 and the highest value of 4 for the R^2 of porosity data was produced from the sigmoid kernel model using the Logarithmic transformation data proving that the general variability of the porosity data was not well explained by the model.

3.4 Data Back Transformation

Results are usually more appropriate to be reported in the original scale of measurement after the analysis is performed on transformed data. Back transformation after the analysis of the transformed data is then required **(28)**. For this reason, the two transformations performed on the original data were back-transformed as follows:

LOGARITHMITIC BACK TRANSFORMATION : A value of 2 was generally deducted from the logarithmically transformed data as a value of 2 was initially added to the logarithmically transformed data to eradicate all negative transformed values, afterwards, the exponent of these results was performed using the ***exp()*** function in the R programming language, the original scale of predicted measurements is obtained and displayed in Table 3 and 4 in appendix

STANDARDIZATION BACK TRANSFORMATION: To transform the standardized data back to the original scale of measurement, a set of functions was declared for each of the individual columns, with the general formula:

$$x_o = [x * \mathbf{max}(x) - x * \mathbf{min}(x) + \mathbf{mi\,n}(x) \quad (\mathbf{10})$$

where max(x) is the maximum value of original data columns, min(x) is the minimum value of original data columns and standardized data is the normalized values in each column. the original scale of predicted measurement is obtained and displayed in Table 5 and 6 in appendix

3.5 Comparison with Artificial Neural Network

Although the SVR models predicted a reasonable set of permeability and porosity data, the precision of the models was compared to the ANN technique. The ANN was chosen for its ability to analyze the non-linear problem by analyzing multi-source inputs by self-learning and producing solutions without hypothesis.

For developing the artificial neural network model, the input data was split into 80% of the 206 data points for training data in creating the model and 20% of the 206 data points for testing data in checking the performance of the model, the input layer has three nodes attributing to the spatial data: longitude, latitude and elevation were declared and an output layer of one node attributing to permeability and porosity respectively was declared for the prediction of the individual petrophysical data.2 hidden layers were selected for the ANN model with four nodes for each layer using the rectified linear unit (RELU) activation function.

The gradient descent algorithm was chosen to fit the model. The selection of the optimum training data set was performed by setting an epoch value and stopping the cycling at the optimum cycle by weighting the MSE value obtained from each cycle. However, the model predicted rather high RMSE values of 89 for the model prediction of permeability and RMSE value of 10 for the model prediction of porosity even after scaling the data to avoid bias from columns with high range values.

4. Conclusion

Three different Support Vector Regression kernels were developed and trained to predict two petrophysical properties:-permeability and porosity. The original data was transformed to yield better results and back-transformed to obtain the original scale of measurement. The performance of the three different kernels used to train the data was assessed. The result obtained points out that the radial kernel of the standardized dataset outperforms the sigmoid kernel and polynomial kernel of all data types in the prediction of the permeability data in the area of study and the radial kernel of the standardized dataset produced the least RMSE value for porosity prediction. The radial kernel permeability prediction accuracy developed from the standardized dataset is an RMSE value of 0.1 and MAE value of 0.5 and the Radial kernel porosity prediction accuracy obtained from the standardized dataset is an RMSE value of 0.2 MAE value of 0.5. However, the low value of the coefficient of determination obtained could be improved by acquiring more data points and integrating additional petrophysical properties such as the gamma-ray log data. Comparison of the SVR technique with the ANN model shows that the SVR technique acquires a better predictive capacity and requires lower data points to produce more reasonable results. Nevertheless, the ANN model could also be improved by acquiring much more data points as it requires a high volume of data for better performance.

5. References

1. ***petrophysics.*** petrowiki. **s.l. : petrowiki, 2015.**

2. ***Variation of Porosity and Permeability of Nahr Umr and Yammama Formations in South of Iraq.*** Al-Sudani, Hussein. **s.l. : Researchgate, 2019.**

3. ***Paper presented in SPE Annual Technical Conference and Exhibition.*** F.G. Paul, C. Luca. **Denver,Colorado : SPE-84387, 2003.**

4. ***Fundamental Objections to the 7th Approximation.*** WEBSTER,R. **2, s.l. : European Journal of Soil Science, Vol. 19 1968.**

5. ***Porosity prediction using cokriging with multiple secondary datasets.*** **s.l.**Hong Xu, Jian Sun, Brian H. Russell, and Kris A. Innanen. **: CREWES Research Report, 2015.**

