DEVELOPING A PREDICTIVE MODEL OF CO_2 FLOODING PROJECT

BACHELOR THESIS

D'Aqnan Marusaha Matthew Pandiangan
12215072

Submitted as partial fulfillment of the requirements for the degree of
BACHELOR OF ENGINEERING
in Petroleum Engineering study program

PETROLEUM ENGINEERING STUDY PROGRAM
FACULTY OF MINING AND PETROLEUM ENGINEERING
INSTITUT TEKNOLOGI BANDUNG
2019

DEVELOPING A PREDICTIVE MODEL OF CO_2 FLOODING PROJECT

BACHELOR THESIS

D'AQNAN MARUSAHA MATTHEW PANDIANGAN
12215072

Submitted as partial fulfillment of the requirements for the degree of
BACHELOR OF ENGINEERING
in Petroleum Engineering study program

Approved by:
Thesis Adviser,

…………………………

Zuher Syihab ST, Ph.D.
NIP. 197206061999031003

Developing a Predictive Model of CO_2 Flooding Project

D'Aqnan Marusaha Matthew Pandiangan* and Zuher Syihab**

Abstract

Economic analysis that perceive both technical and economical parameters plays crucial role in identifying the feasibility of EOR application. This step usually be done before further reservoir investigation being conducted. A model that can fulfill the analysis will greatly help the project feasibility study.

This study focused on creating a predictive model to accurately predict the reservoir performance for 10 years CO_2 flooding project. The model itself constructed by 25 parameters, which affecting both technically and economically in which its value distributed in three ways, discrete real, continuous real, and using a formula and trained in experiments using CMG-CMOST. Experiments that passed the data quality control through several constraints then used as model training and verification data. Net Present Value (NPV) is then used as project's economic objective for the predictive model as it represents the viability of EOR application.

Several methods, both regression and neural network was done to predict the chosen objective function, NPV. 6089 experiments generated by CMG-CMOST used as the proxy material to generate the model. Mainly considering the proxy cumulative error and error distribution, the study showed that multilayer artificial neural network with 20-9-6-1 structured neurons gave the best fitted model, where fitted more than 97% with training validated with verification data, followed by CMG-CMOST generated regression, CMG-CMOST generated one layered radial basis function neural network, and self-approached regression. The predictive model that chosen was expected to generate the project's NPV with confidence level around 80% based on P50 value of proxy verification data.

The novelty of this predictive model helps user to determine, whether to continue the project or not directly to the project's predicted NPV.

Keywords: predictive model, CO_2 flooding, regression, neural network.

Sari

Analisis ekonomi yang memperhitungkan parameter teknis dan ekonomis memainkan peran penting dalam mengidentifikasi kelayakan aplikasi EOR. Langkah ini biasanya dilakukan sebelum penyelidikan reservoir lebih lanjut dilakukan. Model yang dapat memenuhi analisis akan sangat membantu studi kelayakan proyek.

Studi ini berfokus pada pembuatan model prediksi untuk memprediksi kinerja reservoir secara akurat selama 10 tahun proyek CO_2 flooding. Model ini sendiri dibangun oleh 25 parameter, yang mempengaruhi baik secara teknis maupun ekonomi dimana nilainya didistribusikan dalam tiga cara, diskrit, kontinu, dan menggunakan formula serta dilatih dalam percobaan menggunakan CMG-CMOST. Eksperimen yang melewati kontrol kualitas data melalui beberapa batasan kemudian digunakan sebagai model pelatihan dan data verifikasi. Net Present Value (NPV) digunakan sebagai tujuan model prediktif dimana mewakili kelayakan aplikasi EOR.

Beberapa metode, baik regresi dan jaringan saraf dilakukan untuk memprediksi fungsi tujuan yang dipilih, NPV. 6089 percobaan yang dihasilkan oleh CMG-CMOST digunakan sebagai bahan proxy untuk menghasilkan model. Mempertimbangkan error kumulatif proxy dan distribusi error, penelitian menunjukkan bahwa multilayer neural network dengan struktur neuron 20-9-6-1 menghasilkan model terbaik, di mana melebihi 97% dengan data pelatihan dan divalidasi dengan data verifikasi, diikuti oleh CMG-CMOST yang menghasilkan model dengan pendekatan regresi, radial basis function neural network yang dibentuk oleh CMG-CMOST, dan regresi bentukan penulis. Model prediktif yang dipilih diharapkan untuk menghasilkan NPV proyek dengan tingkat kepercayaan sekitar 80% berdasarkan nilai P50 dari data verifikasi proxy.

Inovasi dari adanya model prediksi ini membantu pengguna untuk menentukan apakah akan melanjutkan proyek atau tidak, langsung melalui nilai prediksi NPV sebagai parameter kelayakan proyek.

Kata kunci: Predictive model, injeksi CO_2, regresi, neural network

*) Student of Petroleum Engineering Study Program, Institut Teknologi Bandung, 2015 batch
**) Thesis Advisor in Petroleum Engineering Study Program, Institut Teknologi Bandung

1. Introduction

Continuous decline of oil production that occurs in Indonesia is being one of the main concerns for both oil company and government. Enhanced Oil Recovery (EOR) is one of the solutions for this problem, other than exploration enhancement, well workover, production optimization, and any other perceptible method. For a decade, EOR has shown an increase of oil recovery when applied after the field has passed its peak performance, which helps for prolong the field productivity. Commonly, EOR was done using chemical agent which react to the reservoir fluid and changes its characteristics.

One of common EOR method was CO_2 injection, which shows a positive trend in field oil production. This gas mainly produced from industrial activities not to mention petroleum industry through its daily activities such as flaring as excess gas treatment and oil refineries. CO_2 injection helps reduce the gas house effect which resulted from the increase of CO_2 in the atmosphere.

Many of the CO_2 injection projects that applied in mature field have been highly successful in recovering remaining oil. However, there is no guarantee that CO2 injection project will economically feasible to be applied in some cases. The first steps in developing a depletion plan for a reservoir are to identify the primary factors that will have the greatest impact on the CO_2 flood, which to carry out an assessment of those factors' probable impact on the project's technical and economic success, and to use those analysis result to decide the future of the project.

At early stage of CO_2 project design, scaling techniques are a better means of forecasting CO_2 flood performance than are reservoir simulators. To yield meaningful predictions, reservoir simulators require an accurate characterization of both the reservoir, its operation, and characterizing a reservoir accurately can be a time-consuming and somewhat expensive process. If a scoping analysis shows that a CO_2 flood might be economical and if significant financial risk is involved, then using a reservoir simulator to forecast performance would be a useful and economically justified to proceed to the next step. A predictive model that can present the reservoir performance with high level of accuracy can be used as this method of assessment before characterizing a reservoir accurately, which made from various combination of parameter, both from technical and economic for economic prediction that mainly affects the reservoir performance.

This study objectives are to build, choose the best model, and optimize the predictive model for CO_2 flooding under five-spots injection pattern for 10 years injection based on the minimum R squared error of NPV.

2. Basic Theory

2.1 CO_2 Flooding

CO_2 flooding is an EOR method where carbon dioxide continuously injected into the reservoir increase the recovered oil. CO_2 will react with oil, water, gas and reservoir rock when injected. Due to CO_2 characteristic that soluble in oil and reacts less to water, it contributes to the large contribution of CO_2 as the injector fluid in the EOR method.

Understanding basics of physical interactions and chemical reactions between CO_2 and the reservoir is important for the flooding project, alongside the project's prospect. CO_2 solubility in oil is the major parameter due to its effect in oil viscosity reduction and increases the oil swelling, sequentially, enhance the oil relative permeability and oil mobility, which resulting in the increase the oil recovery (Emera & Sarma, 2006).

There are some reservoir aspects that help to know what problems might happen and cause technical failure in CO_2 flooding (Jarrel, 2003):

- Average reservoir pressure and thermodynamic MMP
  The minimum miscibility pressure (MMP) is the lowest pressure for which a gas can develop miscibility through a multi-contact process with a given reservoir oil at reservoir temperature (Elsharkawy, Poettmann, & Christiansen, 1992). When the contact between CO_2 and oil occurs, the pressure at which little or no reservoir mixing the pressure switch miscibility happens is defined as the thermodynamic minimum miscibility. The problem in CO_2 flooding is how much the reservoir is above the thermodynamic MMP, where too difficult to be determined without a detailed reservoir simulation. Hence, most cases, the average reservoir pressure is used as an estimate of probable success.

- Well pattern
  Pattern itself affects the flooding performance, because in CO_2 flooding. Higher injection-production wells ratio increases its productivity, hence seen that five-spot pattern gave better flooding result. This pattern helps to maintain reservoir average pressure, indirectly indicates whether the flooding will reach its MMP or not. Frequently, CO_2 flood was done after waterflood. When compared, the problems caused from low sweep efficiency and low reservoir pressure in a waterflood will worsen during CO_2 flooding,

- Residual oil saturation
  Low residual oil saturation indirectly tells that there may not enough reserves left for company to gain their project's profit, hence the project will be counted as economically unfeasible.

- Reservoir wettability
  Formation wettability ought to be known before the project of CO_2 flooding being held. For strongly water wet rock, the best mode of CO_2 flooding is continuous injection, rather than using WAG, where have been used successfully in strongly water-wet reservoir. For oil-wet reservoirs, WAG and CO_2 flood is feasible to be done.

- Reservoir heterogeneities
  The degree of reservoir vertical heterogeneity affects CO_2 flooding performance in reservoirs where layer effects are more important than gravity effect. High permeability channels, faults and fractures can cause sweep problems which implies to the reservoir performances.

- Injection well conformance
  Conformance is a measure of where injected fluids are entering the pay zone. Ideally, injected fluids enter the formation only at pay zones and spread out evenly across these zones to avoid early breakthrough.

- Ability to inject and produce fluids at economical rates
  To yield economical fluid injection and production rates, the reservoir must have enough formation flow capacity. If not, the well must be stimulated enough to achieve those rates.

- Gravity effects
  Gravity affects CO_2 flood in two ways, through significant reservoir dip and by fluid crossflow between reservoir.

2.2 CO_2 Flooding Technical Aspects

To create a representative predictive model as a tool for determining the feasibility of CO_2 flooding, variation of representative parameters and its result need to be prepared, where being made from synthetic model in this study. Parameters that used as a guide to determine the feasibility of CO_2 flooding are listed below (Aladasani & Bai, 2010):

a. Oil API Gravity
b. Oil viscosity
c. Oil composition
d. Reservoir pressure
e. Reservoir temperature
f. Heterogeneity
g. Oil saturation
h. Absolute permeability
i. Relative permeability
j. Wetting conditions
k. Source of CO_2
l. Injection fluid composition
m. Surface facility capacity

Detailed information related to those parameters applied in this study shown in Table 1.

