Energy

AIMS Energy, 11(3): 425–453.
DOI: 10.3934/energy.2023023
Received: 28 January 2023
Revised: 04 April 2023
Accepted: 22 April 2023
Published: 08 May 2023

http://www.aimspress.com/journal/energy

Research article

Prediction of the equivalent circulation density using machine learning algorithms based on real-time data

Abdelrahman Kandil[1], Samir Khaled[2], and Taher Elfakharany[2,*]

[1] Department of Petroleum Engineering, Faculty of Engineering and Technology, Future University in Egypt (FUE), Cairo 11835, Egypt
[2] Department of Mining and Petroleum Engineering, Faculty of Engineering, Al-Azhar University, Cairo 11835, Egypt

* **Correspondence**: Email: Abdelrahman.kandil@fue.edu.eg; Tel: +0201023815381.

Abstract: Equivalent circulation density (ECD) is one of the most important parameters that should be considered while designing drilling programs. With increasing the wells' deep, offshore hydrocarbon extraction, the costly daily rate of downhole measurements, operating restrictions, and the fluctuations in the global market prices, it is necessary to reduce the non-productive time and costs associated with hole problems resulting from ignoring and incorrect evaluation of ECD. Therefore, optimizing ECD and selecting the best drilling parameters are curial tasks in such operations. The main objective of this work is to predict ECD using three machine learning algorithms: an artificial neural network (ANN) with a Levenberg-Marquardt backpropagation algorithm, a K neighbors regressor (knn), and a passive aggressive regressor (par). These models are based on 14 critical operation parameters that have been provided by downhole sensors during drilling operations such as annular pressure, annular temperature, and rate of penetration, etc. In the study, 4663 data points were selected and included, where 80% to 85% of the data set has been used for training and validation according to the algorithm, and the remaining data points were reserved for testing. In addition, several statistical tests were used to evaluate the accuracy of the models, including root mean square error (RMSE), correlation coefficient (R^2), and mean squared error (MSE). The results of the developed models show various consistencies and accuracy, while the ANN shows a high accuracy with an R^2 of nearly 0.999 for the training, validation, and testing, as well as the overall of them. The RMSE is 0.000211, 0.000253, 0.00293, and 0.00315 for overall, training, validation, and testing, respectively. This work expands the use of artificial intelligence in the gas and oil industry. The developed ANN model is more flexible in response to challenges, reduces dependence on humans, and thus, reduces the chance of human omission, as well as increasing the efficiency of operations.

Keywords: Equivalent circulation density (ECD); artificial intelligence (AI); drilling wells; artificial neural networks (ANN); real-time data

Abbreviations: AARE: absolute average relative error; AI: artificial intelligence; ANFIS: adaptive network-based fuzzy interference system; ANN: artificial neural networks; ARE: average relative error; ECD: equivalent circulation density; EM: equivalent mud weight; FIS: fuzzy inference system; GA: genetic algorithm; LLSVM: least square support vector machine; Mw: mud weight; PSO: particle swarm optimization; PV: plastic viscosity; Q: mud flow rate; R^2: correlation coefficient; RD: relative deviation; RMSE: root-mean-squared error; ROP: rate of penetration; RPM: total downhole rotating speed in revolutions per minute; SD: standard deviation; SPP: standpipe pressure; TFA: total flow area; TVD: true vertical depth; WOB: weight on bit; YP: yield point

1. Introduction

Nowadays, the process of exploring, drilling, producing, and managing reservoirs for gas and oil is a complex task due to the fact that traditional sources of these resources have already been identified and utilized for numerous years. Petroleum engineers are currently striving to adopt sophisticated technologies like machine learning to make informed decisions that minimize unproductive periods and expenses. Machine learning (ML) has been applied to various studies, primarily from 1993 by Hansen [1], where they used artificial neural networks (ANN) to successfully identify primary reflections from seismic data. In 2007, Kononov et al. [2] developed ANNs to compute travel times for a complete 3D volume model, and petrophysical and geomechanical properties of shale reservoirs. Fahd Syed et al. [3] presented a comprehensive review of artificial intelligence (AI) and ML applications in modeling such properties, as well as a systematic publication count that reflected the increasing interest in this subject.

ML has been utilized in drilling engineering for a long time. In 1990, Arehart [4] used ANNs to determine the grade (state of wear) of a drill bit while drilling. ANNs were applied by Moran [5] to predict the rate of penetration (ROP) in drilling operations, enabling a more accurate estimation of drilling time. In 2022, Sprunger et al.. [6] discussed the benefits of implementing machines in hydraulic fracturing operations in shale gas reservoirs. This study covered the design, interpretation, real-time prediction, and re-frac selection of hydraulic fracturing and concluded that machine learning is an accurate method for dealing with a large amount of fracturing data.

Additionally, ML has been previously applied in petroleum production engineering. Denney [7] utilized ANNs to pinpoint the potential restimulation locations for fractures based on the experience gained from the Red Oak oilfield. This paper [8] detailed the use of AI and ML for artificial lift selection and maintenance, equ ent malfunction detection. Additionally, it explained the workflow and effectiveness of each application.

There are many applications of ML in reservoir engineering. For example, in 1996, Yang and Kim [9] developed ANNs to find rock properties. Kohli & Arora [10] developed an ANN model to predict permeability from well logs. Fahd Syed et al. [11] discussed recent developments in the production performance estimation of shale gas in North America using machine learning, including key input parameters and methodology comparisons of different algorithms. In [12], the performance of shale gas reservoirs was evaluated, alongside understanding how artificial intelligence (AI) and machine learning (ML) were used for this purpose. The text also explained different ML approaches such as supervised and unsupervised methods, random forest, support vector machine, boosting techniques, clustering methods, and artificial network-based architecture. Moreover, Taha et al. [13–15] efficiently used ML algorithms to detect and remove noise from the shale gas production data without removing the main trends of the data profile. The results of these works helped improve the forecasting of the production, and can be applied to any other data.

Equivalent circulation density (ECD) is the total hydrostatic pressure of the mud in dynamic conditions. It displays the annular losses in a mud column [16]. ECD is one of the most important parameters for drilling operations monitoring, particularly for the little window between the formation and the fracture pressure [17]. Three crucial drilling fluid densities include the equivalent mud weight (EMW), ECD, and equivalent drilling fluid static density (ESD), with the first two being particularly significant. Estimating automated or real-time EMW during drilling allows for a quick evaluation for ECD, as well as the volume of the cuttings transported by the drilling fluid, which makes the drilling crew able to track and assess the effectiveness of hole cleaning in real-time. Hole cleaning in drilling operations must be optimized to increase the rate of penetration (ROP) and reduce hole issues. However, achieving optimal hole cleaning needs continuous and proper controlling and development, which are difficult to maintain. Poor hole cleaning can result in stuck pipes, reduced circulation, lower ROP, and increased torque and drag. The duration of the operations may be significantly impacted by these factors, which will ultimately raise the overall cost of drilling operations [18].

Drilling mud can detect the hydrostatic column pressure at any drilled depth. Mud density can either balance or unbalance the section of the wellbore. If the used drilling mud weight and ECD were high, overbalanced drilling occurs, which can significantly improve the control during operation [19]. On the other hand, several problems are associated with overbalanced drilling, such as a poor drilling rate, generating irregular differential pressure between both EMW and ECD with the formation, which elevates the potential for differential sticking, and increased mud chemical additions resulting in costs increments [19]. In addition, it can induce formation damage, particularly in the reservoir area. Hence, there are various advantages for calculating the desired mud weight and ECD overbalance throughout the drilling hole section. The best way to prevent these issues, besides achieving safe well drilling, maximizing drilling efficiency, and improving drilling performance, is understanding drilling fluid density, ECD, drilling fluid treatment, and chemical additives, in addition to continuously monitoring hole cleaning.

Marsh [20] described the idea and concept of drilling fluid properties, such as mechanical analysis, weight, consistency, characteristics of desired drilling fluid, treatment of drilling fluid, and chemical factors or additives. It has been found that the annular pressure losses, wellbore geometry, mud characteristics (density and viscosity), mud pumping rate, downhole pressure and temperature, and cuttings concentration have a significant effect on the ECD during the drilling operations [21,22]. Ataga and Ogbonna [23] mentioned that the ECD and drilling pressure can be obtained through a mathematical model estimation, prediction utilizing artificial intelligence (AI) approaches, and by employing downhole sensors while drilling.

Erge et al. [24] and Rommetveit et al. [25] showed that although the downhole measurement is thought to be reliable and precise for ECD values, its application is uncommon due to the costly daily rate and operating restrictions, such as downhole pressure and temperature, that lead to tool failures.

There are several mathematical models for estimating the ECD in the literature, each with a different fluid type and set of input parameters. In the traditional models, numerous values and restrictions have been assumed for the downhole pressure, temperature, and mud types. The material balance calculation has been used to estimate ECD for the mud compositional study [26,27]. Bybee [28] presented a mathematical equation for determining ECD with the mud static density and other mud-related characteristics. The model also takes the effect of solids concentration in the annulus into account.

The mathematical correlations that have been created are restricted to a few features and disregard many other input variables that affect the ECD. These variables include geometry, fluid rheological characteristics, drill string rotation, downhole pressure, temperature, cuttings dispersion, hole cleaning, swab, and the surge of drill pipe movements in the well; other factors have been neglected, but have an impact on the mud density [29,30]. Ignoring these factors will have an impact on ECD prediction, resulting in incorrect evaluation of ECD, and causing well control issues while drilling [31].

The limitations of the ECD downhole device and the inadequate performance of mathematical correlations have prompted the emergence of a new approach for predicting equivalent circulating density (ECD) from drilling data in the petroleum industry. This approach involves the utilization of ANNs, which comprise the interconnected units or nodes, referred to as artificial neurons. These artificial neurons are designed to receive inputs and generate outputs based on their predetermined activation functions through a preconfigured network, which mimics the neurons found in the biological brain. ANNs have been recognized as one of the most successful artificial intelligence systems to date. The underlying principle of ANNs is rooted in their ability to emulate the learning capability of the nervous system, where the input data or signals received are utilized to generalize findings.

Baranthol et al. [32] performed field sampling measurements and validated their measurements using computer models. In [33], a correlation using AI and an ANN was created to estimate mud density as a function of the drilling fluid type, various pressures, and different temperatures. In [34], a study was conducted on the ECD under extreme heat and pressure to assess the ECD during drilling, in which a temperature profile was created using the CrankNicolson numerical discretizing scheme in the DD-Simulator, which is used to simulate the circulation of the wellbore. Ahmed et al. [30] created a model to predict how drill-string rotation will affect ECD. They discovered that the ECD would be more effective with a greater rotation and pump flow rates.

Elzenary et al. [35] used AI technologies, such as adaptive neuro-fuzzy inference systems (ANFIS) and ANNs, to find a relationship to predict the ECD while drilling and to compare the measured outcomes with these predictions. Ahmadi [36] used the ANFIS, enhanced particle swarm optimization ANFIS (PSO-ANFIS), and least square support vector machine (LLSVM) algorithms to calculate the ECD using only the mud's initial density, temperature, and pressure. Using the initial mud density, pressure, and temperature data, Ahmadi et al.. [37] investigated the prediction of ECD using PSO-ANN, fuzzy inference system (FIS), and a hybrid of genetic algorithm (GA) and FIS (GA-FIS). In terms of the R^2 and average absolute percentage error between the actual and projected values of ECD, the PSO-ANN model demonstrated a good level of prediction ability.

A model for ECD prediction utilizing the two AI techniques, ANN and ANFIS, was presented by Abdelgawad et al. [22]. While the ANFIS model was created using five membership functions with a gaussian membership function (gaussmf) as the input membership function and a linear type as the output membership function, the study provided an ECD-ANN model with one hidden layer with 20 neurons.

The radial basis function was used by Rahmati and Tatar [38] to create an ECD prediction model that demonstrated a strong prediction performance with an R^2 of 0.98 and an AAPE of 0.22. They used three different machine learning algorithms to predict the ECD while drilling horizontal sections using surface drilling data.