6. ***Spatial prediction of soil properties in temperate mountain regions using support vector regression .*** Ballabio, Cristiano. **s.l. : The Global Journal of Soil Science (elsevier), Vol. 151, 2009.**

7. ***Support vector regression machines. .*** Drucker, Harris & Burges, Christopher & Kaufman, Linda & Smola, Alexander & Vapnik, V. **Adv Neural Inform Process Syst, s.l. : ResearchGate, 1997, pp. 779-784.**

8. ***Predict soil texture distributions using an artificial neural network model.*** Zhengyong Zhaoa, Thien Lien Chowb, Herb W. Reesb, Qi Yanga, Zisheng Xingb, Fan-Rui Menga. **65 , s.l. : Elsevier,Vol. computers and electronics in agriculture, 2009.**

9. ***Erdgeschichte – Spurensuche im Gestein. Wissenschaftliche Buchgesellschaft.*** **darmstadt.** Rothe, P. **: s.n. p. 240., 2000.**

10. ***Climate Cycles Documented in a Playa System: Comparison of Geochemical Signatures Derived from Subbasins (Triassic, Middle Keuper, German Basin).*** Lutz Reinhardt, Werner Ricken. **Zentralblatt für Geologie und Paläontologie Teil1, s.l. : Researchgate, Vol. 1999. 315. 2000.**

11. ***Geological Atlas of Western and Central Europe.*** **s.l. : Shell International Petroleum Matschappij, The Hague.**Ziegler, Peter A. **1990.**

12. ***Geology of Baden-Wuerttemberg.*** **stuggart : schweizerbartsche Verlagsbuchhandlung**.Geyer, Otto F. and Gwinner:, Manfred P. **1991.**

13. ***Stratigraphy and nomenclature of the southwest German sandstone kettle.*** Brenner, Klaus. ***Geological State Office of Baden-Württemberg*** **23, 1981.**

14. ***Reservoirarchitecture in a terminal alluvial plain: An outcrop analogue study (Upper triasic, southern Germany).*** HORNUNG, J. und AIGNER, T. **Sedimentology and petrophysics, s.l. : Journal of Petroleum Geology, Vol. 1, pp. 3-30. 2002.**

15. ***Klimagesteuerte Denudations-Sedimentakkumulations-Systeme: Massenbilanzen, Modellierung und Anwendung.*** **s.l.** HINDERER, M. **: Habilitationsschrift Univ. Tübingen, 1999.**

16. ***3-D-Reservoir-Charakterisierung einer Gleithangschichtung im Stubensandstein (Mittlerer Keuper).***Greb, Matthias D**. esslingen : TU Darmstadt, 2006.**

17. ***Reservoirarchitecture in a terminal alluvial plain: An outcrop analogue study (Upper triasic, southern Germany) .*** HORNUNG, J. und AIGNER, T. **Cyclicity, controls and models, s.l. : Journal of Petroleum Geology, 2002, Vol. 1, pp. 151-178. .**

18. ***Statistical Learning Theory.***V, Vapnik. **New York : Wiley, 1998.**

19. ***Support vector networks.*** V, Cortes C. and Vapnik. **s.l. Vol. Machine Learning : springer, 1995.**

20. ***Predicting Time Series with a Local Support Vector Regression Machine.***Fernandez, Rodrigo. **Paris : LIPN, Institut Galilee-UniversiteParis , 1999.**

21. ***Enhancing the performance of elastic graph matching for face authentications by using Support Vector Machines.*** Tefas A., Kotropoulos C., and Pitas I. **s.l. : Researchgate, 2000.**