2.3 CO_2 Flooding Economical Aspects

Net present value will be used as an indicator to determine the feasibility of CO_2 flooding by its economic value. An economic model, scoping the project expected income and expenditure is used to create the function of net present value.

a. Estimated income
   Estimated income is approached from produced oil by its price.

b. Estimated expenditure
   Cost model will be split into two parts, capital expenditure and operational expenditure. The assumption of no inflation happened on dollars is used in this study. Capital expenditure itself covers site characterization cost, cost for one additional well that drilled as injector well, pattern cost, equipment cost, workover for existing well cost, cost for addition of flow line to injector well, and surface facilities completed with injection facilities costs, which complete cost model for each part can be seen in appendix A.

$$CAPEX = 1{,}752{,}433.13 + 371132 * e^{2.4384*10^{-4}*Depth\,(ft)} + 25 * Depth(ft) + 1.2 * Qinj(SCFD) \quad \text{……………………} (1)$$

   Operational expenditure itself covers from project operation and maintenance cost, CO_2 compression cost, CO_2 recycling cost, lifting cost, CO_2 cost, and monitoring and verification cost, which detailed cost model attached in appendix B.

$$OPEX = 27{,}891 + 209.9 * Q_{inj}(MSCF)^{0.5} + 15.25 * Depth\ (ft) + 0.25 * N_p + 5.057 * Q_{inj}(MSCF) \quad \text{……………} (2)$$

2.4 Regression

Based on Sinha, 2013, regression is the method of estimating a relationship from the given data to depict the nature of data set. The relationship constructed from regression analysis can be used for several computations. One of them is to predict the values from several related variables when there is a relation among them.
Basically, regression analysis composed from several stages:

1. Both dependent and independent variables identification.
2. Form of relationship identification. Usually, when it comes to construct the predictive model, linear, exponential and polynomial relationship being used, in scatter diagram between dependent and independent variables.
3. Computation for analysis.

4. Error analysis. This stage ought to be done to know the model performance with actual data set, based on its fitness.

To do the regression analysis, several methods can be used. Linear Regression approaches the regression from the linear relationship among the variables. When the relationship made from one response variable to its regressor, this regression called as simple linear regression. When its relationship constructed from multiple variable to a response variable, it is called multivariate linear regression. The form of linear regression described in this equation:
$y = A_0 + A_1R_1 + A_2R_2 + A_3R_3 + A_4R_4 + A_5R_5$ .. (3)

When there is a quadratic relationship between variables, quadratic regression can be used. In this regression, the relationship between variables is modelled as the n^{th} order polynomial equation. For easier computation, multiple linear regression model can be used to compute the polynomial regression, which modelled below for simple quadratic and quadratic regression.

$$y = A_0 + A_1R_1 + A_2R_2 + A_3R_3 + A_4R_4 + A_5R_5 + B_1R_1^2 + B_2R_2^2 + B_3R_3^2 + B_4R_4^2 + B_5R_5^2 \quad \text{……... (4)}$$

$$y = A_0 + A_1R_1 + A_2R_2 + A_3R_3 + A_4R_4 + A_5R_5 + B_1R_1^2 + B_2R_2^2 + B_3R_3^2 + B_4R_4^2 + B_5R_5^2 + C_1R_1R_2 + C_2R_1R_3 + C_3R_1R_4 + C_4R_1R_5 + C_5R_2R_3 + C_6R_2R_4 + C_7R_2R_5 + C_8R_3R_4 + C_9R_3R_5 + C_{10}R_4R_5 \quad \text{………………………..…… (5)}$$

Regression mainly determined from its error to find its best fit curve, LSV (Least Square Value) and LAV (Least Absolute Value). This study will refer to LSV where the sum of the diversions being squared is taken to be minimum will be used, referred as sum of square of error. When being compared with neural network, regression can’t model a non-linear relation and has high sensitivity to outlier, but its results can be easily interpreted and being understood by user.

2.5 Neural Network

Neural network is a computer system modeled on the human brain and nervous system. Neural network usually gave better model when compared to regression, due to the complexity of its architecture, made from nodes and weights. The architecture consists of nonlinear information that constructed in several layers parallelly, referred as topology. These nonlinear information as the element of network processing called as neural network neurons or nodes, and the interlink between them called as synapse or weights.

To construct a good neural network, a learning algorithm is used to train the neural network, enhancing their performance to utmost ability. Some examples of neural network learning algorithm are Back Propagation, Quasy Newton, Conjugate Gradient, Levenberg Marquardt, etc.

Counted as one of deep learning method, neural network work by summing the inputs between nodes (s_1, s_2, s_3) and scaled by their weight factors (w_1, w_2, w_3) with a processing function that activated from their sum with a chosen activation function (a_f). There is also a radial basis function (RBF) neural networks which typically have three layers: an input layer, a hidden layer with a non-linear RBF activation function and a linear output layer. This neural network is featured in CMG-CMOST, alongside common neural network.

Artificial neural network works well for nonlinear model and its flexibility to determine the results counted as a classification or regression. The problem is its black box, the neurons and its weight, is hard to be understand and the high number of datasets needed to make a good model. This method is trained to fulfill the chosen goal by develop the network by firing the inputs to its desired output. The output then propagated the results and can be trained backward.

There are limitations in this system. The biggest limitation is that the network is only perform outstandingly in its range of data served, where its knowledge can’t be integrated. The noise, or incorrect data that leading to a reduction of network performance hardly can be identified, hence need further study. The network itself will yield a better result when the data are complete and represent the natural behavior that want to be modelled. Incomplete training sets often to lead to indecisive outputs (Pagel, 2017).

3. Methodology

This study is divided in three main steps, reservoir parameter variation, simulation for generated cases, validation and predictive model making and validation. A workflow that briefly explain the study is attached by the end of this paper in figure 1.

3.1 Reservoir Parameter Variation

Reservoir characteristic that being chosen as the affecting parameter in this study listed in table 1. Parameter that being chosen are the main parameter that commonly affects the CO_2 flooding based on Aladasani & Bai (2010).

CMG simulator was used for this study, where Winprop used to create variation of reservoir fluids, Builder-GEM used to model the reservoir, and CMOST used to make the variation of parameter and its value. Javascript programming language was used to help generate needed parameter, such as relative permeability table, capital expenditure, operational expenditure, etc. Value and relation between each parameter were based on reference attached in the same table. To realistically distributes the variation, certain constraints was applied when cases was generated.

a. Initial condition
   The value of initial oil saturation when CO_2 flooding was conducted should be less than the value of residual oil saturation.
b. Depth and reservoir fluid API relationship
   Based on figure 2, this study will ignore cases that have oil density exceeds 35°API in depth less than 6000 ft.
c. Depth and porosity relationship
   For this study, based on figure 3, cases will be ignored when the porosity less than 0.2 in reservoir depth less than 5000 ft, and when porosity exceeds 0.25 in reservoir depth larger than 7500 ft.
d. Porosity and permeability relationship
   Still approached from figure 3, this study will ignore generated cases if porosity was less than 0.15 with absolute permeability exceeds 200 mD and 10mD when porosity value was bigger than 0.3.

3.2 Sensitivity Study

Cases used as predictive model material will be generated based on parameters with random value in range of its possible value and constraints that were explained in the previous section using CMG-CMOST. Results that being observed from the cases are cumulative oil production (N_p) as its technical aspect and net present value (NPV).

3.3 Predictive Model Generation and Validation

Quality control is a must to create a good predictive model. Data collected from the simulation will be analyzed by its possible value in nature. Below listed data quality control conducted to verify the results:

- Parameter value that being discretized are commonly observed in real condition.
- Initial reservoir pressure already reduced by a multiplier (1-DPres) that presents a model which already passed its peak performance.
- Bottom hole pressure constraints (assumed to be two third of reservoir pressure) and rate of production was used in the simulation for production well, and injection rate for injector well.
- Production rate profile should decline by time, because no well constraints given that able to enhance the production in the middle of injection.
- Outlier data that considerably impossible to happen in field scale will be deleted to increase the predictive model accuracy. When oil cumulative production reaches 10000 bbl only (2 barrel/days), the results counted as outlier.

Predictive model will be generated by using multiple methods, in goals to reach the maximum value of its accuracy that being analyzed from its statistic results and tested with available CO_2 flooding results as an assessment of its prediction ability.

4. Case Study

In this study, a one layered, quarter five-spot pattern model was used to predict the reservoir performance, visualized by figure 4. The grid itself varying by its pattern size, which represents 900 ft^2 (30 ft x 30 ft) for each grid. All properties are distributed in a constant value on each grid (for porosity, permeability, pressure, water saturation, and reservoir fluid composition). Reservoir fluid was made by combination of light, medium and heavy component that regressed to desired API, where oil viscosity, density and other reservoir fluid properties approached from general correlations where their phase envelopes can be seen in figure 5.

Performance that predicted was generated only for a pattern, which represented from the bounded reservoir made from the model. Water-oil and gas-oil relative permeability table are being varied and its relative permeability table are approached from Corey's relative permeability model with end points that being assumed to be the same each other were listed below:

- Connate water saturation (S_{wc}) assumed to be equal with critical water saturation (S_{wcrit})
- Irreducible oil saturation (S_{oirw}) assumed to be equal with residual oil saturation (S_{orw})
- Residual oil-gas saturation (S_{org}) assumed to be equal with irreducible oil-gas saturation (S_{oirg})
- Critical gas saturation (S_{gcrit}) assumed to be equal with connate gas saturation (S_{gcon})

Initial reservoir water saturation and pressure are varied, with working bottom hole pressure approached by 2/3 of reservoir pressure. Performance was predicted for 10 years production, where CO_2 will be injected continuously for 10 years long with cost model constructed for a pattern investment and expenses. The result that generated by the predictive model might be categorized as overestimate value, due to homogenous distribution of reservoir volumetric parameter with constant oil price and production expenses (in function of depth, produced oil and injected CO_2) which rarely found in real cases.

The economic approach in this study is written in appendix A and B, where water treatment being neglected in this study. Cashflow balance was made in yearly cashflow, where 10% discount rate applied to calculate the project NPV. Project's income and expenditure was explained before in 2.3. No fiscal system applied, hence the results of NPV itself turns as project's NPV.