In Alkinani et al. [39], drilling parameters were used in addition to the hydraulics and mud properties, such as mud pumping rate, rheology of the mud (density, plastic viscosity, and yield point), total flow area (TFA) for the bit nozzles, revolutions per minute (RPM) for the drill pipe, and weight on bit (WOB), to predict the ECD using an ANN model that had just one hidden layer with 12 neurons.

Alsaihati et al. [40] developed a model using seven drilling parameters as inputs, where the SVM, RF, and FN were used to generate the model. The accuracy of the model that was created varied from an R^2 of 0.95 to 0.99 and the root mean square error (RMSE) ranged from 0.23 to 0.42. The author of this study did not create any machine learning-free ECD equations that could be employed. Gamal et al. [17] used only mechanical drilling parameters, such as mud pumping rate (GPM), ROP, drill string speed in RPM, stand-pipe pressure (SPP), WOB, and drilling torque, to predict the ECD using machine learning approaches such as ANNs and ANFISs, with an R^2 of 0.98 and an average absolute percentage error (AAPE) of 0.3% for the ANN, while the ANFIS recorded an R^2 of 0.96 and an AAPE of 0.7.

These outcomes demonstrated that ANN outperformed the other techniques. The literature makes it obvious that the AI models improved the ECD prediction; however, the models differ in terms of the input parameters, the data fed into the models, and the ECD prediction approaches. One of the issues with multiple studies in the literature is the need to utilizedownhole pressure and temperature as inputs to the prediction models. From an operational perspective, downhole sensors are needed to obtain these parameters with high accuracy for a better ECD prediction, and this will increase operational costs and time for the data collection.

The present study introduces a novel approach in which machine learning algorithms are primarily reliant on real-time drilling parameters from downhole sensor tools, including inclination angle, true vertical depth (TVD), annular temperature, mud flow rate (Q), mud weight (Mw), plastic viscosity (PV), yield point (YP), gel strength (0/10; referring to measurements taken after 10 seconds and 10 minutes, respectively), total flow area (TFA), annular pressure, standpipe pressure (SPP), rate of penetration (ROP), and total downhole revolutions per minute (RPM). To this end, a total of 4663 data points, obtained through measurement-while-drilling (MWD) from an offshore deviated well in Norway, were filtered and utilized to develop models for predicting ECD. The performance of these models was compared against each other, and the model demonstrating the highest accuracy underwent several statistical tests and was compared with other developed models [36].

2. Data description, cleaning, and filtration

The study used actual drilling data that were collected by downhole sensors from a deviated section well during a drilling phase in Norway. The strategy that was followed while conducting this study is depicted in Figure 1. The input data selected based on experts' are shown in Figure 2, which represents inclination angle, annular temperature, annular pressure, TVD, Q, MW, PV, YP, gel (0/10), TFA, SPP, ROP, and RPM.

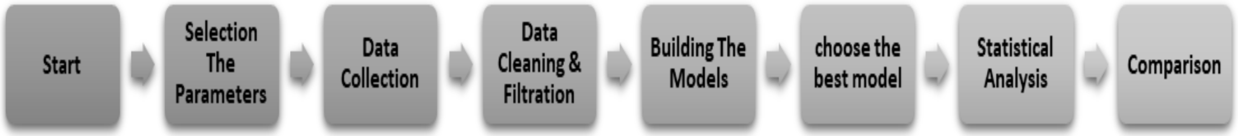

Figure 1. Strategy of the study.

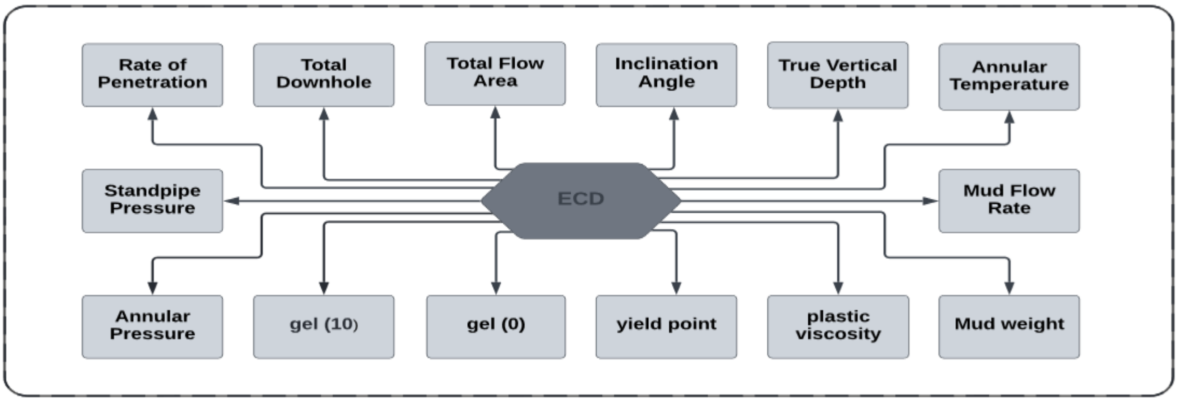

Figure 2. Input parameters.

Key performance indicators (KPIs) have been determined by correlation analysis, as shown in Figure 3 and Table 1. To verify collinearity between the variables, scatterplots of each independent variable were created. According to the heat map, it is clear that some factors such as SSP, TFA, mud flow, and Mw have a variable direct correlation with the ECD, while some of the other parameters (PV, gel 10, inclination, annular pressure, gel 0, hole depth and, RPM) show an inverse correlation with the ECD; alternatively, annular temperature, YP, and ROP show a weak correlation with ECD.

Cleaning and filtering the data produces high-quality input parameters for the model. Therefore, a lot of effort has been done to clean the data, develop the AI model's training procedure, and maximize the correctness of the model results using the trained algorithm. If the accuracy is poor, then a process of retraining should be carried out in order to improve the accuracy and obtain the ideal model parameters for the ECD prediction's performance. Data cleaning is the process of locating and fixing problems with the data, such as data losses, corruptions, and errors, through employing statistical techniques like outlier detection and imputation. One of the cleaning processes that the model applied is the exclusion of the point if all the input parameters do not exist. Figures 4 and 5 show the preprocessed (filtration and removing) data for the inclination angle versus TVD depth before and after processing as an example of the cleaning process which has been done for the input parameters.

Table 1. Correlations between parameters.

Parameters	ECD
Inclination	–0.506
Annular Temp.	–0.059
Mud Flow	0.823
Hole Depth	–0.51
PV (CP)	–0.688
Yp (Pa)	–0.057
Gel 0 (Pa)	–0.335
Gel 10 (Pa)	–0.585
Annular Press.	–0.361
SSP	0.795
ROP	0.096
RPM	–0.256
ECD	1
Mw (Sgr)	0.477
TFA(in^2)	0.894

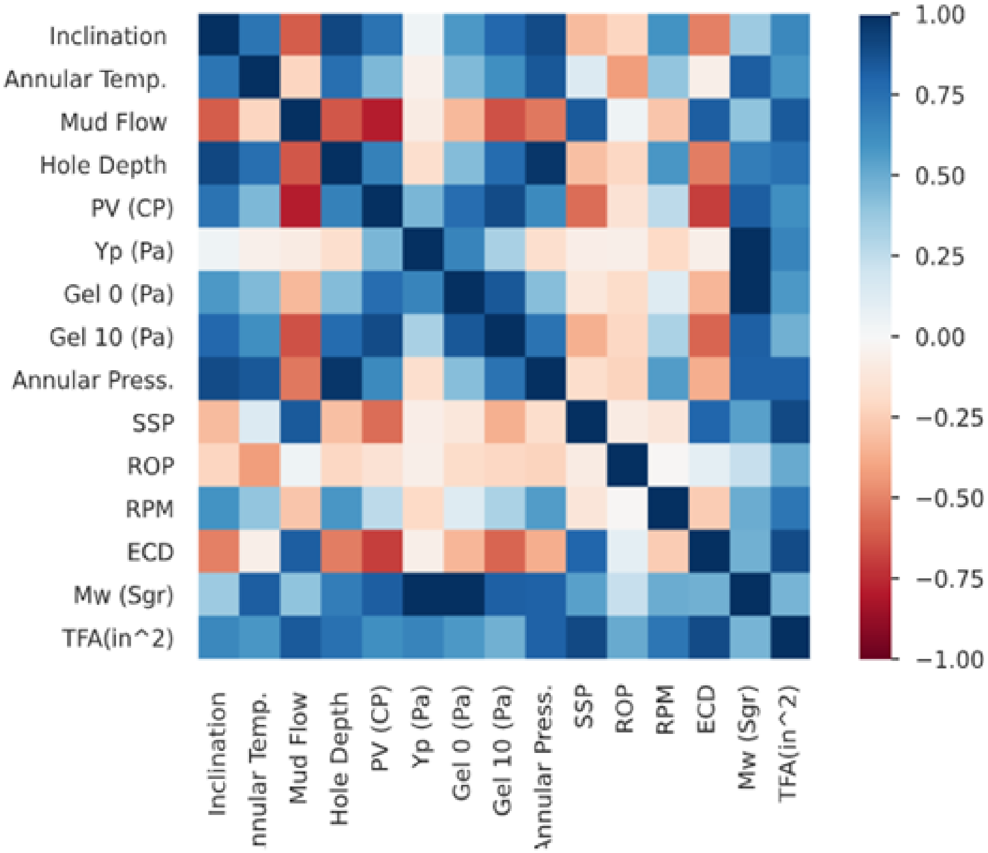

Figure 3. Heatmap correlations.

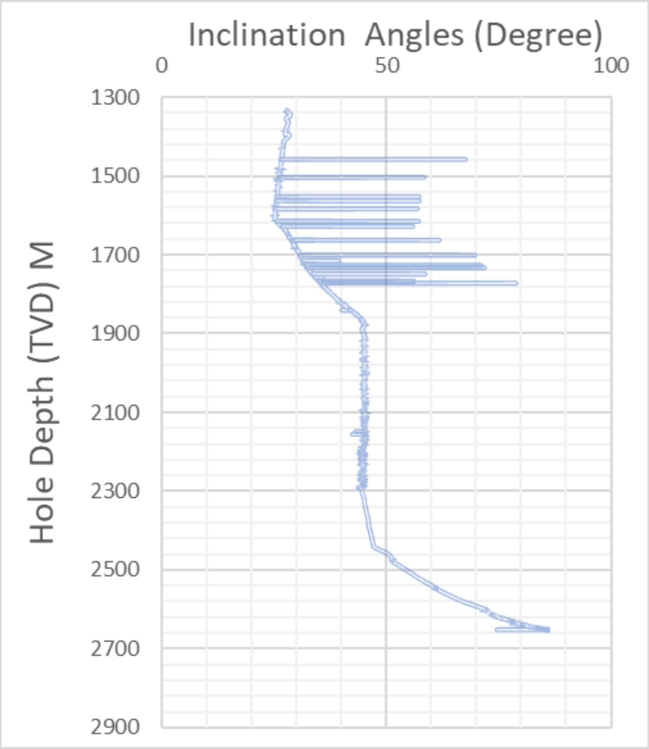

Figure 4. Inclination angle before processing.

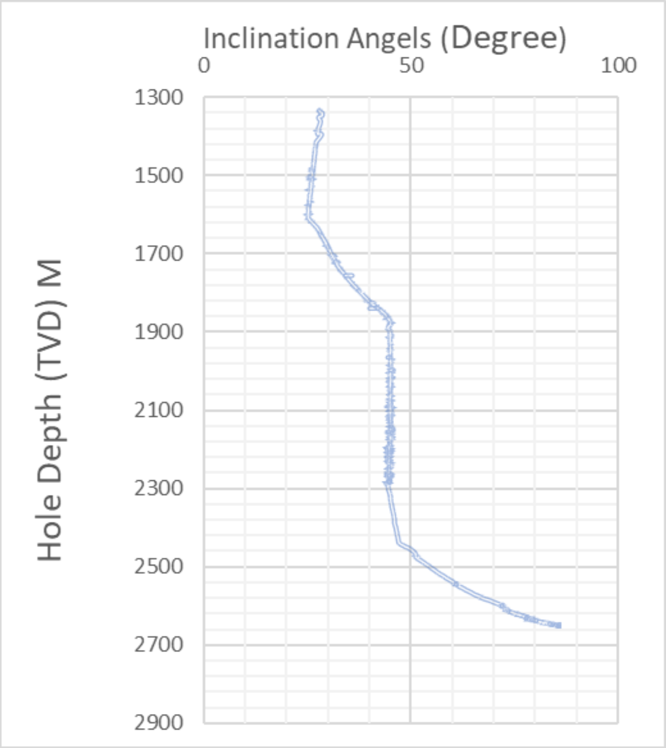

Figure 5. Inclination angle after processing.