22.*The Application of Support Vector Machines to Medical Decision Support: A Case Study* Of Advance Course Artificial Intelligence, Veropoulos K., Cristianini N., and Campbell C. **: ResearchGate, 1999.**
23. *Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study.* A.F. Al-Anazi, I.D. Gates. **computers & geosciences Vol. 36, elsevier ,calgary :, 2010.**
24. *Learning from Data - Concepts, Theory and Methods.*Muller, V. Cherkassky and F.**: John Wiley & Sons,Inc, New Jersey, 2007.**
25. *Support vector regression.* Basak, D., Pal, S., Patranabis, D.C. **Letters and Reviews,Vol. Neural Information Processing. 2007.**
26. *Least Square Support Vector Machine..* Suykens Johan, Van Gestel Tony,De Brabanter Jos ,De Moor Bart,Vandewalle Joos.**: Singapore s.n., 2002.**
27. *Support Vector Machines – An Introduction..* V, Kecman, **: Springer, Berlin , 2005.**
28. *An Introduction to Support Vector Machines .* Cristianini N., Shawe-Taylor John. **s.l. : Cambridge University Press, 200.**
29. *A Practical Guide to Support Vector Classification Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin.* Hsu, Chih-wei .Chang, Chih-chung . Lin, Chih-JenPY. **2003.**
30. *Support Vector Machines(Interface to Libsvm r package e1071).* **s.l.**Meyer, David. **: FH Technikum Wien Austria, 2019.**
31. *The Impact of Preprocessing on Support Vector Regression and Neural Networks in Time Series Prediction.* **s.l.**Sven F. Crone, Jose Guajardo ,Richard Webe. **: DMIN, 2006.**
32. *Handbook of Biological Statistics..* McDonald, John H. **: Sparky House Publishing, Baltimore, Maryland, 1988.**

6. Appendix

```
1.library(e1071) #loading the e1071 package for svm regression
2.library(psych) #loading the psych package for summary statistics and visualization
3.describe(Pphydat) #summary statistics of the data loaded
4.Pphydataper <- Pphydat[,c("long","lat","permeability","elevation")] #calling out the
required  for permeability modelling and prediction
5.Pphydatapor <- Pphydat[,c("long","lat","porosity","elevation")]#declaring the required
varaiables for porosity modelling and prediction
6.normalize <- function(x) {return ((x - min(x)) / (max(x) - min(x)))} # defining the
normalize function for standardization transformation of the varaiables
7.Pphydataper$longnorm<-normalize(Pphydataper$long) #standardizing longitude variable in
pphydataper dataset
8.Pphydataper$latnorm<-normalize(Pphydataper$lat) #standardizing latitude variable in
pphydataper dataset
9.Pphydataper$elevationnorm<-normalize(Pphydataper$elevation) #standardizing elevation
variable in pphydataper dataset
10.Pphydataper$permeabilitynorm<-normalize(Pphydataper$permeability)#standardizing
permeability variable in pphydataper dataset
11.Pphydataper =Pphydataper[,c("longnorm","latnorm","elevationnorm","permeabilitynorm")] #
attributing the standardize variables to Pphydataper
12.describe(Pphydataper) #summary statistics of Pphydataper scaling variables onto [0,1]
13.require(caTools)
14. set.seed(101) # setting the seed for random selection of recurring data points
15.sample = sample.split(Pphydataper$permeabilitynorm,SplitRatio= .8) # splitting the datset
to 80% and 20%
16.trainper = subset(Pphydataper,sample == TRUE) #declaring the 80% of the data for
trainining data (permeability modelling)
17.testper  = subset(Pphydataper,sample == FALSE) #declaring 20% of the data to testing data
(permeability modelling)
18.Pphydatapor$longnormm <-normalize(Pphydatapor$long) #standardizing longitude variable in
pphydatapor dataset
19.Pphydatapor$latnormm <-normalize(Pphydatapor$lat) #standardizing latitude variable in
pphydatapor dataset
20.Pphydatapor$elevationnormm <-normalize(Pphydatapor$elevation) #standardizing elevation
variable in pphydatapor dataset
21.Pphydatapor$porositynormm <-normalize(Pphydatapor$porosity)#standardizing permeability
variable in pphydatapor dataset
22.Pphydatapor <- Pphydatapor[,c("longnormm","latnormm","elevationnormm","porositynormm")] #
declaring the standardize variables to Pphydatapor
23.describe(Pphydatapor)#summary statistics of Pphydatapor scaling variables onto [0,1]
24.require(caTools)
25.set.seed(101) # setting the seed for random selection of recurring data points
26.samplee = sample.split(Pphydatapor$porositynormm,SplitRatio= .8) # splitting the datset
to 80% and 20%
27.trainpor = subset(Pphydatapor,samplee == TRUE) # splitting 80% of the data for trainining
data (porosity modelling)
28.testpor  = subset(Pphydatapor,samplee == FALSE) # splitting 20% of the data to testing
data (porosity modelling)
29.fitpolyperm<-svm(permeabilitynorm~.,data=trainper,type
="epsregression",kernel="polynomial") #training the svm polynomial regression model,using
standardized permeability as response value
30.fitpolyporo<-svm(porositynormm~.,data=trainpor,type ="eps-
regression",kernel="polynomial") #training the svm polynomial regresiion model,using
standardized porosity as response value
31.fitradialperm<-svm(permeabilitynorm~.,data=trainper,type ="eps-
regression",kernel="radial") #training the svm radial regresiion model,using standardized
permeability as response value
32.fitradialporo=svm(porositynormm~.,data=trainpor,type="epsregression",kernel="radial")#tra
ining the svm radialmodel,using standardized porosity as response value
```