5. Result and Discussion

25 parameters in which its value distributed in three ways, discrete real, continuous real, and using a formula are being trained in 6089 experiments using

CMG-CMOST. 2940 experiments that passed the constraints and quality checked were used as predictive model material, where 2663 of them used for generates the proxy and 277 experiments used for verification. Outliers were already ignored, for both oil cumulative production and NPV, which shown in figure 6. Predictive model was made using various of methods, listed below.

a. Regression

- Self-approached regression

  Linear and generalized linear models are useful in a wide variety of application (McCullargh and Nelder, 1989), and it is hoped that there will be a simple approach that can relate all the parameter and accurately approach the NPV value. Since there's no proper rule to choose a parameter that ought to be regressed with the objective function, author approached the performance in 5 main aspects that tested to have trend directly with NPV.

  1. Oil in Place

     Oil in place approached by reservoir volumetric in place which calculated by the water saturation by the start of CO_2 flood.

     $$R_1 = V = 43560\,\frac{Ah\emptyset(1-S_w)}{B_o} \quad\text{.....................(6)}$$

  2. Oil Rate

     Oil rate is approached from Darcy's radial flow equation, stated below. Trial and error done in this study and proxy showed better results when R_2 approached in function of $\ln(q_o)$.

     $$R_2 = \ln(q_o) = \ln\left(\frac{kk_{ro}h}{\mu_o}\frac{P_{res}-BHP}{ln\left(\frac{r_e}{r_w}\right)}\right) \quad\text{......(7)}$$

  3. Depletion time

     Depletion time is approached from the value of oil in place divided by Darcy's radial flow equation, which shown by equation below. Trial and error done in this study and proxy showed better results when R_3 approached in function of log(t).

     $$R_3 = \log(t) = \log\left(\frac{43560\,\frac{Ah\emptyset(1-S_w)}{B_o}}{7.08\frac{kh}{\mu}\frac{P_{res}-P_{BHP}}{ln\left(\frac{r_e}{r_w}\right)}}\right) \quad\text{....(8)}$$

  4. Amount of Injected CO_2

     This parameter is approached from the ratio of reservoir bulk volume with daily injected CO_2.

     $$R_4 = V_{CO_2} = \frac{43560\,Ah\emptyset}{Q_{inj}(MMSCFD)} \quad\text{.....................(9)}$$

  5. Profit

     Profit is generated by amount of income that has been reduced by capital and operational expenditure, which calculated based on volumetric amount of hydrocarbon in reservoir in initial condition. Trial and error done in this study and proxy showed better results when R_5 approached in function of $P/10^6$.

     $$R_5 = \frac{Profit}{1000000} = (Oil\ Price * R_1 - CAPEX - OPEX)/10^6 \quad\text{...(10)}$$

  Those 5 parameters counted as R_1, R_2, R_3, R_4,and R_5 and used for regression using equation 3, 4 and 5 explained before and generate the predicted NPV in million USD. The result which shown in figure 7 and listed in table 2 shows that this method gave mediocre results statistically when compared to other methods that being used, even though the highest degree of regression already being used. Not to remember that this result neglects most of parameter that construct gas and water relative permeability table which might impacts the value of NPV. Verification have been done for this method, and yet the results still gave lower results than the other methods.

- CMG-generated regression

  CMG-CMOST itself generates a regression in function of linear, simple quadratic and quadratic approach. CMG-CMOST also reduces amount of experiments by itself that insignificantly affects the proxy that being made. All parameters that being made and distributed in function of discrete real and continuous real contributes in the proxy and being illustrated in its Sobol Analysis to know the magnitude of impact for each parameter to the objective function, NPV.

  Based on the results shown in figure 8 and listed in table 2, which function is in quadratic regression due its accuracy is better than linear and simple quadratic, this model gave better result than the 1st method statistically, considering it covers all parameter and giving a relatively more accurate result. A consideration about this proxy model is its complexity due to a large amount of parameter (21 parameters) in which there's no sort of rule that theoretically allows a parameter to be reduced due to its insignificant effect to the objective functions. Verification had been done for this method and its accuracy counted as one of the best methods after artificial neural network, which results can be seen at table 3.

b. Neural Network

   CMG-CMOST also capable to generate a proxy with neural network mode, both in radial basis function and artificial modified architecture.

- Radial Basis Function (RBF) Neural Network

  CMG-CMOST generates RBF neural network in a hidden layer. Based on the results that showed in figure 9 and table 2, RBF Neural Network gave an extraordinary fit result, with less than 10^{-6}

relative error for all trained experiments. It is a little bit leery for a proxy model, since most of RBF processes are unsupervised which increases the possibility of overfitted neural network. The verification using verification data showed that this method accuracy is lower than CMG generated quadratic regression, listed in table 3.

- Multilayer Artificial Neural Network
  To exert neural network ability, a supervised learning for making a proxy model was done using Alyuda NeuroIntelligence. Remembering its ability to only perform by limit of 20 parameters and 2000 cases, reduction was done with combining area (A) and thickness (h) as a new parameter called volume. DPres, that used to define reservoir pressure also changed directly to Pres as its parameter. 2940 experiments done selected randomly to fulfill the application limit. Data partition was done randomly resulting in 1681 records being used as training and validation set (84.05%), and another 319 records used for testing set (15.95%). All parameters categorized as numeric data.

  A sensitivity test using back propagation training algorithm was done to find the optimum number of neurons and layers. The common value of hidden neuron proposed by Jha, 2009 was followed, where it should be between 1 and number of parameters. It is discovered that the optimum one layered neural network performance reached by using 9 neurons, showed in figure 10.

  Based on Alvarez, 2006, 2 hidden layers are commonly used and usually when it comes more than 2 layers of neural network, it does not affect the result. By that, another sensitivity was done to choose the best fit number of neurons for 2nd layer of neural network. Based on figure 11, most of neurons studied overfit based on its absolute error value, but not on 6 neurons which gave the best results. A neural network with 20-9-6-1 architecture chosen as the best neural network design for proxy model with R square value of 0.9272 before optimization.

To choose the best proxy to estimate the value of project's NPV, basic statistics results related to estimated value by the proxy was used. Mainly weighting on proxy fitness and error distribution as the decision parameter, tabulated in table 2 and 3, multilayer artificial neural network chosen as the predictive model to be used as this study results.

Figure 12 show the neural network that already been optimized by trained in 40,000 iterations to meet its optimum performance with Back Propagation learning algorithm to find its optimum performance.

6. Conclusion

Based from the study, it can be concluded that multilayer artificial neural network gave the best accuracy to predict the net present value of a bounded 5-pattern CO_2 flooding continuously for 10 years. This result can be used as a screening method for CO_2 flooding economic viability in terms for 1 pattern, with accuracy higher than 97% based from its proxy data average error, with mainly 50% of verified data approached with value of error less than 20%.

The value of larger number of patterns can be approached by multiplying the result with its number of patterns. An adjustment, especially in the project capital expenditure, if not being cost modeled, can be approached by reducing the value of predicted NPV with additional expenses. All generated proxy model made from this study attached on appendix C.

7. Recommendation

To construct better proxy model for CO_2 flood economic screening, a further study related to parameter reductions and combination for simpler proxy model is needed. It is hoped that a facile, yet accurate model can be constructed using easier method and can be trained easily. Clustering the data based on a range of parameter might help in better prediction.

Reconstructing better artificial neural network can be done with further adjustment in each weighted neuron can help it reaches higher accuracy to predict the project's NPV. Changing its its activation function of each nodes (step function, sigmoid function) also might help. Developing neural network using other software may also help for modifying its architecture, nodes, weight and other neural network component to its best performance.

Finally, for further development, another proxy model to approach CO_2 flood performance by its economic value indicators other than NPV also can be constructed, like internal rate of return and payout time. Reconstruction of other predictive model to approach NPV in different project life time is highly recommended. When it comes to the cost model used in this study, reduction of cost assumption (e.g. fluctuation of oil price) can help the proxy to approach better project's NPV. It is recommended to develop another predictive model for other EOR method.

8. Acknowledgement

The author would like to express his gratitude to several people listed below for:

1. The author's beloved parents, sister and brother for endless love, support and prayer.
2. Zuher Syihab, S.T., Ph.D., as the thesis advisor for providing invaluable guidance, comments and suggestions constantly motivating author to work harder to reach the best results.

3. Muhammad Bhakti Setiawan for his knowledge for the author to learn CMG utilization to its utmost capability.
4. Lecturers, officers and employees in Petroleum Engineering Study Program ITB, for their knowledge, support, and facilities the author received during his study in ITB.
5. Computer Modelling Group Ltd. that provides CMG educational licenses for this study.
6. Author’s friends and relations for all support, motivation and joy to the author during his study in Petroleum Engineering ITB.

9. References

Advanced Resources. (2006). *Cost Model for CO_2-Based Enhanced Oil Recovery (CO_2-EOR)*. DOE/FE-HQ.

Aladasani, A., B. Bai. (2010). *Recent Developments and Updated Screening Criteria of Enhanced Oil Recovery Techniques*. SPE International Oil and Gas Conference and Exhibition, China, 8-10 June, Beijing, 2010.

Alvarez, E. (2006). *Artificial Neural Networks*. Mexico: Instituto Tecnologico de Toluca.

Danish Energy Agency. (2017). *Vietnam Energy Outlook Report 2017*.

Ehrenberg, S.N. and Nadeau, P.H. *Sandstone vs. Carbonate Petroleum Reservoir: A Global Perspective on Porosity-Depth and Porosity-Permeability Relationships*. AAPG Bulletin, v. 89, no. 4 (April 2005), pp. 435–445.

Elsharkawy, A. M., Poettmann, F. H., & Christiansen, R. L. (1992). *Measuring Minimum Miscibility Pressure: Slim-Tube or Rising-Bubble Method?* SPE/DOE Enhanced Oil Recovery Symposium. Tulsa: Society of Petroleum Engineers.

Emera, M. K., & Sarma, H. K. (2006). *A Genetic Algorithm-Based Model to Predict CO2-oil Physical Properties for Dead and Live Oil*. Petroleum Society's 7th Canadian International Petroleum Conference (57th Annual Technical Meeting). Calgary: Petroleum Society.

Jarrell, P.M. (2002). *SPE Monograph Vol. 22: Practical Aspects of CO_2 Flooding*. Texas: SPE.