3. Statistical analysis for parameters

The statistics features are crucial to the theoretical, practical, and long-term development of machine learning. These features can be used in problem framing, data comprehension, preparation, model evaluation, configuration, model selection, model presentation, and model predictions. Tables 2 and 3 show the statistical parameters of the whole data set. The first and second rows show the maximums and minimums of the data set, which shows the range of the selected data. The standard deviation is a measurement of dispersion, or how the data are distributed around the mean. The degree to which the data deviates from the normal distribution is measured by skewness. Any skewness is generally undesirable in models since it leads to a variance in estimation.

Table 2. Data statistical distribution.

	TFA	Annular Pressure	SPP	ROP	RPM	Q	ECD	ECD
	(in^2)	(kPa)	(kPa)	(m/h)	(rpm)	(l/min)	(g/cm^3)	(Ib/ft^3)
Maximum	1.51	35313.56	28702.32	54.649	186.14	4529.96	1.563	97.62
Minimum	1.075	18022.89	11568.64	0.6797	2.75	2017.81	1.319	82.39
Skewness	–0.6749	–0.6538	–1.413	1.227	–4.657	–0.6883	–0.559	–0.559
Median	1.51	30757.51	24239.07	16.867	160	4489.99	1.500	93.656
Average	1.362	29893.45	24436.38	17.482	162.92	4116.84	1.447	90.376
Kurtosis	–1.537	–0.5951	5.396	2.751	28.273	–1.236	–1.630	–1.630
Range	0.435	17290.67	17133.67	53.97	183.39	2512.15	0.2439	15.230
Mode	1.51	34818.52	24015.06	16.867	159	4489.99	1.500	93.65
Standard deviation	0.2053	4270.19	2544.86	7.445	15.178	499.26	0.0791	4.943

Table 3. Data statistical distribution.

	Inclination	Annular	TVD	Mw	PV	Yp	Gel 0	Gel 10
	dega	temperatures (°C)	(m)	(Sp.gr)	(CP)	(Pa)	(Pa)	(Pa)
Maximum	85.74	99	2651.881	1.51	37	20.1	8.6	10
Minimum	25.14	47	1333.353	1.39	22	6.7	3.3	3.8
Skewness	0.7369	–0.7436	–0.1084	–4.105	–0.3435	–0.4195	–1.643	–0.7464
Median	45.02	86	2122.542	1.51	32	13.9	7.6	8.6
Average	49.801	84.472	2124.218	1.504	32.971	13.796	7.437	8.4449
Kurtosis	–0.5924	0.1373	–1.127	16.562	0.5800	6.6708	4.272	0.7163
Range	60.6	52	1318.527	0.12	15	13.4	5.3	6.2
Mode	45.08	94	-------	1.51	37	13.9	8.1	10
Standard deviation	18.434	10.713	385.349	0.0218	3.0250	1.6091	0.9310	1.317

4. Models development

Three models have been developed by using three machine learning algorithms. The first model used the passive aggressive regressor (par) algorithm. For the second model, the K neighbors regressor (knn) algorithm has been applied. The final model adopted the ANN with a Levenberg-Marquardt backpropagation algorithm.

4.1. Passive aggressive regressor algorithm

This model was developed using a Python code. The total data points have been split into 80% for training and validation and 20% for the testing model. The selected sets of data were changed in an alternating pattern ten times, and the result of these folds is shown in Table 4.

Table 4. Results of folds for (Par).

Fold	MAE	MSE	RMSE	R^2
0	0.0588	0.0045	0.0672	0.282
1	0.0603	0.0047	0.0686	0.2814
2	0.0585	0.0044	0.0662	0.2961
3	0.0658	0.0054	0.0738	0.1051
4	0.0649	0.0053	0.0729	0.1681
5	0.0617	0.0049	0.0701	0.1705
6	0.0521	0.0036	0.0596	0.4115
7	0.0579	0.0044	0.0666	0.261
8	0.0624	0.0049	0.0701	0.2253
9	0.0739	0.0069	0.0831	–0.2551
Mean	0.0616	0.0049	0.0698	0.1946

Figure 6 shows the amount of loss in the value of residuals between the predicted ECD and the actual value of ECD. Notice that the residual values between the real and expected values by this model are very large. In addition, the accuracy of this model is very weak and unreliable, with a flawed value for the R^2, which is close to 25%, as shown in Figure 7.

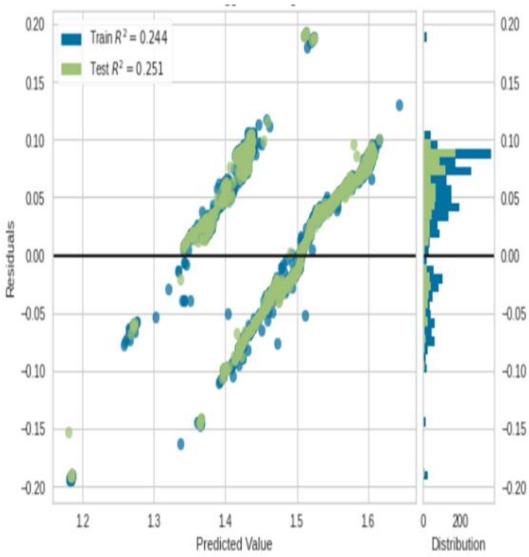

Figure 6. Residuals for passive-aggressive regressor.

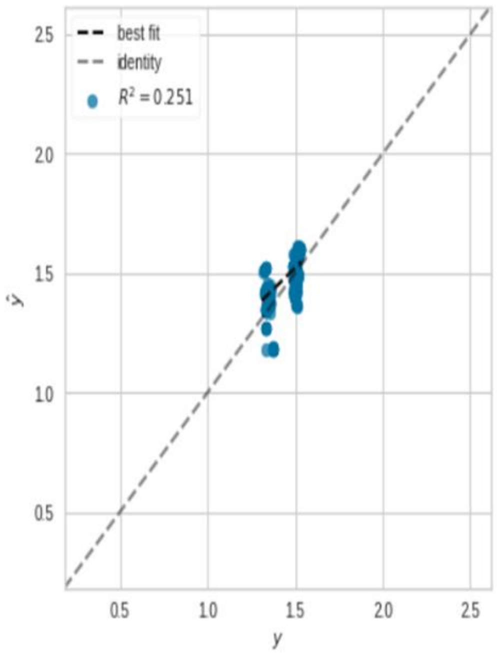

Figure 7. Prediction error for passive-aggressive regressor.

4.2. K Neighbors regressor algorithm

This model was also built using a Python script. The total data points have been divided into 80% for training and validation and 20% for testing. The selected sets of data were alternatively altered ten times, as shown in Table 5.

Table 5. Results of folds for knn.

Fold	MAE	MSE	RMSE	R^2
0	0.0015	0.0001	0.0088	0.9877
1	0.0022	0.0001	0.0103	0.9839
2	0.0009	0	0.0017	0.9995
3	0.0015	0	0.0047	0.9963
4	0.0013	0	0.003	0.9986
5	0.0013	0	0.0041	0.9971
6	0.0014	0	0.0039	0.9975
7	0.0011	0	0.0028	0.9987
8	0.0012	0	0.0024	0.9991
9	0.0011	0	0.0031	0.9982
Mean	0.0014	0	0.0045	0.9957

According to Figure 8, it is clear that the residual values between the predicted and actual ECD is centered around the values ranges from 0.05 to –0.05 (i.e., very small), and that this model can be trusted in the prediction for ECD.

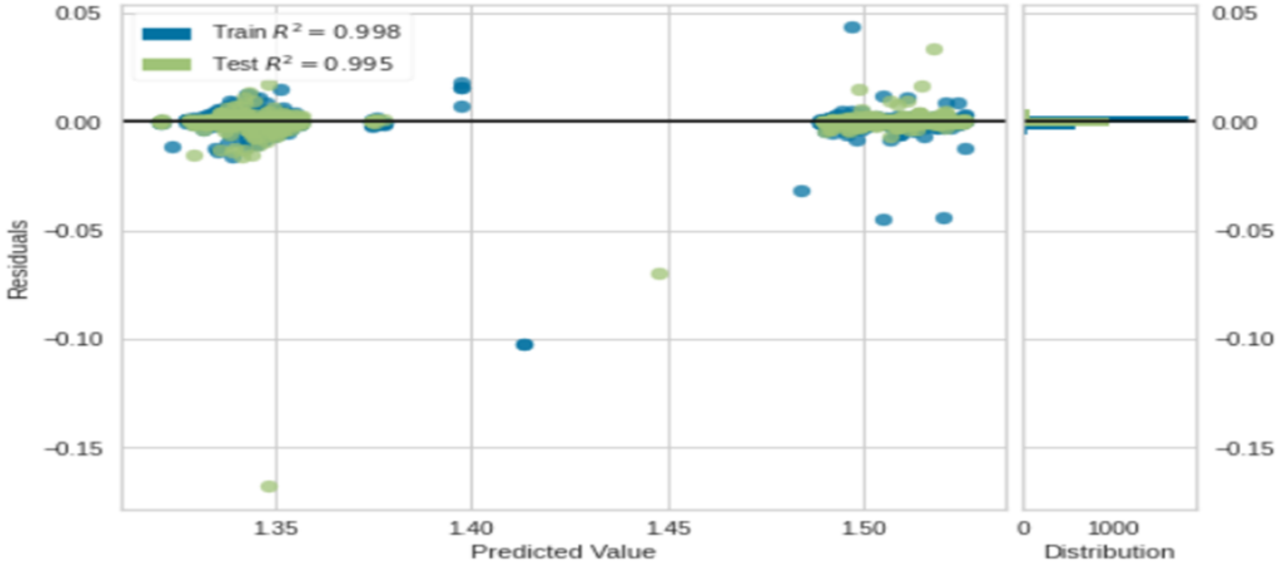

Figure 8. Residuals for K neighbors regressor.

Additionally, Figure 9 shows a high value of the R^2 for the values of ECD expected by this model.

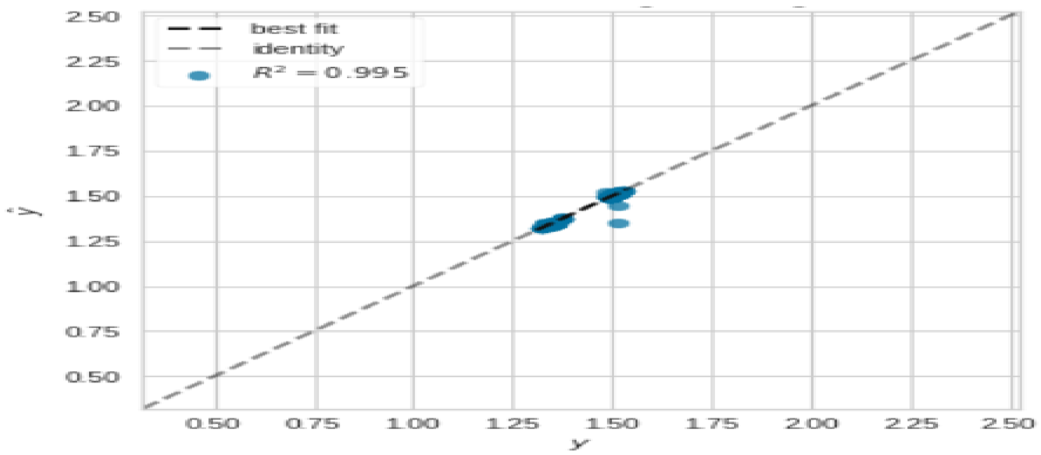

Figure 9. Prediction error for K Neighbors Regressor.