Figure 15: line of code written in the R scripting language for the SVR technique.

```
33.fitsigmoidperm=svm(permeabilitynorm~.,data=trainper,type="epsregression",kernel
="sigmoid")#training the svm sigmoid regresiion model,using standardized
permeability as response value
34.fitsigmoidporo<- svm(porositynormm~.,data=trainpor,type ="eps-
regression",kernel="sigmoid") #training the svm sigmoid regresiion model,using
standardized porosity as response value
35.predictperm<- predict(fitpolyperm,testper) #predicting permeability using
trained svm polynomial model
36.predictpermrad <-predict(fitradialperm,testper)#predicting permeability using
trained svm radial model
predictpermsigmoid <-predict(fitsigmoidperm,testper)#predicting permeability using
trained svm sigmoid model
37.predictporo<- predict(fitpolyporo,testpor) #predicting porosity using trained
svm polynomial model
38.predictpororad <-predict(fitradialporo,testpor)#predicting porosity using
trained svm radial model
39.predictporosigmoid <-predict(fitsigmoidporo,testpor)#predicting porosity using
trained svm sigmoid model
40.plot(testper$permeabilitynorm,predictperm)#plotting the observed test data
against predicted permeability data from polynomial regression model
41.plot(testper$permeabilitynorm,predictpermrad)#plotting the observed test data
against predicted permeability data from radial regression model
42.plot(testper$permeabilitynorm,predictpermsigmoid)#plotting the observed test
data against predicted permeability data from sigmoid regression model
43.plot(testpor$porositynormm,predictporo)#plotting the observed test data against
predicted permeability data from polynomial regression model
44.plot(testpor$porositynormm,predictpororad)#plotting the observed test data
against predicted permeability data from radial regression model
45.plot(testpor$porositynormm,predictporosigmoid)#plotting the observed test data
against predicted sigmoid data from sigmoid regression model
46.cor(testper$permeabilitynorm,predictpermrad)#plotting the observed test data
against predicted permeability data from radial regression model
47.cor(testper$permeabilitynorm,predictperm)#plotting the observed test data
against predicted permeability data from polynomial regression model
48.cor(testper$permeabilitynorm,predictpermsigmoid)#plotting the observed test
data against predicted permeability data from sigmoid regression model
49.cor(testpor$porositynormm,predictporo)#plotting the observed test data against
predicted porosity data from polynomial regression model
50.cor(testpor$porositynormm,predictpororad)#plotting the observed test data
against predicted porosity data from radial regression model
51.cor(testpor$porositynormm,predictporosigmoid)#plotting the observed test data
against predicted porosity data from sigmoid regression model
52.library(Metrics)
53.RMSE(testper$permeabilitynorm, predictpermrad)#obtaining the
rootmeansquarederror of the radialkernel model
54.RMSE(testper$permeabilitynorm, predictpermsigmoid)#obtaining the
rootmeansquarederror of the sigmoidkernel model
55.RMSE(testper$permeabilitynorm, predictperm)#obtaining the rootmeansquarederror
of the polynomialkernel model
56.RMSE(testpor$porositynormm,predictporo)#obtaining the rootmeansquarederror of
the polynomialkernel model
57.RMSE(testpor$porositynormm,predictpororad)#obtaining the rootmeansquarederror
of the radiallkerenel model
```

Figure 16 : line of code written in the R scripting language for the SVR technique.