Jha, G. 2007. *Artificial Neural Networks and Its Applications*. Journal of Indian Agricultural Research Institute, vol. 110, no. 012, pp. 41–49.

Kantzas, A. (2001). *Residual Gas Saturation Revisited*. SPE Reservoir Evaluation & Engineering, September 2001, pp. 467-476

Kennaird, T. (1988). *Residual Oil Saturations Determined by Core Analysis*. OSEA-88175: 7th Offshore South East Asia Conference. Singapore.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models, 2nd ed*. London: Chapman & Hall.

Pagel, J.F. and Kirshtein, P. (2017). *Machine Dreaming and Consciousness*. Amsterdam : Elsevier Science.

Rangel, A. et al. (2017). *Geochemical Assessment of the Colombian Oils Based on Bulk Petroleum Properties and Biomarker Parameters*. Marine and Petroleum Geology 86 pp. 1291-1309, September 2017.

Saad, N., Cullick, A.S., Honarpour, M.M. (1995). *Effective Relative Permeability in Scale-Up and Simulation*. SPE Rocky Mountain Regional: Low-Permeability Reservoirs Symposium, 20-22 March 1995.

Sinha, P. (2013). *Multivariate Polynomial Regression in Data Mining: Methodology, Problems and Solutions*. International Journal of Scientific & Engineering Research Vol. 4. India.

VITTRANS. (2000). *Technical Report No.3 Transport Cost and Pricing in Vietnam*. Almec Corp: Vietnam.

Wei, Ning. (2015). *Economic Evaluation on CO_2 EOR of Onshore Oil Fields in China*. International Journal of Greenhouse Gas Control 37, 2015 pp. 170-181.

10. Nomenclature

μ = viscosity (cp)
μ_o = oil viscosity (cp)
A = area (acre)
A_n = constant
a_f = activation function
ANN = Artificial Neural Network
BHP = bottom hole pressure (psi)
B_n = constant
B_o = oil formation volume factor (rb/STB)
$CAPEX$ = Capital Expenditure
C_n = constant
h = thickness (ft)
k = absolute permeability (mD)
k_{ro} = oil relative permeability
MMP = Minimum Miscible Pressure
NN = Neural Network
N_p = cumulative oil produced (STB)
NPV = Net Present Value
$OPEX$ = Operational Expenditure
P_{res} = reservoir pressure (psi)
Q_{inj} = injection rate
q_o = oil rate (STB/d)
RBF = Radial Basis Function
r_e = reservoir radius (ft)
R_n = regressor
r_w = well radius (ft)
S_{gcon} = connate gas saturation
S_{gcrit} = critical gas saturation
s_n = summing inputs
S_{oirg} = irreducible oil-gas saturation
S_{oirw} = irreducible oil saturation
S_{org} = residual gas saturation
S_{orw} = residual oil saturation

S_w = water saturation
S_{wc} = connate water saturation
S_{wcrit} = critical water saturation
t = time
V = volume (ft^3)
V_{CO2} = volumetric injected CO_2
WAG = Water Alternating Gas
w_n = weight factors
ρ_o = oil density (lb/ft^3)
ϕ = porosity

List of Figures

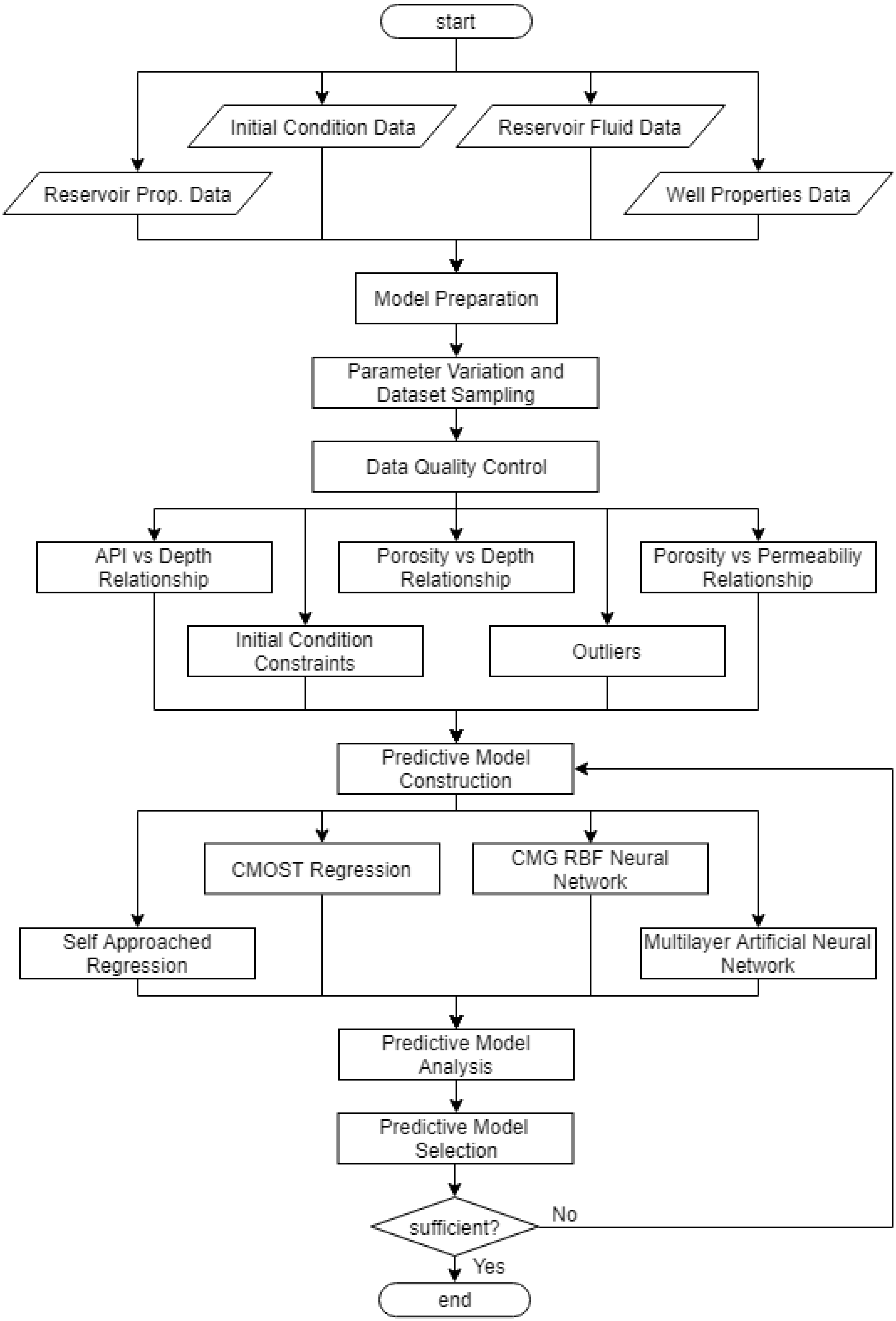

Figure 1. Study Workflow

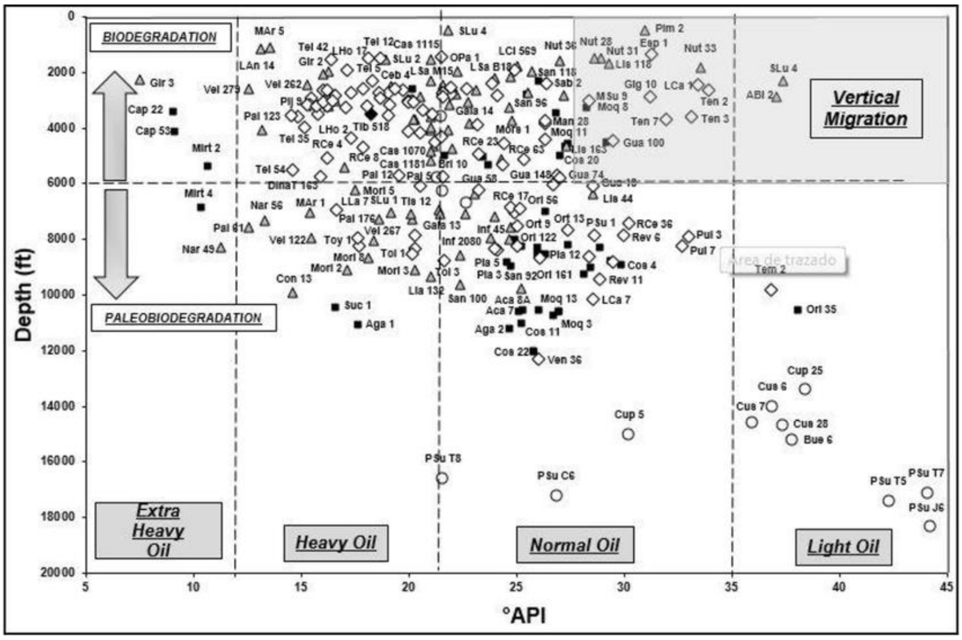

Figure 2. Oil API vs. Depth Distribution (Rangel, 2017)

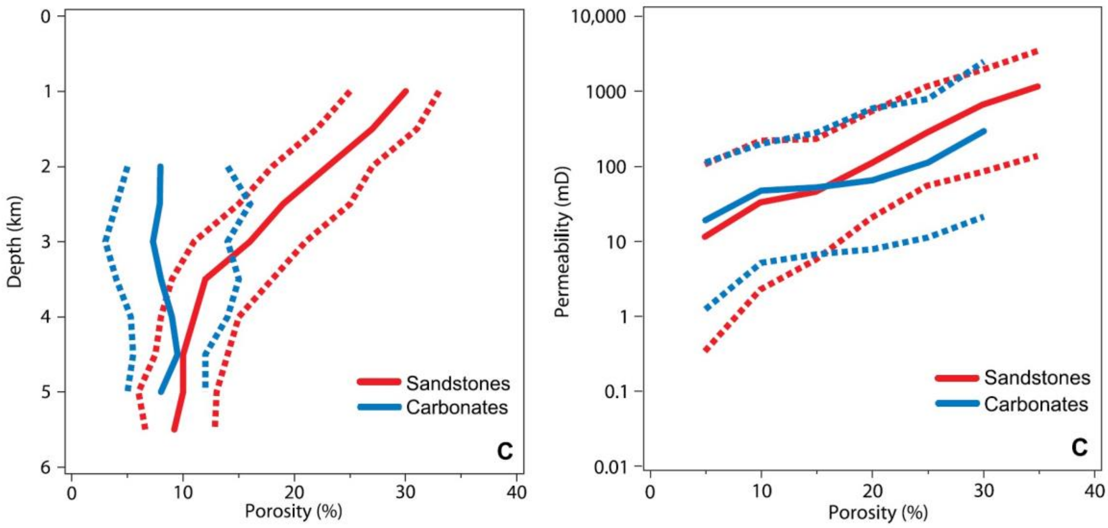

Figure 3. Porosity, Permeability, and Depth Relationships (Ehrenberg, 2005)