4.3. Artificial neural network (ANN)

This model was created using a MATLAB script. A total of 4,663 points were obtained after data cleaning and filtering. The input data set is a function of the depth, which is a representation of the true vertical well depth. If either the input or output data for neural networks are very small, very large, or not normally distributed, data scaling should be performed [41,42]. Therefore, the input parameters were normalized using Eq 1 to be between 0 and 1, while the output parameter (ECD) was transferred according to Eq 2, for a faster and more efficient training process.

$$Xi, normalized = \frac{X_i - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

where i is the index number for each input parameter: inclination angle, annular temperature, annular pressure, TVD, Q, MW, PV, YP, gel (0/10), TFA, SPP, ROP, and RPM.

$$Yt = \log(y+1) \tag{2}$$

where, y is the value of equivalent circulation density and Yt is the transformed value of the equivalent circulation.

4.3.1. Building the model

The total dataset was randomly partitioned into three subsets: 70% for training, 15% for validation, and 15% for testing purposes. The training data were utilized to develop the ANN model, with the target output being used to adjust the weights of each input. The calibration of these weights was performed using the feedforward backpropagation algorithm, which utilizes error backpropagation through the network to adjust and fine-tune the weights. It should be noted that the occurrence of overtraining (either overfitting or underfitting) in the model can diminish its ability to provide accurate estimations when applied to diverse conditions. Therefore, the training process is halted when the network's generalization reaches a certain threshold, and the verification set (data not employed to construct the network) is employed to assess the model's generalization capabilities. Subsequently, the accuracy of the ANN model is evaluated using the testing set. The architecture of the ANN model is illustrated in Figure 10.

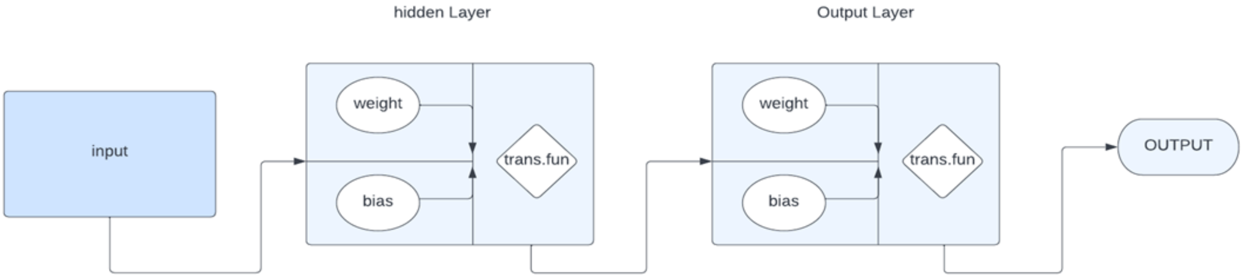

Figure 10. ANN Architecture.

A MATLAB script was created to go over 3,840 possible combinations of different parameters (transfer function, training algorithm, and the number of neurons in the hidden layer(s)) as shown in Table 6 for one hidden layer network, and developed for two hidden layer networks to reach more than 230,000 possible combinations with different parameters, referred to in Table 7.

The approach adopted in developing the model is depicted in Figure 11, which commences with a basic model comprising of a single hidden layer. The number of neurons is initially incremented, followed by the modification of the training algorithm, activation function of the hidden layer, and ultimately the activation function of the output layer. Subsequently, all possible combinations of the single hidden layer are explored, following which, the model is upgraded to two hidden layers, with modifications made to the number of neurons in the first hidden layer, the number of neurons in the second hidden layer, the training algorithm, activation function of the first hidden layer, the activation function of the second hidden layer, and the activation function of the output layer, in that order. Each combination is evaluated using the RMSE (Eq 3) and the R^2 (Eq 4). The combination yielding the lowest RMSE and highest R^2 was selected. The model and algorithm were then constructed based on this process, with the evaluation continuing until a satisfactory level of RMSE and R^2 was attained (i.e., R^2 overall > 0.9995 and RMSE < 0.0005). The program is designed to minimize the time required to construct and train the model.

Table 6. Different combinations of one hidden layer.

Parameters	Range
Number of neurons	1–60
Training algorithm	Trainlm (Levenberg-Marquardt Backpropagation) Trainrp (Resilient Backpropagation) Traingdx (Gradient Descent with Momentum and Adaptive Learning Rate Backpropagation) Trainbr (Bayesian Regularization Backpropagation)
Activation function of hidden layer	Logsig-Purelin-Radbas-Tansig
Activation function of output layer	Logsig-Purelin-Radbas-Tansig

Table 7. Different combinations of two hidden layers.

Parameters	Range
Number of neurons of frist layer	1–30
Number of neuron of second layer	1–30
Training algorithm	Trainlm-Trainrp-Traingdx-Trainbr
Activation function of frist hidden layer	Logsig-Purelin-Radbas-Tansig
Activation function of second hidden layer	Logsig-Purelin-Radbas-Tansig
Activation function of output layer	Logsig-Purelin-Radbas-Tansig

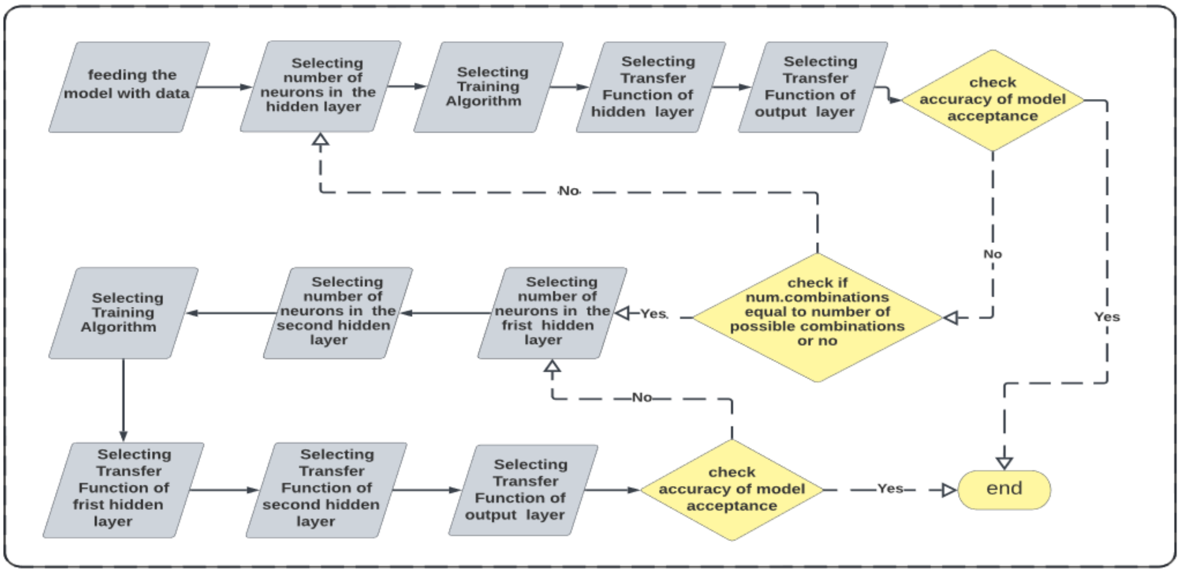
feeding the model with data
Selecting number of neurons in the hidden layer
Selecting Training Algorithm
Selecting Transfer Function of hidden layer
Selecting Transfer Function of output layer
check accuracy of model acceptance
No
No
Yes
check if num.combinations equal to number of possible combinations or no
Yes
Selecting number of neurons in the frist hidden layer
Selecting number of neurons in the second hidden layer
Selecting Training Algorithm
No
Selecting Transfer Function of frist hidden layer
Selecting Transfer Function of second hidden layer
Selecting Transfer Function of output layer
check accuracy of model acceptance
Yes
end

Figure 11. The strategy following in developing the modeling.

$$RMSE = \left(\frac{1}{N} * \sum_{j=1}^{N}(E^{Predicted} - E^{actual})^2\right)^5 \tag{3}$$

$$R^2 = \left\{1 - \frac{\sum_{j=1}^{N}\left(E^{\,actual} - E^{\,predicted}\right)^2}{\sum_{j=1}^{N}\left(E^{\,actual} - E^{\,mean\ actual}\right)^2}\right\} \tag{4}$$

$$\text{loss Function} = \left(\frac{1}{\text{N}}\sum_{\text{j}=1}^{\text{N}}\left(E^{\,actual} - E^{\,predicted}\right)^2\right) \tag{5}$$

where N is the number of data points tested, $E^{Predicted}$ is the predicted equivalent circulation density, E^{actual} is the corresponding to the actual ECD, and $E^{\,mean\ actual}$ is the average of the actual ECD.

4.3.2. Results and discussion

The ANN model was created to predict the ECD according to Eq 6 with one hidden layer. The number of neurons in this hidden layer was optimized to be 34 according to Figures 12 and 13. The structure of this model is shown in Figure 14. The Levenberg-Marquardt backpropagation optimization was used as a training algorithm. TANSIG and LOGSIG were selected to be the hidden and output layers activation functions, respectively. This model displayedhigh accuracy and certainty when compared with the other two developed models.

$$E = Exp\left[\left[\sum_{j=1}^{j} W_{2j}\, TANSIG\left(\sum_{K=1}^{K} W_{1k} * X_K + b_{1j}\right)\right] + b_2\right] - 1 \tag{6}$$

where Exp is the exponential function, j is the number of neurons in the hidden layer, W_{2j} is the weight of the resulting layer, k is the number of input parameters, $W1$ is the weight of the hidden layer, Xj is the standard input parameters (inclination angle, TVD, annular temperature, Q, MW, PV, YP, gel (0/10), TFA, annular pressure, SPP, ROP, and RPM), $b1$ is the bias of the hidden layer, and $b2$ is the bias of the resulting layer. The weights and bias values are illustrated in Appendix A.

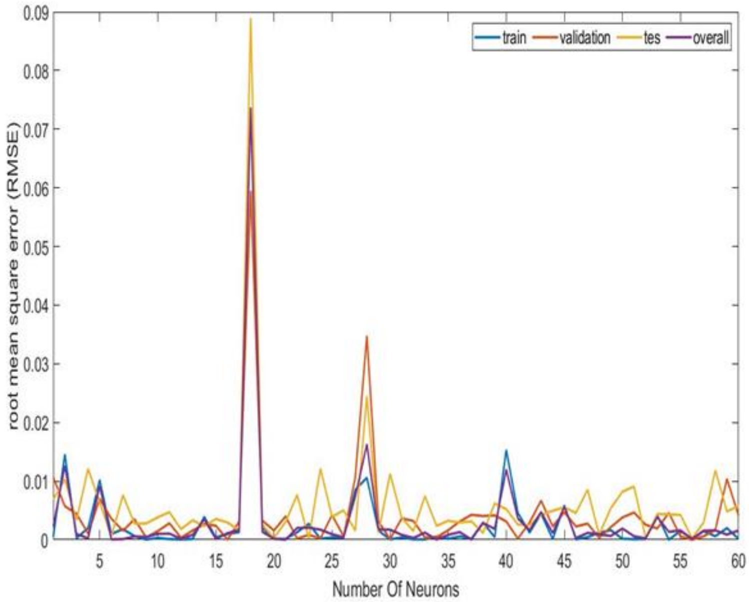

Figure 12. Selecting number of neurons according RMSE.

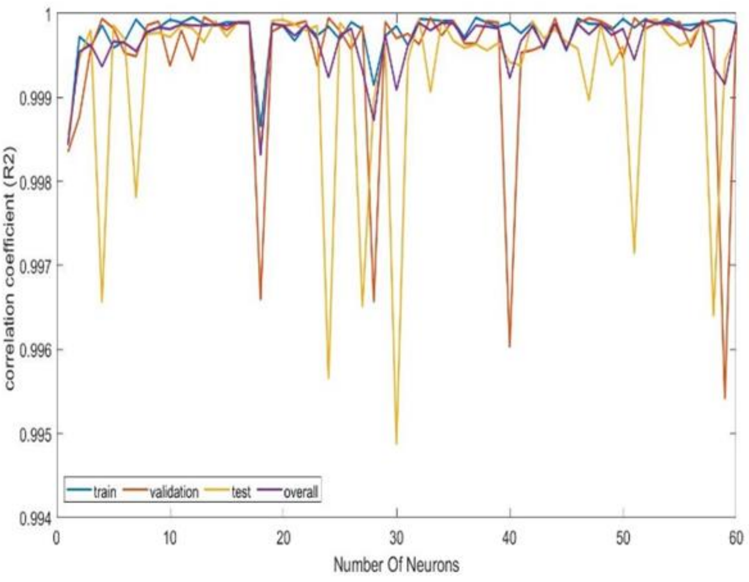

Figure 13. Selecting number of neurons according to the R^2.