longitude	latitude	elevation	Initial permeability values of testing data	Predicted Permeability using polynomial Kernel	Predicted Permeability using radial Kernel	Predicted Permeability using sigmoid Kernel
9.32159	48.7734	428.622	1.498	1.712913357	1.209042946	1.85327004
9.32159	48.7734	421.764	3.06	2.068490701	2.74125652	81.4984414
9.32159	48.7734	418.35	1.133	2.235781688	4.376668624	136.040725
9.32123	48.7737	434.924	42.459	0.953173085	0.811554364	0.00261719
9.32123	48.7737	430.543	1.188	1.408055475	1.375907753	0.17723188
9.32123	48.7737	422.924	41.976	2.041692933	4.254955582	37.3789904
9.32123	48.7737	421.799	32.595	2.105212129	5.072968967	60.0011172
9.32147	48.7734	434.815	0.237	1.100855488	0.765536369	0.00316557
9.32147	48.7734	430.893	0.36	1.423599535	1.110999249	0.14397895
9.32147	48.7734	430.486	0.444	1.455042547	1.160162582	0.20778225
9.32147	48.7734	430.226	0.662	1.474868192	1.193225107	0.26173114
9.32147	48.7734	429.67	1.105	1.516555601	1.268558198	0.42466695
9.32147	48.7734	428.24	0.228	1.619122695	1.495126212	1.38099536
9.32277	48.7748	435.451	22.434	0.501342457	0.21217022	25672.2715
9.32277	48.7748	420.427	8.18	1.750603838	1.753183379	0.00157129
9.32277	48.7748	419.817	0.134	1.785827721	2.036767948	0.00104944
9.32203	48.7734	409.813	37.714	3.171613253	6.153613603	0.46584974
9.32459	48.7757	401.06	0.08	10.14061023	5.001331902	20165.5459
9.32459	48.7757	395.467	10.828	34.16114977	1.439785157	26370960.3
9.32277	48.7746	409.508	9.337	2.261485102	34.9911844	9.57E-05
9.32277	48.7746	409.076	244.068	2.277079871	37.2102185	0.00010104
9.32277	48.7746	407.201	260.767	2.361261442	45.57356161	0.0001513
9.32277	48.7746	392.451	1.855	6.753779542	6.682990086	3.36126419
9.32277	48.7746	389.574	0.234	10.74673844	3.229312337	51.032698
9.32173	48.7732	400.935	36.244	10.01818583	12.36627612	8.70E-05
9.32341	48.7753	415.234	30.511	1.6065035	11.00221067	0.0010741
9.32341	48.7753	403.826	11.915	3.221192855	15.18535405	3.36469784
9.32419	48.7752	409.605	13.17	2.412387639	83.340524	0.33552924
9.32419	48.7752	409.101	285.542	2.524150047	81.91053177	0.50980185
9.32419	48.7752	407	6.956	3.086345908	70.64627947	3.51721402
9.32419	48.7752	405.002	32.118	3.819452024	54.96539292	28.5974421
9.32419	48.7752	403.98	35.17	4.301077883	46.50699833	91.1868149
9.32419	48.7752	402.999	4.661	4.854104287	38.73567526	291.921231
9.32308	48.7755	410.416	1.034	1.370706306	2.60616551	0.00389057
9.32308	48.7755	406.274	1.066	1.596763805	2.758396499	0.09086811
9.32308	48.7755	405.777	1.033	1.623688573	2.718767523	0.1445521
9.32308	48.7755	404.705	1.043	1.682750123	2.595272217	0.41601061
9.32308	48.7755	404.057	1.069	1.719476723	2.498681817	0.81619681
9.32308	48.7755	396.927	1.077	2.27950857	1.133627594	5130.25015
9.32167	48.7732	415.459	3.721	2.769059429	4.897726796	81.0759753
9.32167	48.7732	413.642	21.479	2.991391173	6.091224389	42.8879423
9.32167	48.7732	410.25	17.731	3.626274764	8.794364796	5.44023801

Table 6 : showing the initial values of testing data and the back transformed logarithmic predicted Permeability data