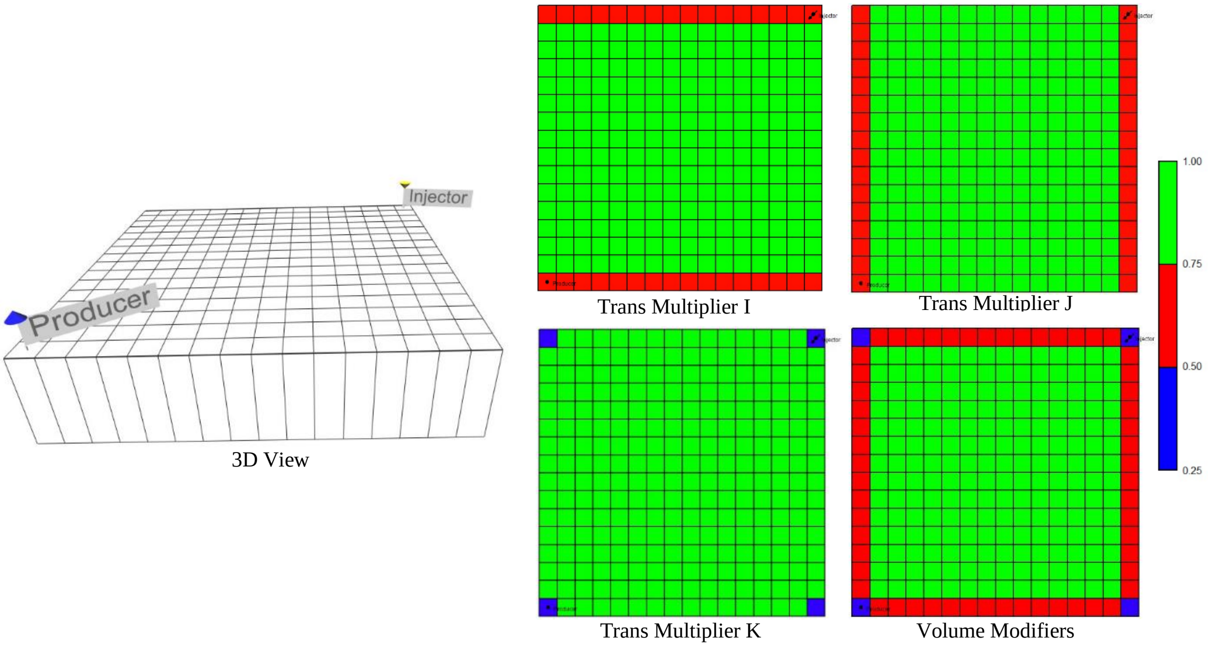

Figure 4. ¼ five-pattern Model

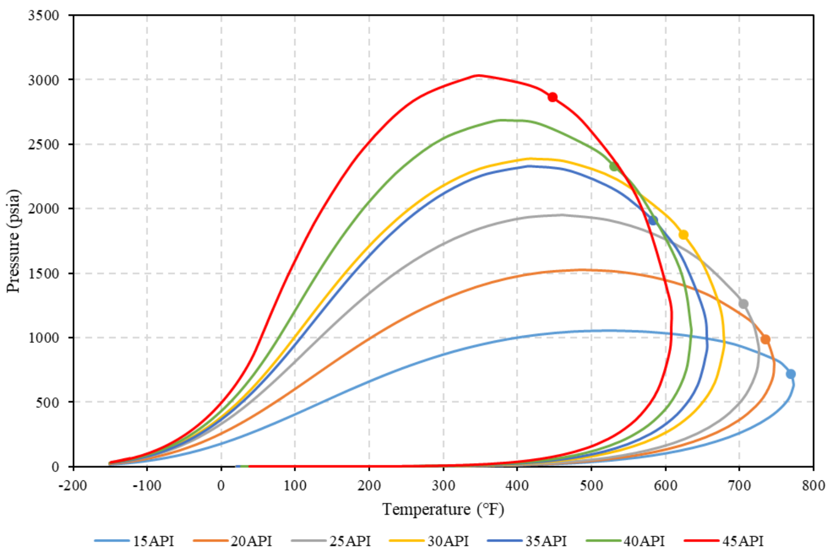

Figure 5. Reservoir Fluid Phase Envelope

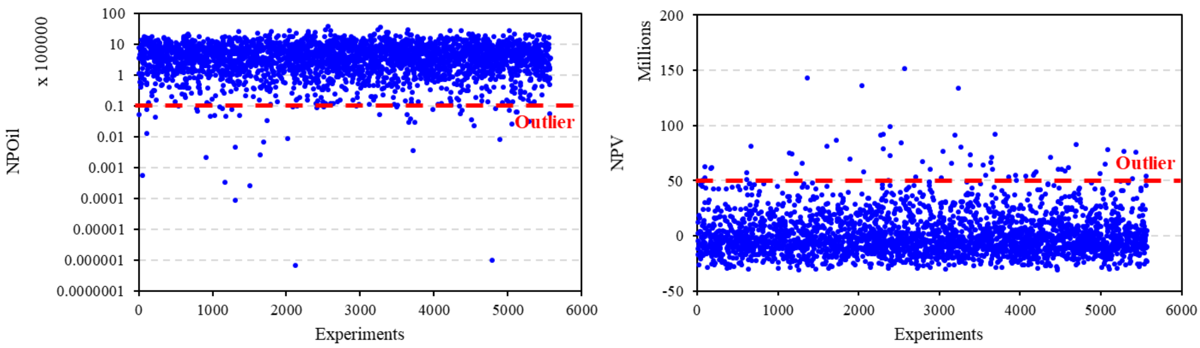

Figure 6. Experiments Results Distributions

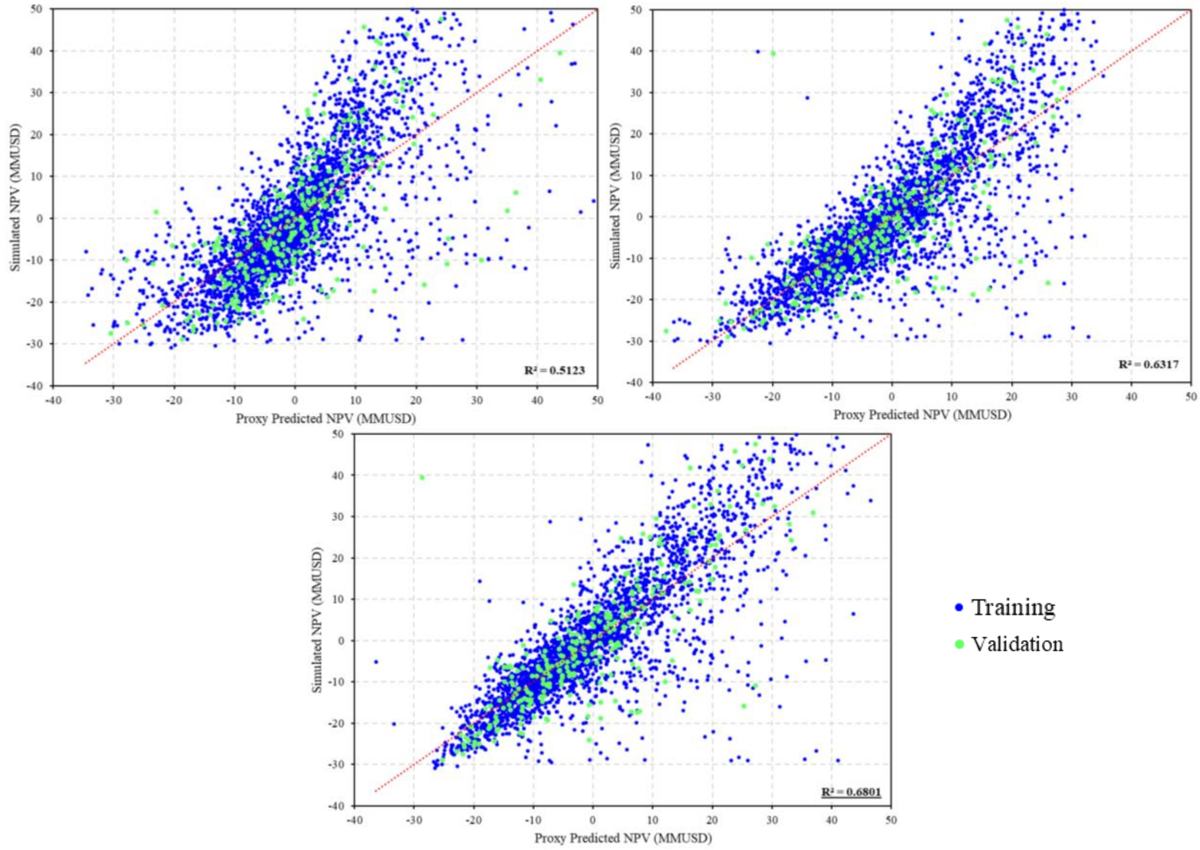

Figure 7. Self-Approach Regression Results
(Upper Left: Linear Regression, Upper Right: Simple Quadratic Regression, Middle: Quadratic Regression)