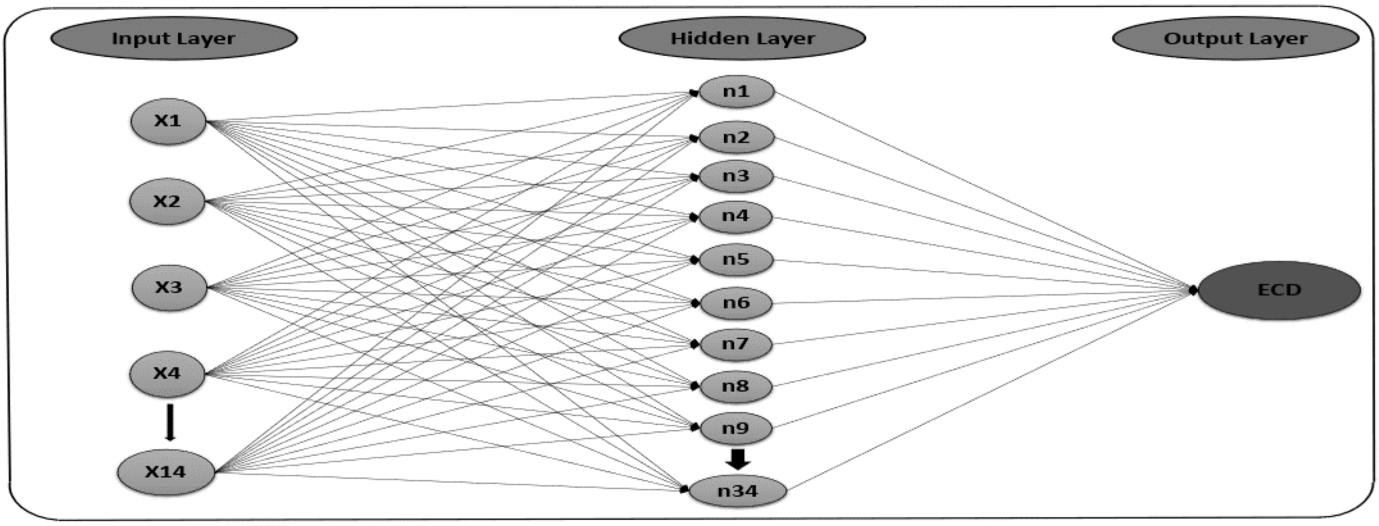

Figure 14. Structure of the model.

The results showed a strong accuracy for the model in terms of R^2, RMS, and loss function (Eq 5) for the training validation, testing, and overall, as shown in Table 8 and Figures 15–17.

Table 8. R^2, RMSE and loss function results.

	Training	Validation	Testing	Overall
Correlation coefficient (R^2)	0.9999	0.9996	0.9999	0.9998
Root mean square error (RMSE)	0.000253	0.00293	0.00315	0.000211
Loss function	0.001981	0.009738	0.001611	0.003088

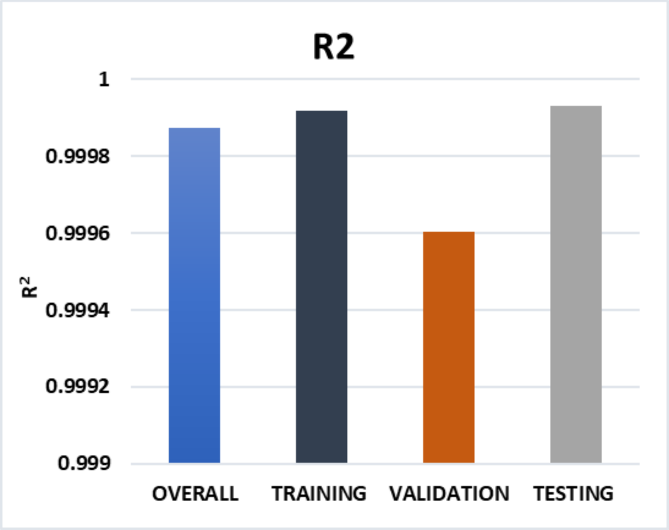

Figure 15. Correlation coefficient.

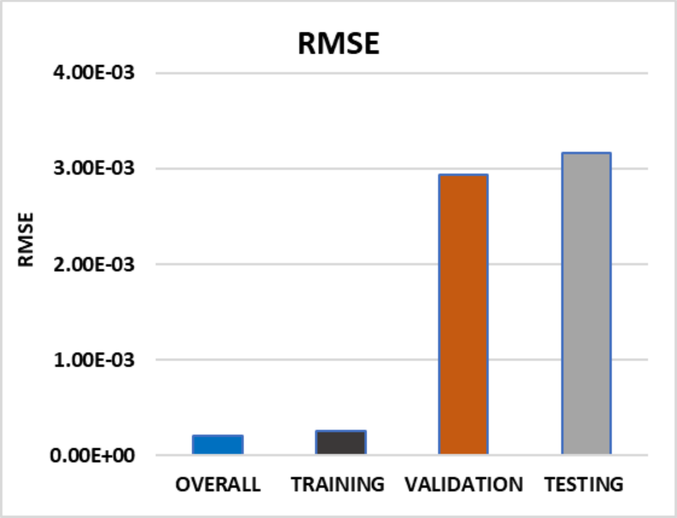

Figure 16. Root mean square error.

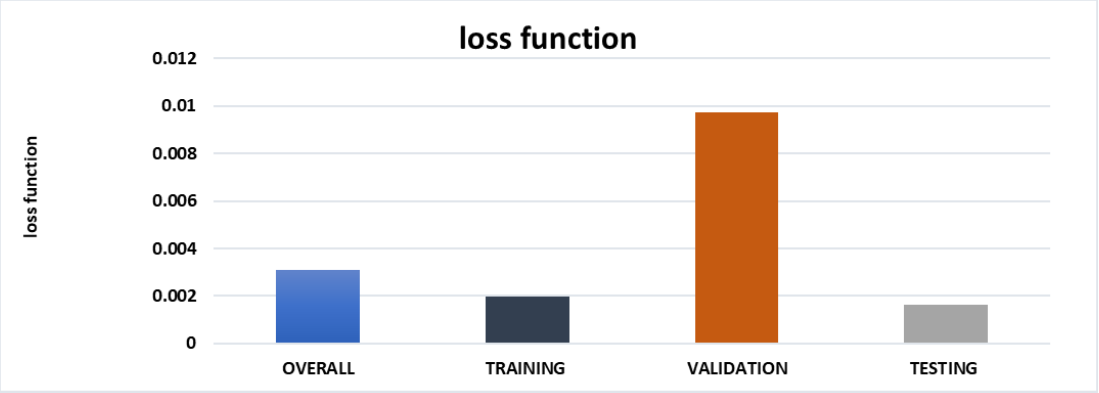

Figure 17. Loss function.

Figure 18 represents the cross-plots for the ANN model results for the model training, validation, testing, and overall processes for estimating ECD values. The plots reveal that there is no unacceptable overestimation or underestimation of the anticipated value. Furthermore, they demonstrated the models' remarkable accuracy performance of predictions.

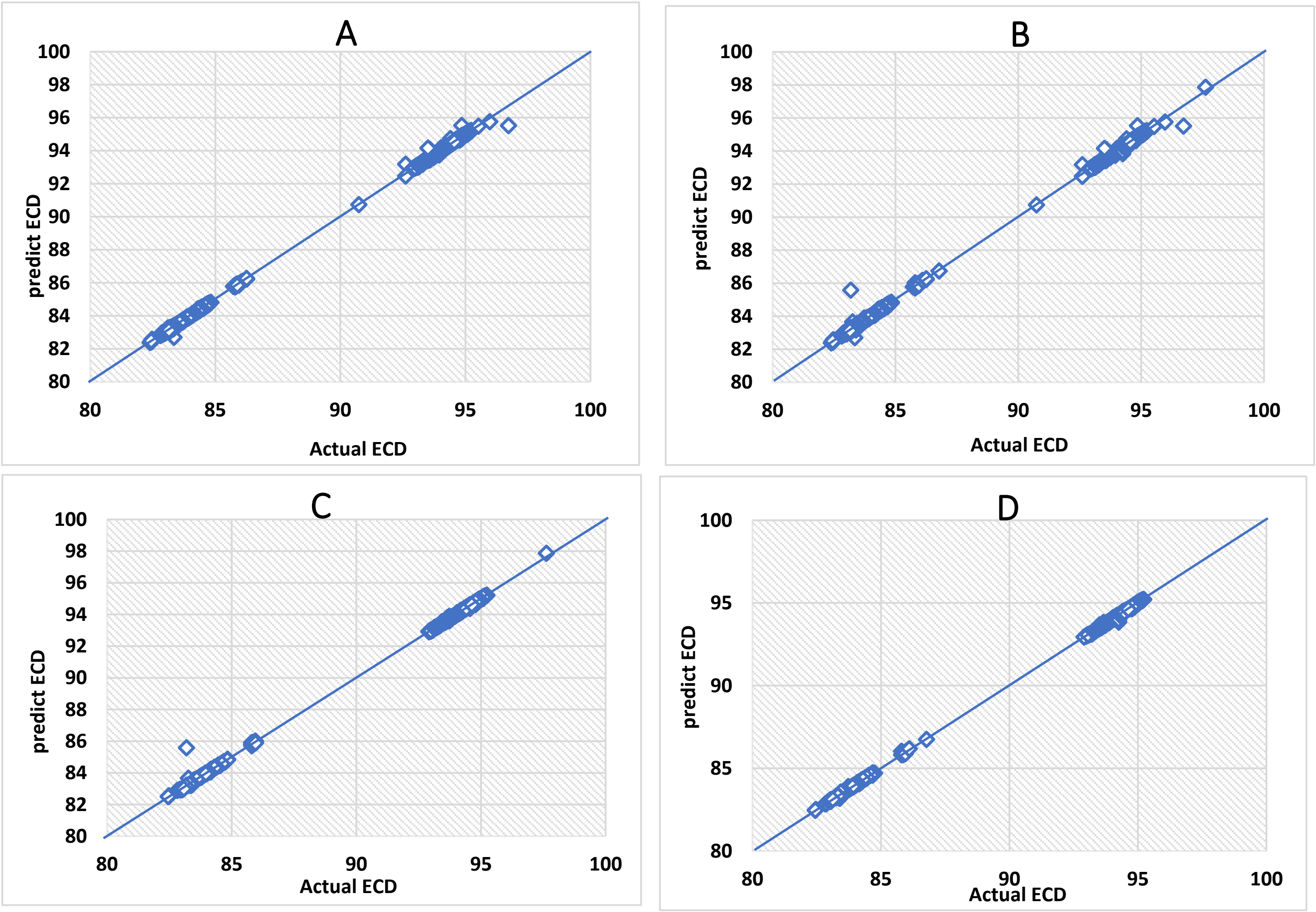

Figure 18. Predict ECD with actual ECD (A) for training, (B) for overall, (C) for validation, (D) for testing.

Figure 19 displays the agreement between the actual and predicted ECD values along the number of data points which have been used in the overall, training, validation, and testing of the ANN model.