longitude	latitude	elevation	Initial porosity values of testing data	Predicted Porosity using polynomial Kernel	Predicted Porosity using radial Kernel	Predicted Porosity using sigmoid Kernel
9.32159	48.7734	428.622	22.6	1.712913357	1.209042946	1.85327004
9.32159	48.7734	421.764	36.231	2.068490701	2.74125652	81.4984414
9.32159	48.7734	418.35	8.401	2.235781688	4.376668624	136.040725
9.32123	48.7737	434.924	19.88	0.953173085	0.811554364	0.00261719
9.32123	48.7737	430.543	10.568	1.408055475	1.375907753	0.17723188
9.32123	48.7737	422.924	18.96	2.041692933	4.254955582	37.3789904
9.32123	48.7737	421.799	24.994	2.105212129	5.072968967	60.0011172
9.32147	48.7734	434.815	21.905	1.100855488	0.765536369	0.00316557
9.32147	48.7734	430.893	13.735	1.423599535	1.110999249	0.14397895
9.32147	48.7734	430.486	24.847	1.455042547	1.160162582	0.20778225
9.32147	48.7734	430.226	22.852	1.474868192	1.193225107	0.26173114
9.32147	48.7734	429.67	20.48	1.516555601	1.268558198	0.42466695
9.32147	48.7734	428.24	21.348	1.619122695	1.495126212	1.38099536
9.32277	48.7748	435.451	21.978	0.501342457	0.21217022	25672.2715
9.32277	48.7748	420.427	21.372	1.750603838	1.753183379	0.00157129
9.32277	48.7748	419.817	16.636	1.785827721	2.036767948	0.00104944
9.32203	48.7734	409.813	16.334	3.171613253	6.153613603	0.46584974
9.32459	48.7757	401.06	19.993	10.14061023	5.001331902	20165.5459
9.32459	48.7757	395.467	20.758	34.16114977	1.439785157	26370960.3
9.32277	48.7746	409.508	16.161	2.261485102	34.9911844	9.57E-05
9.32277	48.7746	409.076	13.86	2.277079871	37.2102185	0.00010104
9.32277	48.7746	407.201	17.022	2.361261442	45.57356161	0.0001513
9.32277	48.7746	392.451	26.173	6.753779542	6.682990086	3.36126419
9.32277	48.7746	389.574	6.818	10.74673844	3.229312337	51.032698
9.32173	48.7732	400.935	23.218	10.01818583	12.36627612	8.70E-05
9.32341	48.7753	415.234	24.277	1.6065035	11.00221067	0.0010741
9.32341	48.7753	403.826	13.739	3.221192855	15.18535405	3.36469784
9.32419	48.7752	409.605	14.848	2.412387639	83.340524	0.33552924
9.32419	48.7752	409.101	20.542	2.524150047	81.91053177	0.50980185
9.32419	48.7752	407	14.907	3.086345908	70.64627947	3.51721402
9.32419	48.7752	405.002	14.409	3.819452024	54.96539292	28.5974421
9.32419	48.7752	403.98	41.886	4.301077883	46.50699833	91.1868149
9.32419	48.7752	402.999	11.23	4.854104287	38.73567526	291.921231
9.32308	48.7755	410.416	20.045	1.370706306	2.60616551	0.00389057
9.32308	48.7755	406.274	25.478	1.596763805	2.758396499	0.09086811
9.32308	48.7755	405.777	8.072	1.623688573	2.718767523	0.1445521
9.32308	48.7755	404.705	10.646	1.682750123	2.595272217	0.41601061
9.32308	48.7755	404.057	13.815	1.719476723	2.498681817	0.81619681
9.32308	48.7755	396.927	11.562	2.27950857	1.133627594	5130.25015
9.32167	48.7732	415.459	10.836	2.769059429	4.897726796	81.0759753
9.32167	48.7732	413.642	23.799	2.991391173	6.091224389	42.8879423
9.32167	48.7732	410.25	11.854	3.626274764	8.794364796	5.44023801

Table 7 : showing the initial values of testing data and the back transformed logarithmic predicted Permeability data