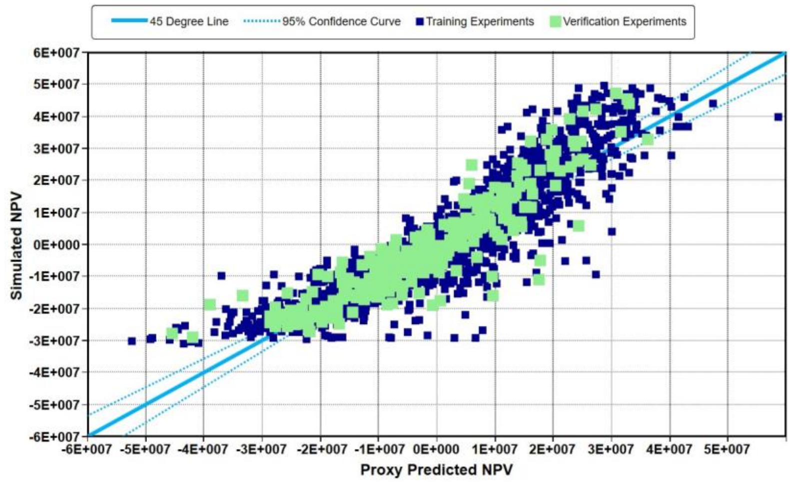

Figure 8. CMG Quadratic Regression Results

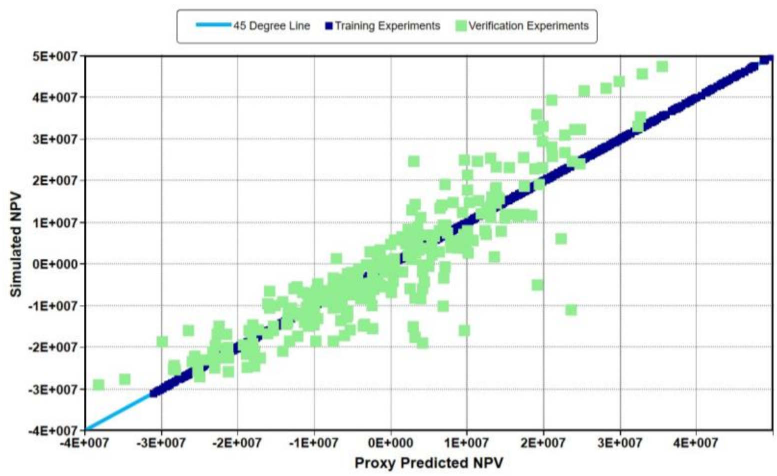

Figure 9. RBF Neural Network Results

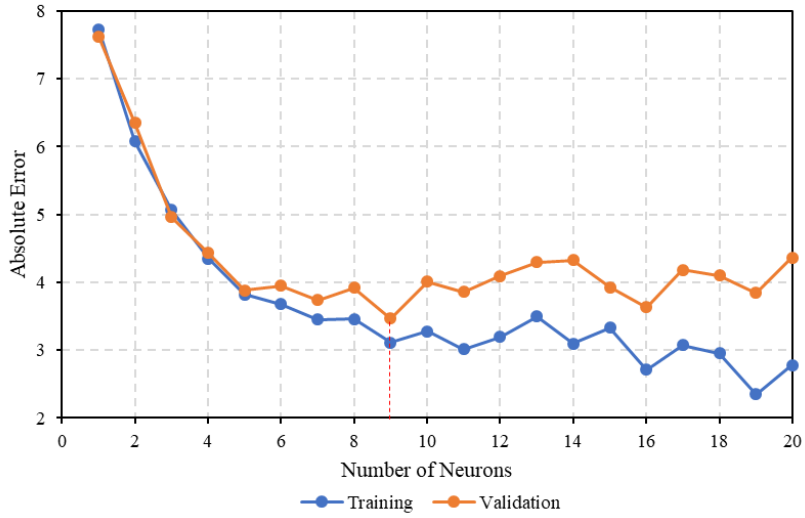

Figure 10. Sensitivity Number of Neurons for 1st Layer Neural Network

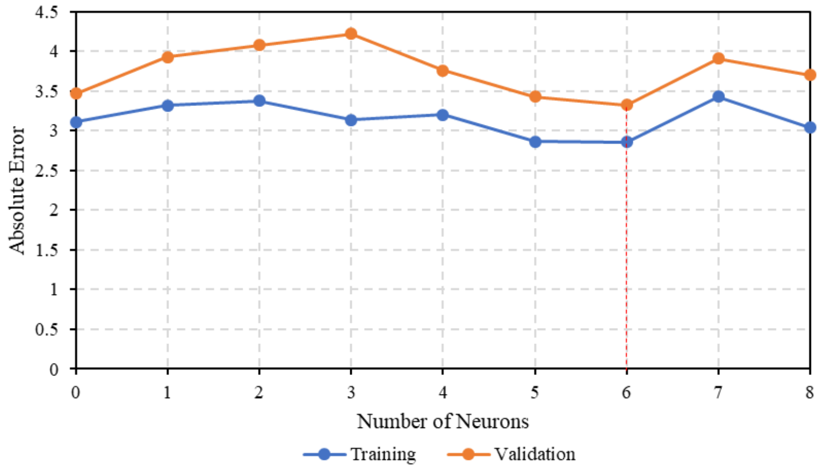

Figure 11. Sensitivity Number of Neurons for 2nd Layer of Neural Network

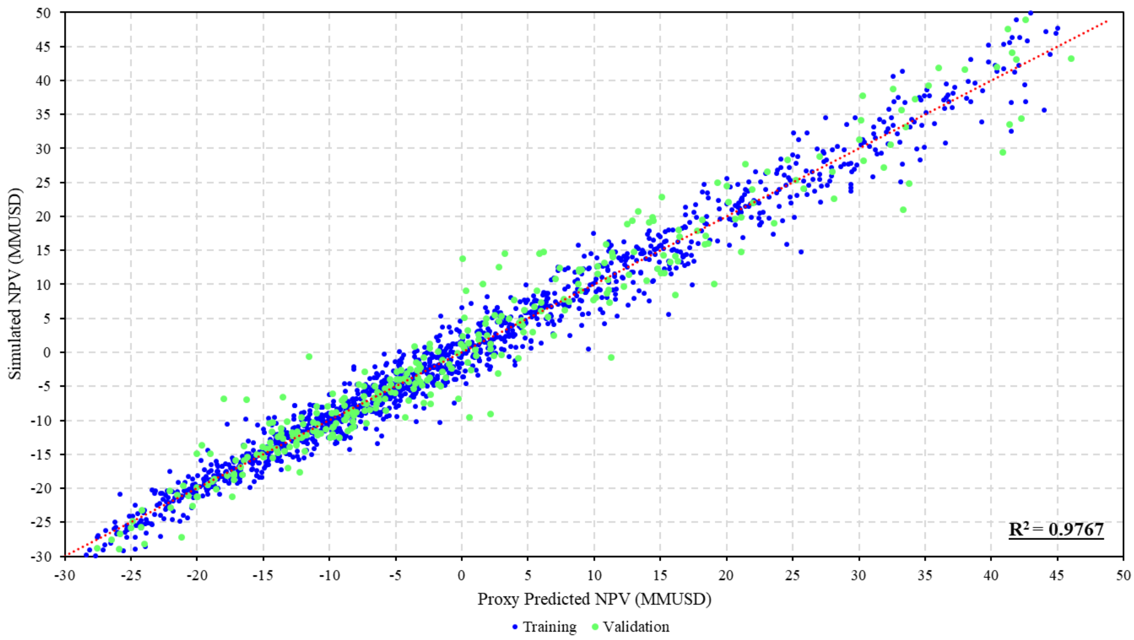

Figure 12. 20-9-6-1 Neural Network Results

List of Tables

Table 1. Parameters Used

Properties	No	Parameters	Variation	Unit	Distribution	Notes/References
Fluid	1	API	25-45	°API	Discrete	Jarrel, 2002
Rock-Fluid	2	Krgcl	0.6-0.9	-	Uniform	
	3	Krocw	0.7-0.85	-	Uniform	Saad, 1995
	4	Krwiro	0.2-0.5	-	Uniform	Saad, 1995
	5	Ng	2-4	-	Uniform	CMG Template
	6	Nog	2-4	-	Uniform	CMG Template
	7	Now	2-4	-	Uniform	CMG Template
	8	Nw	2-4	-	Uniform	CMG Template
	9	Sgcon	0.05-2	-	Uniform	
	10	Sorg	0.1-0.4	-	Uniform	Kantzas, 2001
	11	Sorw	0.1-0.3	-	Uniform	Kennaird, 1988
	12	Swcon	0.2-0.35	-	Uniform	Kennaird, 1988
	13	OilWaterTable		-	Formula	Generalized Corey
	14	GasWaterTable		-	Formula	Generalized Corey
Reservoir	15	Depth	2500-8500	ft	Triangle	Aladassani, 2010 & Taber, 1996
	16	Permi	1.5-300	mD	Triangle	Aladassani, 2010 & Taber, 1996
	17	POR	0.03-0.3	-	Triangle	Aladassani, 2010 & Taber, 1996
	18	N	22-44	-	Discrete	Grid for 10-40 acre
Initial Conditions	19	DPres	0.1-0.6	-	Uniform	
	20	Pres		psi	Formula	Pres=0.012*ρ_o*Depth*(1-DPres)
	21	h	25-200	ft	Uniform	Aladassani, 2010 & Taber, 1996
	22	Swinit	0.4-0.75	-	Normal	Jarrel, 2002
Wells & Recurrent	23	BHPProd		psi	Formula	BHP = 2/3*Pres
	24	Qinj	400-2000	MSCF	Uniform	
Economic	25	OilPrice	40-90	USD	Uniform	2017-2019 Oil Price Statistic

Table 2. Proxy Result Statistical Review

Statistics	Regression				Neural Network	
	Linear	Simple Quadratic	Quadratic	CMG Quadratic	CMG RBF NN	Multilayer ANN
R square	51.14%	63.03%	68.01%	81.97%	99.99%	97.67%
Max Absolute Error	56.69	62.25	70.07	36.76	1.20E-11	10.891
Avg Absolute Error	8.37	6.81	6.11	4.91	2.25E-12	1.808
1st Quartile	36%	23%	19%	17%	0%	5%
2nd Quartile	61%	46%	40%	36%	0%	13%
3rd Quartile	105%	100%	89%	73%	0%	32%

Table 3. Proxy Verification Statistical Review

Statistics	Regression				Neural Network	
	Linear	Simple Quadratic	Quadratic	CMG Quadratic	CMG RBF NN	Multilayer ANN
Max Absolute Error	40.80	59.30	68.12	28.59	34.54	13.76
Avg Absolute Error	8.29	6.47	5.99	4.87	4.93	2.57
1st Quartile	37%	21%	18%	17%	17%	7%
2nd Quartile	67%	49%	42%	35%	39%	19%
3rd Quartile	103%	106%	95%	71%	76%	38%

Appendix A. Capital Expenditure Cost Model

Cost Model Assumption:

- 4 producer wells that fulfill the distance for 5-spot pattern injection existed.
- An injector well needs to be drilled to the same depth of producer wells.
- Only injector well needs flowline to connect the surface facilities with wellhead.
- Injection rate held constant for 10 years.