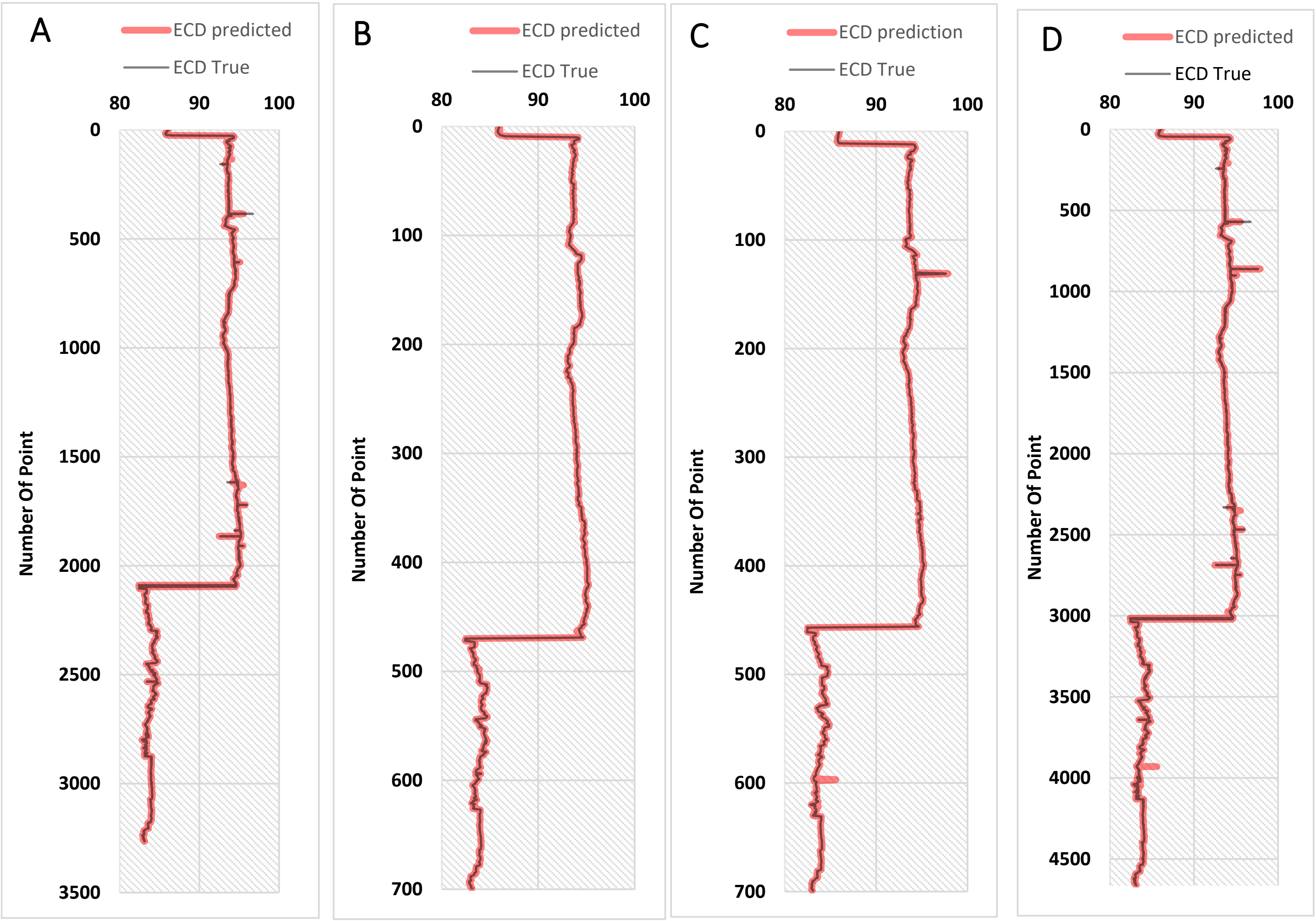

Figure 19. (A) for training, (B) for testing, (C) for validation, (D) for overall.

4.3.3. Statistical analysis for model results

To test the model, several statistical comparisons between the actual and anticipated data have been made. The absolute average relative error (AARE), standard deviation (SD), average relative error (ARE), relative deviation (RD), and mean square error (MSE) are used to evaluate the model based on the following Eqs (7–11). Table 9 and Figures 20–24 show the corresponding values of those statistical features.

$$\mathrm{AARE} = \left(\frac{1}{\mathrm{N}} * \sum_{\mathrm{j=1}}^{\mathrm{N}} \left|\frac{E^{predicted} - E^{actual}}{\mathrm{E_j\ actual}}\right|\right) * 100 \tag{7}$$

$$\mathrm{SD} = \left(\frac{1}{\mathrm{N\text{-}1}} * \sum_{\mathrm{j=1}}^{\mathrm{N}} \left(\frac{\mathrm{E^{Predicted}\text{-}E^{actual}}}{\mathrm{E^{actual}}}\right)^2\right)^{.5} \tag{8}$$

$$ARE = \left(\frac{1}{N} * \sum_{\mathrm{j}=1}^{N} \frac{E^{predicted} - E^{actual}}{E^{\ actual}}\right) * 100 \tag{9}$$

$$\mathrm{RD} = \frac{E^{predicted} - E^{actual}}{E^{\ actual}} * 100 \tag{10}$$

$$\mathrm{MSE} = \left(\frac{1}{N} * \sum_{j=1}^{N} (E^{Predicted} - E^{actual})^2\right) \tag{11}$$

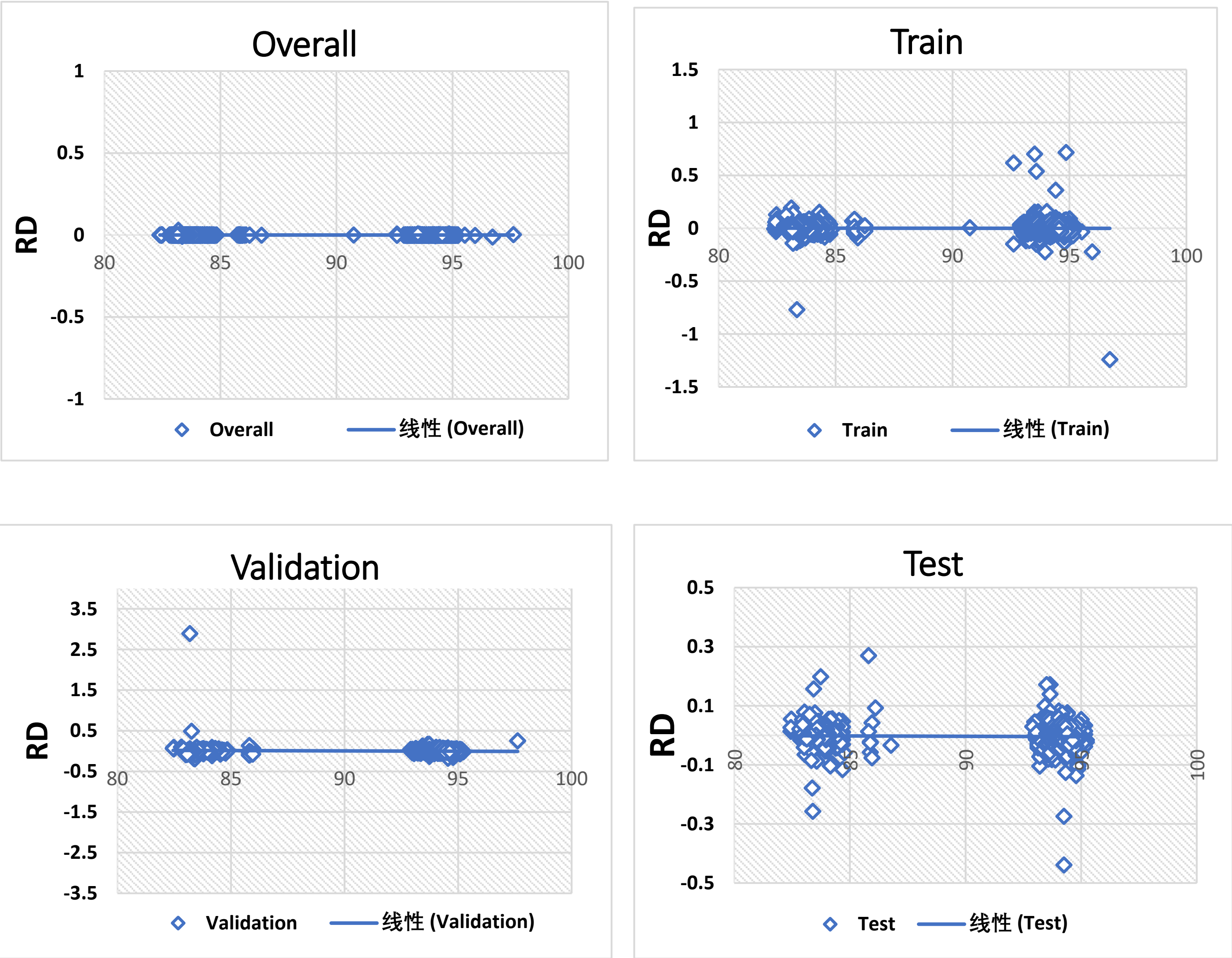

Figure 20. RD versus actual ECD for training, testing, validation, and overall.

Table 9. Statistical test of the model.

	ARE	AARE	SD	MSE
Overall	–0.0001286	0.027362	0.000631	0.003089
Training	–0.0002421	0.025952	0.000482	0.001981
Validation	0.00369188	0.032496	0.001175	0.009738
Testing	–0.0034194	0.028814	0.000443	0.001610

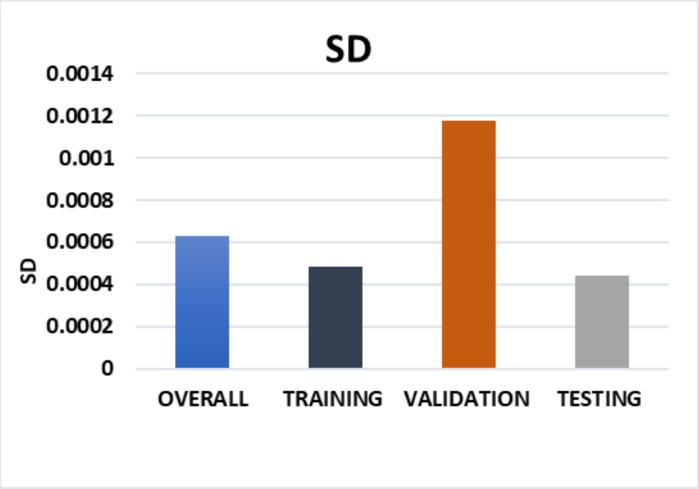

Figure 21. Standard deviation (SD) for overall, training, validation, and testing dataset.

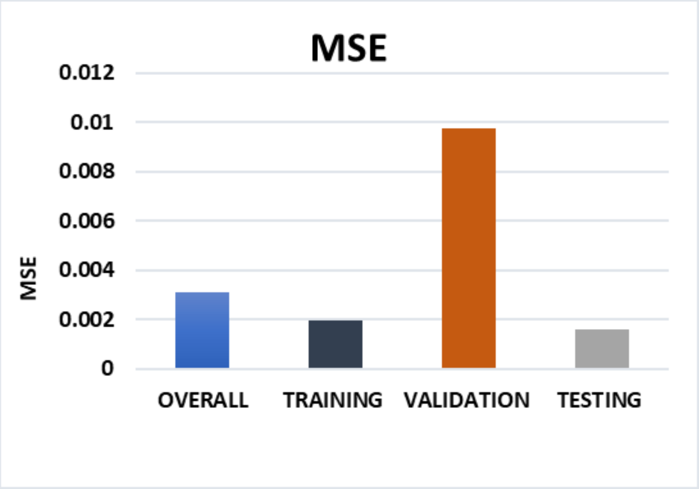

Figure 22. Mean square error (MSE) for overall, training, validation, and testing dataset.

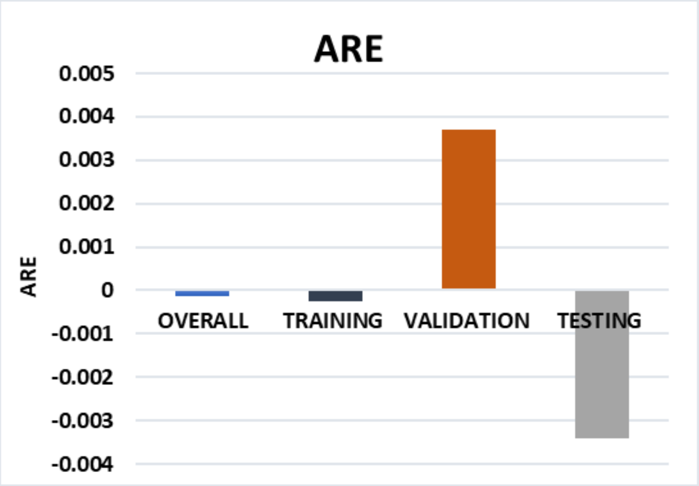

Figure 23. Average relative error for overall, training, validation, and testing dataset.