longitude	latitude	elevation	Initial permeability values of testing data	Predicted Permeability using polynomial Kernel	Predicted Permeability using radial Kernel	Predicted Permeability using sigmoid Kernel
9.32153	48.7734	430.407	4.923	1.712913357	1.209042946	1.85327004
9.32153	48.7735	428.622	1.498	2.068490701	2.74125652	81.4984414
9.32153	48.7735	421.764	3.06	2.235781688	4.376668624	136.040725
9.32153	48.7735	419.844	46.568	0.953173085	0.811554364	0.00261719
9.32191	48.7732	433.353	2.888	1.408055475	1.375907753	0.17723188
9.32191	48.7732	425.394	0.506	2.041692933	4.254955582	37.3789904
9.32191	48.7732	417.937	7.174	2.105212129	5.072968967	60.0011172
9.32158	48.7734	430.893	0.36	1.100855488	0.765536369	0.00316557
9.32158	48.7734	430.612	17.373	1.423599535	1.110999249	0.14397895
9.32158	48.7734	430.051	4.856	1.455042547	1.160162582	0.20778225
9.32158	48.7734	429.67	1.105	1.474868192	1.193225107	0.26173114
9.32158	48.7734	428.823	0.226	1.516555601	1.268558198	0.42466695
9.32158	48.7734	428.406	0.273	1.619122695	1.495126212	1.38099536
9.32347	48.7744	437.1	49.475	0.501342457	0.21217022	25672.2715
9.32347	48.7744	434.672	18.272	1.750603838	1.753183379	0.00157129
9.32347	48.7744	429.784	0.241	1.785827721	2.036767948	0.00104944
9.32347	48.7744	419.817	0.134	3.171613253	6.153613603	0.46584974
9.32146	48.7738	423.732	0.011	10.14061023	5.001331902	20165.5459
9.32146	48.7738	412.311	0.101	34.16114977	1.439785157	26370960.3
9.32459	48.7757	402.854	5.703	2.261485102	34.9911844	9.57E-05
9.32459	48.7757	398.046	0.395	2.277079871	37.2102185	0.00010104
9.32459	48.7757	397.29	2.739	2.361261442	45.57356161	0.0001513
9.32459	48.7757	396.512	125.66	6.753779542	6.682990086	3.36126419
9.32459	48.7757	395.467	10.828	10.74673844	3.229312337	51.032698
9.32312	48.7744	389.574	0.234	10.01818583	12.36627612	8.70E-05
9.32128	48.7736	406.05	2.235	1.6065035	11.00221067	0.0010741
9.32128	48.7736	405.12	5.431	3.221192855	15.18535405	3.36469784
9.32128	48.7736	401.569	18.19	2.412387639	83.340524	0.33552924
9.32409	48.7748	410.183	68.468	2.524150047	81.91053177	0.50980185
9.32397	48.7754	409.605	13.17	3.086345908	70.64627947	3.51721402
9.32397	48.7754	409.101	285.542	3.819452024	54.96539292	28.5974421
9.32397	48.7754	408.001	7.234	4.301077883	46.50699833	91.1868149
9.32397	48.7754	405.002	32.118	4.854104287	38.73567526	291.921231
9.32397	48.7754	402.999	4.661	1.370706306	2.60616551	0.00389057
9.32436	48.7746	411.11	3.23	1.596763805	2.758396499	0.09086811
9.32436	48.7746	409.239	86.463	1.623688573	2.718767523	0.1445521
9.32436	48.7746	408.556	11.654	1.682750123	2.595272217	0.41601061
9.32436	48.7746	405.777	1.033	1.719476723	2.498681817	0.81619681
9.32436	48.7746	404.057	1.069	2.27950857	1.133627594	5130.25015
9.32128	48.7736	415.459	3.721	2.769059429	4.897726796	81.0759753
9.32128	48.7736	412.209	7.769	2.991391173	6.091224389	42.8879423
9.32128	48.7736	401.734	54.66	3.626274764	8.794364796	5.44023801

Table 8 : showing the initial values of testing data and the back transformed standardized predicted Permeability data