Items	Equation for evaluation (USD)	References
Site	$C_{site} = 30{,}000\ USD/km^2$	Wei, 2015
Drilling	$C_{drilling} = 127{,}100 * e^{2.4384*10^{-4}*Depth\,(ft)} + 530.7$	Wei, 2015
Injection Well – Pattern Cost	$C_{injector} = N_{well} * 100{,}100 * e^{0.0008}$	Wei, 2015
Producer Well – Pattern Cost	$C_{producer} = N_{wells} * 3200 * e^{0.975}$	Wei, 2015
Lease – Pattern Cost	$C_{lease} = 32000 * e^{0.0299}$	Wei, 2015
Equipment – Pattern Cost	$C_{equipment} = 29000 * e^{2.81*10^{-4}}$	Wei, 2015
Pattern Cost	$C_{pattern} = C_{injector} + C_{producer} + C_{lease} + C_{equipment}$	Wei, 2015
Workover	$C_{WO} = N_{wells}\left[0.48 * C_{drilling} + 0.5 * \frac{C_{producers}}{N_{wells}}\right]$	Wei, 2015
Flowline	$C_{flowline} = N_{flowline} * 43600 * \sqrt{\frac{7389}{280*N_{well}}}$	Wei, 2015
Surface Facilities	$C_{surface} = 0.1 * C_{pattern} + 1{,}308{,}500$	Jarrel, 2002 Wei, 2015
Injection Facilities	$C_{injection} = 1.2 * Q_{inj}(SCF)$	Advanced Resource, 2006

$$CAPEX = 1{,}338{,}500 + 127{,}100 * e^{2.4384*10^{-4}*Depth\,(ft)} + 530.7 + 166{,}420.33 + N_{wells}\left[0.48 * C_{drilling} + 0.5 * \frac{C_{producer}}{N_{wells}}\right] + N_{flowline} * 43600 * \sqrt{\frac{7389}{280*N_{well}}} + 0.1 * C_{pattern} + 23000 + 25 * Depth(ft) + 1.2 * Qinj\ (SCF)$$

$$CAPEX = 1{,}338{,}500 + 127{,}100 * e^{2.4384*10^{-4}*Depth\,(ft)} + 530.7 + 166{,}420.33 + 4\left[0.48 * 127{,}100 * e^{2.4384*10^{-4}*Depth\,(ft)} + 530.7 + 0.5 * \frac{8483.73}{4}\right] + 1 * 43600 * \sqrt{\frac{7389}{280}} + 0.1 * 166{,}420.33 + 1.2 * Qinj\ (SCF)$$

$$CAPEX = 1{,}752{,}433.13 + 371132 * e^{2.4384*10^{-4}*Depth\,(ft)} + 25 * Depth(ft) + 1.2 * Q_{injmax}\ (SCFD)$$

Appendix B. Operational Expenditure Cost Model

Cost Model Assumption:

- Injection rate held constant for 10 years.
- 1 metric ton CO_2 equals 19274.47 SCF.
- Approach done in yearly cashflow.

Items	Equation for evaluation (USD)	References
Daily O&M	$O\&M_{daily} = N_{well} * 7{,}596$	Wei, 2015
Consumables O&M	$O\&M_{consumables} = N_{well} * 20{,}295$	Wei, 2015
Surface O&M	$O\&M_{surface} = N_{well} * \left[15{,}420 * \left(\frac{m_{CO_2}}{280 * N_{well}}\right)^{0.5}\right]$	Wei, 2015
Subsurface O&M	$O\&M_{subsurface} = N_{well} * 5{,}669 * \frac{Depth\ (meter)}{1{,}219}$	Wei, 2015
CO_2 Compression Cost	$C_{compression} = 13\ m_{CO_2}$	Wei, 2015
CO_2 Recycling Cost	$C_{recycle} = 23.66\ m_{CO_2}$	Wei, 2015
Lifting Cost	$C_{lifting} = 0.25 * N_{oil}$	Advanced Resource, 2006
CO_2 Cost	$C_{CO_2} = 1.05 * Q_{inj}(MSCF)$	Danish EA, 2017
CO_2 Transport Cost	$C_{CO_2\ transport} = 2.1 * Q_{inj}(MSCF)$	VITTRANS, 2000
Monitoring and verification	$C_{M\&V} = 0.1\ m_{CO_2}$	Wei, 2015

$$OPEX = N_{well} * 7{,}596\ USD + N_{well} * 20{,}295 + N_{well} * \left[15{,}420 * \left(\frac{m_{CO_2}}{280 * N_{well}}\right)^{0.5}\right] + N_{well} * 5{,}669 * \frac{Depth\ (meter)}{1{,}219} + 13 * m_{CO_2} + 23.66\ m_{CO_2} + 0.25 * N_{oil} + 1.05 * Q_{inj} + 2.1 * Q_{inj} + 0.1\ m_{CO_2}$$

$$OPEX = 27{,}891 + 209.9 * Q_{inj}(MSCF)^{0.5} + 15.25 * Depth\ (ft) + 0.25 * N_{oil} + 5.057 * Q_{inj}(MSCF)$$

Appendix C. Generated Predictive Model

In this appendix, listed all predictive model made in this study. To be noted that some model generates NPV in millions USD and some of them in USD.

a. Self-Approached Regression

Here listed the constants for equation 3, 4 and 5. This method generates NPV in millions USD.

Constant	Linear Eq. 3	Sim Quad Eq. 4	Quad Eq. 5
A0	-43.594	-31.854	32.4750
A1	-0.001	0.000	-0.0001
A2	2.788	-0.471	-6.4940
A3	6.694	11.037	-14.6203
A4	-0.009	-0.018	0.0077
A5	0.043	0.060	0.0243
B1	-	0.000	0.0000
B2	-	0.166	0.2677
B3	-	-0.951	1.5958
B4	-	0.000	0.0000
B5	-	0.000	-0.0001

Constant	Linear Eq. 3	Sim Quad Eq. 4	Quad Eq. 5
C1	-	-	0.0001
C2	-	-	-0.0001
C3	-	-	0.0000
C4	-	-	0.0000
C5	-	-	1.1300
C6	-	-	-0.0003
C7	-	-	0.0005
C8	-	-	-0.0029
C9	-	-	0.0020
C10	-	-	0.0001

b. CMG Generated Regression

Below tabulated the equation, parameter and coefficient to generate the value of NPV in USD.

$$NPV = Coeff_1 * Term_1 + Coeff_2 * Term_2 \ldots\ldots\ldots + Coeff_{n-1} * Term_{n-1} + Coeff_n * Term_n$$

Term	Coefficient
Intercept	-1.3674E+08
API	5.4232E+05
N	1.5915E+06
h	1.9901E+05
POR	2.3533E+08
PERMI	5.5750E+04
Depth	4.1629E+03
Krocw	2.3936E+08
Krwiro	-1.8320E+07
Krgcl	-1.1524E+08
Nw	-3.0415E+06
Now	-4.8773E+05
Ng	-4.7968E+06
Nog	-5.4968E+06
Sgcon	1.8862E+07
Sorg	-1.4060E+08
Sorw	1.1442E+08
Swcon	-6.3699E+07
DPres	1.8405E+07
Swinit	1.9606E+08
OilPrice	-1.1328E+05
Qinj	-2.8518E+01
API*API	-2.8870E+04
API*N	1.2402E+04
API*POR	7.1726E+05
API*PERMI	-3.6492E+02
API*Depth	1.8434E+02
API*Krocw	-7.3169E+05
API*Krgcl	2.4370E+05
API*Nw	-4.2181E+04
API*Ng	5.9598E+04
API*Sgcon	1.0157E+06
API*Sorw	-7.0012E+05
API*Swcon	-2.0772E+06
API*DPres	-1.3166E+06
API*Swinit	2.1254E+06
API*OilPrice	-5.5257E+03
API*Qinj	-3.4258E-01
N*N	-1.6089E+04
N*POR	-6.5220E+05
N*Depth	-4.1388E+01
N*Krgcl	3.2516E+05
N*Nw	1.1136E+05
N*Now	-1.2292E+05
N*Sorw	-1.0520E+06
N*Swcon	1.9369E+06
N*Swinit	-2.4907E+06
N*OilPrice	5.3364E+03
N*Qinj	4.6695E-01
h*h	-4.8216E+02
h*POR	1.1481E+05
h*Depth	6.0522E+00
h*Nw	1.6895E+04
h*Now	-1.3491E+04
h*Ng	6.6113E+03
h*Sgcon	8.3018E+04
h*Sorw	-2.0175E+05

Term	Coefficient
h*Swcon	2.3973E+05
h*DPres	-5.9094E+04
h*Swinit	-4.4590E+05
h*OilPrice	1.1469E+03
h*Qinj	6.1942E-02
POR*POR	-2.8451E+08
POR*Depth	-5.1493E+03
POR*Nw	4.5763E+06
POR*Now	-8.0969E+06
POR*Sorw	-1.5553E+08
POR*Swcon	1.5781E+08
POR*Swinit	-3.0552E+08
POR*OilPrice	8.3500E+05
POR*Qinj	5.4217E+01
PERMI*Krocw	-5.7105E+04
Depth*Depth	-7.2044E-01
Depth*Nw	-3.0458E+02
Depth*Now	2.6667E+02
Depth*Ng	-3.6334E+02
Depth*Sgcon	-8.8966E+03
Depth*Sorg	6.1162E+03
Depth*Swcon	4.5250E+03
Depth*DPres	5.5997E+03
Depth*Swinit	-5.0058E+03
Depth*OilPrice	2.3533E+01
Depth*Qinj	1.5561E-03
Krocw*Krocw	-1.5156E+08
Krocw*Nog	4.8725E+06
Krocw*OilPrice	2.8718E+05
Krwiro*Krgcl	1.9705E+07
Krgcl*Krgcl	5.8072E+07
Nw*OilPrice	3.0552E+04
Nw*Qinj	-5.3674E-01
Now*Now	6.4177E+05
Now*Nog	3.9062E+05
Now*Sorw	4.1269E+06
Now*OilPrice	-2.9907E+04
Ng*Ng	5.5646E+05
Ng*DPres	4.3368E+06
Ng*Qinj	6.5219E-01
Sgcon*DPres	4.4561E+07
Sorg*Sorg	2.3960E+08
Sorg*OilPrice	3.2050E+05
Sorw*DPres	-2.5659E+07
Sorw*Swinit	-5.9637E+07
Swcon*DPres	-5.4394E+07
Swcon*OilPrice	3.0362E+05
DPres*DPres	-3.4313E+07
DPres*Swinit	4.1284E+07
DPres*OilPrice	-1.9751E+05
DPres*Qinj	-8.1253E+00
Swinit*Swinit	-4.1578E+07
Swinit*OilPrice	-8.4479E+05
Swinit*Qinj	-8.5033E+00
OilPrice*Qinj	1.0767E-01
Qinj*Qinj	-3.0948E-06