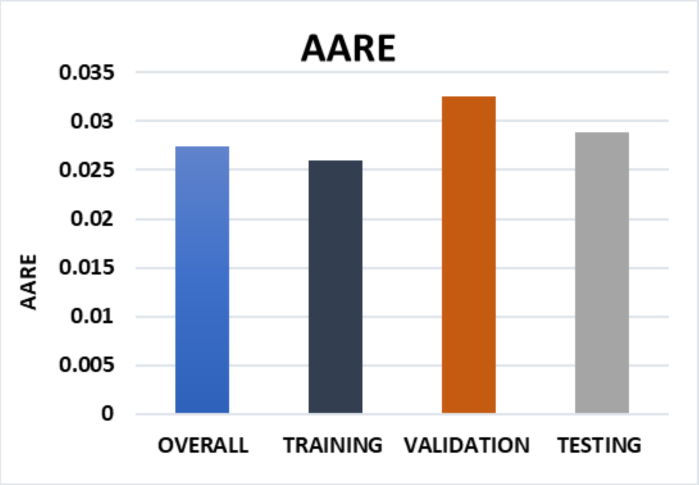

Figure 24. Absolute average relative error for overall, training, validation, and testing dataset.

4.3.4. Comparison with other models

When compared to the other models, the developed ANN model has promising results. The comparison was based on the input parameters feeding the models, the number of data points used for creating these models, and the sources of this data. Additionally, this comparison takes into account the types of ML algorithms that were used to develop models. Furthermore, the results of several statistical tests were used to evaluate the accuracy of models. The developed model shows a high accuracy when compared to other models. This comparison is shown in Table 10.

Table 10. Comparing between the developed ANN model and other models.

	ML algorithm	Inputs parameters	Number of data point	Source of dataset	R^2	Statistical test
Ahmadi et al.[21]	FIS	initial density pressure	664		0.7273	67.0907 AARE
	PSO-ANN	temperature		from the literature	0.9964	0.0001374 AARE
	GA-FIS				0.9397	0.091 AARE
Abdelgawad et al.[22]	ANN	drill pipe pressure	2376	Surface measurements	0.9982	0.2237 AARE
	ANFIS	mud weight rate of penetration			0.98	0.22 AARE
Rahmati & Tatar [38]	radial basis function	Temperature initial density type of mud pressure	884	from the literature	0.99	0.00000166 MSE
Alsaihati et al. [40]	functional network (FN)	flow rate pressure weight on bit drill string speed rate of penetration torque hook load standpipe pressure true vertical depth	3567	surface real-time transmitter sensors	0.99	0.45 RMSE
	Random Forests				0.99	0.42 RMSE
	Support Vector Machines				0.97	0.58 RMSE
Our newly developed model	ANN	Mud Flow rate Mud weight plastic viscosity yield point gel (0/10) total flow area Annular Pressure Standpipe Pressure Rate of Penetration Inclination angle Total Downhole Annular Temperature	4663	real-time downhole sensors	0.9999	RMSE 0.00315

5. Conclusions

An ANN model was developed using MATLAB to predict the ECD based on a large dataset of data points obtained from downhole sensors during drilling operations. Preprocessing operations such as data cleaning, filtering, normalization, and transformation were carried out to prepare the data for training, validation, and testing of the model. Statistical tests were conducted to evaluate the performance of the developed model.

The statistical tests conducted include R^2, RMSE, ARE, AARE, SD, and MSE for the training, validation, testing, and overall datasets. The results of the tests indicate that the developed model does not suffer from neither underfitting nor overfitting. Moreover, the predicted ECD values from the developed model agrees well with the actual ECD values.

To further evaluate the performance of the developed model, the trend of the predicted data and actual data points along the number of the datasets was also analyzed. Additionally, the overall model performance was compared to other models based on input parameters, type of machine learning algorithm, number of data points, and results of statistical tests used to evaluate these models. The developed model showed a higher accuracy in predicting the ECD than other models considered.

In conclusion, the developed ANN model shows a promising solution for the optimization and automation of the drilling process by directly predicting the ECD. This can lead to minimizing hazards and reducing risks associated with drilling operations.

6. Future work recomendation

Future research should investigate the performance of the developed model on datasets from different wells in various regions to confirm its capability to accurately predict ECD under different circumstances. In case of any deficiencies in the model's predictive ability, the model can be updated by including data from multiple wells located in different places to improve the training process and enhance its predictive capability.

Moreover, a parametric study can be conducted to examine the significance of each input element used to feed the model and to determine the effects of each input on the model's ability to predict the true value of ECD. This will help in understanding the relative importance of each input variable and may aid in improving the accuracy of the model predictions.

Overall, these future investigations will further enhance our understanding of the developed model and enable its application to a wider range of drilling operations, leading to an improved optimization and automation of the drilling process with reduced hazards and risks.

Conflict of interest

The authors declare that there is no conflict of interest in publishing this paper.

References

1. Hansen KV (1993) Neural networks for primary reflection identification. *SEG Technical Program Expanded Abstracts*, 242–245. https://doi.org/10.1190/1.1822450
2. Kononov A, Gisolf D, Verschuur E (2007) Application of neural networks to travel-times computation. *2007 SEG Annual Meeting, San Antonio, Texas,* SEG-2007-1785. Available from: https://onepetro.org/SEGAM/proceedings-abstract/SEG07/All-SEG07/SEG-2007-1785/95760.
3. Syed FI, AlShamsi A, Dahaghi AK, et al. (2022) Application of ML & AI to model petrophysical and geomechanical properties of shale reservoirs—A systematic literature review. *Petroleum* 8: 158–166. https://doi.org/10.1016/j.petlm.2020.12.001
4. Arehart RA (1990) Drill-bit diagnosis with neural networks. *SPE Comput Appl* 2: 24–28. https://doi.org/10.2118/19558-PA

5. Moran D, Ibrahim H, Purwanto A, et al. (2010) Sophisticated ROP prediction technologies based on neural network delivers accurate drill time results. *IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition,* SPE-132010-MS, OnePetro. https://doi.org/10.2118/132010-MS
6. Sprunger C, Muther T, Syed FI, et al. (2022) State of the art progress in hydraulic fracture modeling using AI/ML techniques. *Model Earth Syst Environ* 8: 1–13. https://doi.org/10.1007/s40808-021-01111-w
7. Denney D (2000) Artificial neural networks identify restimulation candidates. *J Pet Technol* 52: 44–45. https://doi.org/10.2118/0200-0044-JPT
8. Syed FI, Alshamsi M, Dahaghi AK, et al. (2022) Artificial lift system optimization using machine learning applications. *Petroleum* 8: 219–226. https://doi.org/10.1016/j.petlm.2020.08.003
9. Yang H-S, Kim N-S (1996) Determination of rock properties by accelerated neural network. *2nd North American Rock Mechanics Symposium, Montreal, Quebec, Canada,* ARMA-96-1567, OnePetro. Available from: https://onepetro.org/ARMANARMS/proceedings/NARMS96/All-NARMS96/ARMA-96-1567/130834.
10. Kohli A, Arora P (2014) Application of artificial neural networks for well logs. *European Assoc Geosci Eng*. https://doi.org/10.3997/2214-4609-pdb.395.IPTC-17475-MS
11. Syed FI, Alnaqbi S, Muther T, et al. (2022) Smart shale gas production performance analysis using machine learning applications. *Pet Res* 7: 21–31. https://doi.org/10.1016/j.ptlrs.2021.06.003
12. Syed FI, Muther T, Dahaghi AK, et al. (2021) AI/ML assisted shale gas production performance evaluation. *J Petrol Explor Prod Technol* 11: 3509–3519. https://doi.org/10.1007/s13202-021-01253-8
13. Yehia T, Khattab H, Tantawy M, et al. (2022) Improving the shale gas production data using the angular-based outlier detector machine learning algorithm. *J Univ Shanghai Sci Technol* 24: 152–172. Available from: https://jusst.org/wp-content/uploads/2022/08/Improving-the-Shale-Gas-Production-Data-Using-the.pdf.
14. Yehia T, Khattab H, Tantawy M, et al. (2022) Removing the outlier from the production data for the decline curve analysis of shale gas reservoirs: A comparative study using machine learning. *ACS Omega* 7: 32046–32061. https://doi.org/10.1021/acsomega.2c03238
15. Yehia T, Wahba A, Mostafa S, et al. (2022) Suitability of different machine learning outlier detection algorithms to improve shale gas production data for effective decline curve analysis. *Energies* 15: 8835. https://doi.org/10.3390/en15238835
16. Haciislamoglu M (1994) Practical pressure loss predictions in realistic annular geometries. *SPE Annual Technical Conference and Exhibition, OnePetro.* https://doi.org/10.2118/28304-MS
17. Gamal H, Abdelaal A, Elkatatny S (2021) Machine learning models for equivalent circulating density prediction from drilling data. *ACS Omega* 6: 27430–27442. https://doi.org/10.1021/acsomega.1c04363
18. Arshad U, Jain B, Ramzan M, et al. (2015) Engineered solution to reduce the impact of lost circulation during drilling and cementing in rumaila field, Iraq. *International Petroleum Technology Conference, Doha, Qatar,* OnePetro. https://doi.org/10.2523/IPTC-18245-MS
19. Al-Rubaii MM, Al-Nassar FY, Al-Harbi S (2022) A new real time prediction of equivalent circulation density from drilling surface parameters without using pwd tool. *SPE Symposium: Unconventionals in the Middle East - From Exploration to Development Optimisation, Manama, Bahrain,* SPE-209945-MS, OnePetro. https://doi.org/10.2118/209945-MS

20. Marsh HN (1931) Properties and treatment of rotary mud. *Transac AIME* 92: 234–251. https://doi.org/10.2118/931234-G
21. Zhang H, Sun T, Gao D, et al. (2013) A new method for calculating the equivalent circulating density of drilling fluid in deepwater drilling for oil and gas. *Chem Technol Fuels Oils* 49: 430–438. https://doi.org/10.1007/s10553-013-0466-0
22. Abdelgawad KZ, Elzenary M, Elkatatny S, et al. (2019) New approach to evaluate the equivalent circulating density (ECD) using artificial intelligence techniques. *J Petrol Explor Prod Technol* 9: 1569–1578. https://doi.org/10.1007/s13202-018-0572-y
23. Ataga E, Ogbonna J (2012) Accurate estimation of equivalent circulating density during high pressure high temperature (HPHT) drilling operations. *Nigeria Annual International Conference and Exhibition,* OnePetro. https://doi.org/10.2118/162972-MS
24. Erge O, Vajargah AK, Ozbayoglu ME, et al. (2016) Improved ECD prediction and management in horizontal and extended reach wells with eccentric drillstrings. *IADC/SPE Drilling Conference and Exhibition,* OnePetro. https://doi.org/10.2118/178785-MS
25. Rommetveit R, Ødegård SI, Nordstrand C, et al. (2010) Drilling a challenging HPHT well utilizing an advanced ECD management system with decision support and real time simulations. *IADC/SPE Drilling Conference and Exhibition, New Orleans, Louisiana, USA, February 2010,* SPE-128648-MS. https://doi.org/10.2118/128648-MS
26. Hoberock LL, Thomas DC, Nickens HV (1982) Here's how compressibility and temperature affect bottom-hole mud pressure. *Oil Gas J* 80: 12. Available from: https://www.osti.gov/biblio/5213477.
27. Peters EJ, Chenevert ME, Zhang C (1990) A model for predicting the density of oil-based muds at high pressures and temperatures. *SPE Drilling Eng* 5: 141–148. https://doi.org/10.2118/18036-PA
28. Bybee K (2009) Equivalent-circulating-density fluctuation in extended-reach drilling. *J Pet Technol* 61: 64–67. https://doi.org/10.2118/0209-0064-JPT
29. Hemphill T, Ravi K, Bern PA, et al. (2008) A simplified method for prediction of ECD increase with drillpipe rotation. *SPE Annual Technical Conference and Exhibition,* OnePetro. https://doi.org/10.2118/115378-MS
30. Ahmed R, Enfis M, Miftah-El-Kheir H, et al. (2010) The effect of drillstring rotation on equivalent circulation density: Modeling and analysis of field measurements. *SPE Annual Technical Conference and Exhibition, Florence, Italy,* SPE-135587-MS, OnePetro. https://doi.org/10.2118/135587-MS
31. Costa SS, Stuckenbruck S, Fontoura SA, et al. (2008) Simulation of transient cuttings transportation and ECD in wellbore drilling. *Europec/EAGE Conference and Exhibition,* OnePetro. https://doi.org/10.2118/113893-MS
32. Baranthol C, Alfenore J, Cotterill MD, et al. (1995) Determination of hydrostatic pressure and dynamic ecd by computer models and field measurements on the directional HPHT Well 22130C-13. *SPE/IADC Drilling Conference, Amsterdam, Netherlands,* OnePetro. https://doi.org/10.2118/29430-MS
33. Osman EA, Aggour MA (2003) Determination of drilling mud density change with pressure and temperature made simple and accurate by ANN. *Middle East Oil Show,* OnePetro. https://doi.org/10.2118/81422-MS