longitude	latitude	elevation	Initial porosity values of testingdata	Predicted Porosity using polynomial Kernel	Predicted Porosity using radial Kernel	Predicted Porosity using sigmoid Kernel
9.32153	48.7734	430.407	5.394	1.712913357	1.209042946	1.85327004
9.32153	48.7735	428.622	22.6	2.068490701	2.74125652	81.4984414
9.32153	48.7735	421.764	36.231	2.235781688	4.376668624	136.040725
9.32153	48.7735	419.844	10.123	0.953173085	0.811554364	0.00261719
9.32191	48.7732	433.353	7.792	1.408055475	1.375907753	0.17723188
9.32191	48.7732	425.394	5.257	2.041692933	4.254955582	37.3789904
9.32191	48.7732	417.937	7.198	2.105212129	5.072968967	60.0011172
9.32158	48.7734	430.893	13.735	1.100855488	0.765536369	0.00316557
9.32158	48.7734	430.612	19.95	1.423599535	1.110999249	0.14397895
9.32158	48.7734	430.051	24.469	1.455042547	1.160162582	0.20778225
9.32158	48.7734	429.67	20.48	1.474868192	1.193225107	0.26173114
9.32158	48.7734	428.823	12.211	1.516555601	1.268558198	0.42466695
9.32158	48.7734	428.406	18.002	1.619122695	1.495126212	1.38099536
9.32347	48.7744	437.1	15.594	0.501342457	0.21217022	25672.2715
9.32347	48.7744	434.672	25.142	1.750603838	1.753183379	0.00157129
9.32347	48.7744	429.784	14.827	1.785827721	2.036767948	0.00104944
9.32347	48.7744	419.817	16.636	3.171613253	6.153613603	0.46584974
9.32146	48.7738	423.732	24.484	10.14061023	5.001331902	20165.5459
9.32146	48.7738	412.311	17.662	34.16114977	1.439785157	26370960.3
9.32459	48.7757	402.854	15.242	2.261485102	34.9911844	9.57E-05
9.32459	48.7757	398.046	12.738	2.277079871	37.2102185	0.00010104
9.32459	48.7757	397.29	20.699	2.361261442	45.57356161	0.0001513
9.32459	48.7757	396.512	26.046	6.753779542	6.682990086	3.36126419
9.32459	48.7757	395.467	20.758	10.74673844	3.229312337	51.032698
9.32312	48.7744	389.574	6.818	10.01818583	12.36627612	8.70E-05
9.32128	48.7736	406.05	13.598	1.6065035	11.00221067	0.0010741
9.32128	48.7736	405.12	14.581	3.221192855	15.18535405	3.36469784
9.32128	48.7736	401.569	17.168	2.412387639	83.340524	0.33552924
9.32409	48.7748	410.183	20.553	2.524150047	81.91053177	0.50980185
9.32397	48.7754	409.605	14.848	3.086345908	70.64627947	3.51721402
9.32397	48.7754	409.101	20.542	3.819452024	54.96539292	28.5974421
9.32397	48.7754	408.001	13.291	4.301077883	46.50699833	91.1868149
9.32397	48.7754	405.002	14.409	4.854104287	38.73567526	291.921231
9.32397	48.7754	402.999	11.23	1.370706306	2.60616551	0.00389057
9.32436	48.7746	411.11	25.071	1.596763805	2.758396499	0.09086811
9.32436	48.7746	409.239	24.772	1.623688573	2.718767523	0.1445521
9.32436	48.7746	408.556	24.166	1.682750123	2.595272217	0.41601061
9.32436	48.7746	405.777	8.072	1.719476723	2.498681817	0.81619681
9.32436	48.7746	404.057	13.815	2.27950857	1.133627594	5130.25015
9.32128	48.7736	415.459	10.836	2.769059429	4.897726796	81.0759753
9.32128	48.7736	412.209	6.95	2.991391173	6.091224389	42.8879423
9.32128	48.7736	401.734	25.059	3.626274764	8.794364796	5.44023801

Table 9: showing the initial values of testing data and the back transformed standardized predicted

```
"importing the required libraries"
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import tensorflow as tf
import sklearn
from sklearn import preprocessing
from sklearn import metrics
from sklearn.metrics import r2_score
from sklearn.model_selection import train_test_split
from keras.models import Sequential
from keras.layers import Dense
from keras.optimizers import Adam
from keras.callbacks import EarlyStopping

datf = pd.read_excel(r'/Users/og/Desktop/Workbook1.xlsx') #importing the excel spread
sheet containing the input feature into python using pandas
datf1= pd.read_excel(r'/Users/og/Desktop/Workbook2.xlsx')#importing the excel spread
sheet containing the predictor/output feature into python using pandas
datf2= pd.read_excel(r'/Users/og/Desktop/Workbook3.xlsx')#importing the excel spread
sheet containing the predictor/output feature into python using pandas
x_train, x_test, y_train, y_test = train_test_split(datf, datf1, test_size=0.2,
random_state=37)#splitting the data at random into training and testing dataset and
setting test size to 20%
x_train = preprocessing.scale(x_train) #scaling the training set to avaoid high range
variables having bias effect on model
x_test = preprocessing.scale(x_test) #scaling the test set to avaoid high range
variables having bias effect on model
model = Sequential()
model.add(Dense(4, input_shape=(3,), activation = 'relu'))#building the neural
network
model.add(Dense(4, activation='relu'))
model.add(Dense(4, activation='relu'))
model.add(Dense(1,))
model.compile(loss='mean_squared_error', optimizer='Adam')#optimizing parameters
earlystopper = EarlyStopping(monitor='val_loss', min_delta=0, patience=15, verbose=1,
mode='auto')
history = model.fit(x_train, y_train, epochs = 103, validation_split = 0.2,shuffle =
True, verbose = 0,callbacks = [earlystopper])
pred = model.predict(x_test)
print(pred)
score = np.sqrt(metrics.mean_squared_error(pred,y_test))
print("Final score (RMSE): {}".format(score))
print(pred)
```

Figure 17: line of code written in the Python scripting language for ANN.