34. Harris O, Osisanya S (2005) SPE 97018 evaluation of equialent circulating density of drilling fluids under high-pressure/high-temperature conditions. *SPE Annual Technical Conference and Exhibition, Dallas, Texas,* SPE-97018-MS. https://doi.org/10.2118/97018-MS
35. Elzenary M, Elkatatny S, Abdelgawad KZ, et al. (2018) New technology to evaluate equivalent circulating density while drilling using artificial intelligence. *SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition,* OnePetro. https://doi.org/10.2118/192282-MS
36. Ahmadi MA (2016) Toward reliable model for prediction drilling fluid density at wellbore conditions: A LSSVM model. *Neurocomputing* 211: 143–149. https://doi.org/10.1016/j.neucom.2016.01.106
37. Ahmadi MA, Shadizadeh SR, Shah K, et al. (2018) An accurate model to predict drilling fluid density at wellbore conditions. *Egyptian J Pet* 27: 1–10. https://doi.org/10.1016/j.ejpe.2016.12.002
38. Rahmati AS, Tatar A (2019) Application of Radial Basis Function (RBF) neural networks to estimate oil field drilling fluid density at elevated pressures and temperatures. *Oil Gas Sci Technol—Rev IFP Energies nouvelles* 74: 50. https://doi.org/10.2516/ogst/2019021
39. Alkinani HH, Al-Hameedi AT, Dunn-Norman S, et al. (2019) Data-Driven neural network model to predict equivalent circulation density ECD. *SPE Gas & Oil Technology Showcase and Conference, Dubai, UAE,* OnePetro. https://doi.org/10.2118/198612-MS
40. Alsaihati A, Elkatatny S, Abdulraheem A (2020) Real-time prediction of equivalent circulation density for horizontal wells using intelligent machines. *ACS Omega* 6: 934–942. https://doi.org/10.1021/acsomega.0c05570
41. Saeedi A, Camarda KV, Liang J-T (2007) Using neural networks for candidate selection and well performance prediction in water-shutoff treatments using polymer gels—A field-case study. *SPE Prod Oper* 22: 417–424. https://doi.org/10.2118/101028-PA
42. Vickers NJ (2017) Animal communication: When I’m calling you, will you answer too? *Curr Biol* 27: R713–R715. https://doi.org/10.1016/j.cub.2017.05.064

Appendix

Table S1. Weights and bias for Eq 5.

W1k								
j	K = 1	K = 2	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8
1	–0.3162879	–0.2205999	–0.1534146	–0.8122943	0.27684096	0.05946083	0.7057737	0.71837888
2	0.47074242	0.73567669	–0.1841122	0.1609439	–0.544297	–0.0361249	–0.6537549	0.3031463
3	0.57633503	0.53445329	0.11806691	–0.1717782	0.59016485	0.19424577	0.48176406	–0.0461223
4	–0.0723974	–0.7846182	0.69251662	0.5806464	0.3763547	–0.1862909	–0.6627106	0.34436077
5	0.87507287	0.06187093	–0.8490168	–0.3597439	0.05914917	0.6146153	–0.284121	–0.7702518
6	0.40072408	0.40108283	–0.6232822	0.16766818	–0.7951167	0.10796626	0.08864881	–0.3711264
7	–0.3565583	–0.362438	–0.4915954	0.45567143	–0.1270272	0.71201101	0.16808018	–0.582408
8	–0.4841982	–0.2587741	0.2097799	–0.6868848	–0.2181581	–0.386522	–0.819447	–0.0927483
9	–0.5658372	0.22493358	0.44300055	0.62606126	0.03950263	0.74454768	0.32491909	–0.1755511
10	–0.1525185	–0.1471221	0.15335513	1.66202992	–0.0257545	–0.1961542	–0.7738434	0.98744373
11	0.75635418	–0.1192605	0.89637027	0.02366081	–0.0251318	–0.2480005	–0.6917067	–0.6878273
12	–0.1907664	0.75014404	–0.547985	–0.7208037	–0.4015344	–0.8123173	–0.3796249	–0.6414277
13	0.2753617	–0.2807452	0.05779137	0.53305334	0.38355258	0.43781204	0.47715008	0.25132789
14	0.32609773	0.8282494	0.63751415	–0.3671818	0.43665158	–0.2433486	0.57235532	0.51912847
15	–0.376222	–0.6754491	0.03837612	–1.0586986	–0.7997034	–0.3048866	0.39603667	–0.5164811
16	–0.8784547	0.4117126	–0.0640326	2.0966004	1.57688111	0.09710967	–0.2008817	0.71642072
17	0.62278114	–1.2958257	0.38730381	–0.2957597	–1.1229059	0.54408656	–0.0424252	–0.1548535
18	–0.7949088	–0.4885607	0.763839	0.06843448	–0.1143591	0.45146181	0.51566339	–0.3727209
19	–0.3384047	0.28692669	–0.3886945	–0.1566388	0.20261898	0.28846773	0.45120118	0.53359223
20	–0.3055754	–0.5643252	0.00307304	–0.3453724	0.04010683	–0.2098721	0.05480218	0.99288825
21	0.06691694	–0.1800266	–0.0600536	–0.8364574	–0.4739945	0.30177119	–0.4583955	0.54871822
22	0.04089296	–0.3349337	0.8298507	0.62799788	–0.2536319	0.23980499	–0.7075206	0.28613335
23	–0.4694525	0.39965768	–0.115026	–0.2003004	0.14254142	0.70132383	0.41762119	–0.7554541
24	–0.0109694	–0.1814243	0.7821351	–0.5992469	–0.0454298	0.28113596	–0.1445081	0.50279332
25	–0.5191615	–0.5264875	–0.0983493	0.02037143	–0.771881	0.49064904	0.56669374	0.33009342
26	0.02945113	0.76945086	0.69429051	–0.0484746	0.24434191	0.13584118	0.15775373	–0.297542
27	0.37873374	–0.175608	0.25532596	1.56574329	0.65630888	–0.1292866	0.4695617	–0.9068194
28	0.45529626	0.42168405	–0.1693115	–0.151697	–0.3223661	0.52806949	–0.2442731	0.60403059
29	0.61868226	0.0877681	0.15734532	–0.2107422	0.67465835	–0.0321478	0.62465293	0.10003356
30	0.63842067	0.42063418	0.12350438	–0.9856083	–0.0890697	0.34159343	–0.3233713	0.60799357
31	0.59697932	0.04134443	0.2079904	–0.1645776	0.37573412	0.10818018	0.68644479	–0.1616589
32	–0.4885499	0.56509531	–0.6121336	–0.2376948	–0.2009806	–0.3118941	0.79601477	–0.77041
33	0.66641887	0.73621151	0.30987086	0.76136742	0.32619628	0.03254868	0.57595334	0.31551128
34	0.32259532	–0.4036518	0.83293577	–0.3719892	–0.6168266	–0.5279925	0.18995618	–0.0639327

Table S2. Supplement weights and bias for Eq 5.

W1k									
j	K = 9	K = 10	K = 11	K = 12	K = 13	K = 14	W2j	b1j	b2
1	0.444576	0.228892	0.115168	–0.59008	0.365763	0.306	–0.855302	1.91280078	–0.1362576
2	0.665774	–0.11296	–0.48544	0.734796	0.494432	–0.45338	–0.1072826	–1.8223261	
3	–0.37956	0.337142	–0.37756	–0.39069	0.277306	–0.35669	–0.0122904	–1.9480858	
4	–0.25055	0.446847	–0.25027	0.726612	–0.12881	0.200836	0.00112854	1.48904892	
5	0.488812	–0.06977	0.37861	0.315209	0.235882	–0.05906	0.26012078	–1.5324478	
6	–0.57338	0.377488	0.725849	–0.05323	0.056374	0.241381	–0.4420826	–1.5196056	
7	–1.00097	–0.63218	–0.15221	0.580496	0.182629	0.346311	–0.3132041	1.16872836	
8	–1.08034	–0.31026	0.040105	0.384844	–0.28442	0.02589	–0.3468976	1.05896888	
9	0.07828	0.593716	–0.45376	–0.0032	0.512764	–0.08888	0.27052089	1.00651785	
10	0.188522	–0.5075	–1.80733	–0.34945	0.085652	–0.46312	–1.1990717	1.01097797	
11	0.493513	–0.52638	0.283106	–0.36449	0.135945	–0.061	0.30775044	–0.9157728	
12	0.367344	–0.71684	0.463303	0.180616	–0.23133	–0.38203	0.80772526	–0.5568318	
13	–0.30468	0.329222	–1.16385	–0.44313	0.42754	–0.14355	–0.2288981	–0.5073333	
14	0.393266	0.857027	0.285484	–0.60618	0.046845	–0.11731	0.21328815	–0.2853488	
15	0.008057	–0.19876	2.335414	0.099667	–0.08183	0.287767	1.05414841	0.22572716	
16	0.335907	–0.38603	–2.41665	–0.08655	–0.0741	–0.35272	2.15735625	0.55374166	
17	–0.75494	–0.21653	–0.5382	–0.25327	0.335888	–0.65718	0.93083093	–0.6535756	
18	0.031721	0.341946	0.515567	–0.40836	0.593648	0.188823	0.0232836	0.07812335	
19	–0.37288	–1.19062	0.638179	0.257811	–0.27746	0.543206	–0.6433608	–0.2591042	
20	0.278492	0.413251	1.198806	–0.13147	0.270162	0.631891	1.00662778	0.21204428	
21	–0.70493	–0.34051	–0.40315	–0.36569	0.645062	0.483876	0.20386899	0.29928962	
22	–0.27857	0.820464	0.522807	0.415914	0.003446	–0.67671	–0.0980896	–0.4400881	
23	–0.64655	–0.26815	0.917162	–0.45081	–0.11074	–0.17491	–0.2582537	–0.6009535	
24	0.878195	–0.10783	–0.34463	0.447027	0.015446	–0.70335	0.21573361	–1.1874293	
25	0.132877	–0.02074	–0.6973	0.192938	0.601554	–0.15136	–0.1368108	–0.9032459	
26	–0.06464	0.959172	0.542966	0.551888	–0.44545	–0.21166	0.17658593	0.8048832	
27	0.276286	0.849763	–0.818	0.074856	0.132929	0.478235	–0.9751743	–0.7695462	
28	–1.09264	0.14061	–0.87705	–0.0679	0.1268	0.187739	0.79765868	1.18595977	
29	0.39855	1.128146	–0.20816	–0.08315	0.462093	0.777869	–0.5527538	1.29356814	
30	0.270055	–0.10548	0.760275	–0.85866	0.23381	0.637327	0.69111327	1.28966962	
31	0.167	0.882473	0.797602	0.080592	–0.21153	0.692714	–0.4571694	1.69406201	
32	0.047511	–0.44231	–0.45206	–0.59048	0.038518	0.137825	–0.1452332	–1.5819047	
33	–0.6309	0.201291	–0.32657	–0.6889	–0.13256	0.334145	–0.3605838	1.68747166	
34	–0.51563	0.036465	–0.60704	0.489215	–0.73146	–0.47758	0.094864	1.73944829	

AIMS Press