c. CMG RBF Neural Network

Here attached the CMG RBF Function to be used in excel VBA and generates NPV in USD.

```
Function NormalizeArray(npar As Integer, arr() As Double) As Double()
Dim resultArray() As Double
Dim a, b, c, d As Double
ReDim resultArray(npar)
Dim i As Integer
For i = 2 To npar + 1
resultArray(i - 1) = (arr(i - 1) - (Worksheets(2).Cells(i, 1) + Worksheets(2).Cells(i, 2))
/ 2) / (Worksheets(2).Cells(i, 1) - Worksheets(2).Cells(i, 2)) / 2
c = Worksheets(2).Cells(i, 1)
d = Worksheets(2).Cells(i, 2)
a = arr(i - 1)
b = resultArray(i - 1)
Next i
Debug.Print resultArray(npar)
NormalizeArray = resultArray()
End Function

Function GetYMax() As Double
Dim weightsNumber, yMaxIndex As Integer
Dim yMax As Double
yMaxIndex = Worksheets(2).Range("C1").End(xlDown).row
yMax = Worksheets(2).Cells(yMaxIndex, 3).value
GetYMax = yMax
End Function

Function Distance(npars As Integer, arr1() As Double, arr2() As Double) As Double
Dim res As Double
res = 0
Dim i As Integer
If Worksheets(2).Cells(2, 6) = "" Then
    For i = 1 To npars
        res = res + (arr1(i) - arr2(i)) * (arr1(i) - arr2(i))
    Next i
Else
    For i = 1 To npars
        res = res + Worksheets(2).Cells(i + 1, 6).value * (arr1(i) - arr2(i)) * (arr1(i) -
arr2(i))
    Next i
End If
res = Sqr(res)
Distance = res
End Function

Function Power(rad As Double) As Double
Dim res As Double
If rad < 0.000000001 Then
    res = 0
Else
    res = 0.01 * Exp(1.5 * Log(rad))
End If
Power = res
End Function

Function RBFInterpolate(npars As Integer, npt As Integer, pars() As Double) As Double
Dim lastval As Double
lastval = 0
Dim curVec() As Double
ReDim curVec(npt + 1)
For i = 2 To npt + 1
        Dim curPar() As Double
        ReDim curPar(npars)
        Dim curNormPar() As Double
        ReDim curNormPar(npars)
        Dim j As Integer
        For j = 2 To npars + 1
            curPar(j - 1) = Worksheets(1).Cells(i, j).value
        Next j
        curNormPar = NormalizeArray(npars, curPar)
        curVec(i - 1) = Power(Distance(npars, pars, curNormPar))
        xval = curVec(i - 1)
Next i
curVec(npt + 1) = 1
For i = 2 To npt + 2
    lastval = lastval + curVec(i - 1) * Worksheets(2).Cells(i, 3)
Next i
RBFInterpolate = lastval
End Function
```

```
Function CMG_RBF(rng As Range) As Double
    Dim interRes As Double
    Dim npars As Integer
    Dim lastColumn As Long
    lastColumn = Sheet1.Cells(1, Columns.count).End(xlToLeft).Column - 1
    npars = lastColumn - 2
    If rng.Columns.count <> npars Then
        MsgBox ("Range must have " & npars & " parameters")
        interRes = 0
    Else
        Dim curPar() As Double
        ReDim curPar(npars)
        Dim cell As Range
        Dim im, iErr, i As Integer
        im = 1
        iErr = 0
        For Each cell In rng
            If IsNumeric(cell.value) = False Or cell.value = "" Then
                    iErr = iErr + 1
            Else
                    curPar(im) = cell.value
                    im = im + 1
            End If
        Next cell

        Dim curNormPar() As Double
        ReDim curNormPar(npars)
        If iErr = 0 Then
            curNormPar = NormalizeArray(npars, curPar)
            Dim yMax As Double
            yMax = GetYMax()
            Dim yMaxIndex As Integer
            yMaxIndex = Worksheets(2).Range("C1").End(xlDown).row
            interRes = RBFInterpolate(npars, yMaxIndex - 3, curNormPar)
            interRes = interRes * yMax
        Else
            interRes = 0
            MsgBox ("Must have a valid numeric parameter value in each cell, iErr = " &
iErr)
        End If
    End If
    CMG_RBF = interRes
End Function

Sub btnS()
    Dim npars, nrows As Integer
    Dim lastColumn As Long
    lastColumn = Sheet1.Cells(1, Columns.count).End(xlToLeft).Column - 1
    npars = lastColumn - 2
    Dim tit As String
    tit = Worksheets(2).Cells(4, 5).value
    nrows = Worksheets(2).Cells(2, 4).value + Worksheets(2).Cells(3, 4).value + 1
    Dim str1, str2, str3, str4, str5, str6, str7 As String
    str1 = "Proxy Type: Radial Basis Functions"
    str2 = "Objective Function: " & tit
    str3 = "Number of Parameters: " & npars
    str4 = "Formula Syntax: CMG_RBF(PARAMETER_RANGE)"
    str5 = "Calculation Method:"
    str6 = "1. Type or copy/paste " & npars & " parameter values as a row into the
corresponding parameter range cells below line " & nrows & "."
    str7 = "2. type '=CMG_RBF(PARAMETER_RANGE)' or copy/paste (as a function) the cell ("
& nrows & ", " & lastColumn & ") into the cell corresponding to the objective function ("
& tit & ")"

    MsgBox (str1 & vbNewLine & str2 & vbNewLine & str3 & vbNewLine & str4 & vbNewLine &
vbNewLine & str5 & vbNewLine & vbNewLine & str6 & vbNewLine & str7)
End Sub

Sub Tester(row As Integer, col As Integer)
        Dim btn As Button
        Application.ScreenUpdating = False
        ActiveSheet.Buttons.Delete
        Dim t As Range
           Set t = ActiveSheet.Range(Cells(row, col + 2), Cells(row, col + 3))
           Set btn = ActiveSheet.Buttons.Add(t.Left, t.Top, t.Width, t.Height)
           With btn
              .OnAction = "btnS"
              .Caption = "Description"
              .Name = "Description"
           End With
        Application.ScreenUpdating = True
End Sub
```

d. Multilayer Artificial Neural Network

Here attached weight factors generated for 20-9-6-1 neural network to produce expected value of NPV in MMUSD. All layers used sigmoid/logistic function as their activation function.

Input Nodes	1st Layer Nodes								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
API	-0.5962	-0.0631	0.4109	-0.4990	0.2124	4.2485	1.7934	0.5576	1.5794
PERMI	-0.1053	0.0139	0.0032	-0.0665	0.0064	0.2047	-0.0360	0.0349	0.1437
POR	0.1452	0.0900	-3.2706	-0.0759	0.1522	-0.0301	-2.0497	0.2077	-0.3225
AH	0.1368	3.5700	-0.6419	-0.0095	0.3045	0.6908	-2.7184	0.0073	-0.5872
Depth	-0.0346	-0.0609	0.1045	0.0392	-0.0422	-0.6127	0.1041	0.1150	0.1652
Krocw	-0.0781	0.0125	-0.0392	0.0973	-0.0075	-0.0160	0.0687	-0.0055	0.0306
Krwiro	0.0782	-0.0473	-0.0066	-0.1090	0.0261	-0.0323	-0.0052	0.0647	0.0610
Krgcl	-0.0167	-0.0315	0.0177	0.0017	-0.0085	0.1669	-0.0564	-0.0171	-0.0024
Nw	-0.1701	0.0084	0.0545	0.1229	-0.0011	0.1129	0.1845	-0.0491	-0.0042
Now	0.1674	-0.0116	-0.0039	0.0654	-0.0210	-0.2029	0.1920	-0.0047	-0.0488
Ng	-0.0822	-0.0290	0.0171	0.1301	0.0130	-0.2860	-0.1439	-0.0006	-0.0226
Nog	0.1072	0.0228	-0.0280	-0.2028	0.0254	0.1718	0.2288	0.0542	0.0990
Sgcon	-0.0991	0.0148	0.0309	0.1372	-0.0382	-0.3477	-0.0669	-0.0835	-0.1300
Sorg	0.1277	-0.0305	0.0415	-0.1364	0.0238	0.2419	0.1329	0.0219	-0.0168
Sorw	0.1083	0.0678	-0.0962	0.1158	-0.0188	-0.2896	-0.1044	0.0815	0.1603
Swcon	-0.2465	-0.0131	0.0496	0.2491	0.0482	0.0652	0.4366	0.0743	0.1276
Pres	-0.9149	0.1881	-0.1726	2.0282	-0.1585	-9.5906	-2.6720	-0.4763	-0.6301
Swinit	1.2481	-0.1111	0.0770	-1.5222	-0.1482	-0.6819	0.0071	0.0416	0.3670
OilPrice	0.1868	-0.2591	0.2235	-0.1642	0.5745	0.0553	-0.0966	0.6058	0.0941
Qinj	-0.2335	-0.2075	0.2941	0.2560	-0.3917	-0.4027	1.9485	0.2433	0.0374
Bias	-1.2255	4.3655	-4.5952	1.4605	0.7717	-9.2116	4.2722	0.3352	-0.3323

1st Layer Nodes	2nd Layer Nodes					
	1	**2**	**3**	**4**	**5**	**6**
1	-2.481	13.386	-0.778	1.189	2.146	1.657
2	-3.379	-8.931	5.095	3.658	-2.539	1.141
3	3.035	10.047	0.588	-2.901	0.536	-0.410
4	2.026	9.321	-0.058	-2.612	-0.257	0.303
5	0.661	-3.575	-4.712	-1.543	4.811	5.053
6	-3.142	2.106	-0.611	3.169	3.420	0.504
7	6.549	-0.262	0.602	-6.075	-2.740	-1.465
8	-2.344	-5.499	6.430	4.619	-7.424	-4.948
9	-3.094	8.141	-3.201	1.454	4.029	1.910
Bias	-1.144	3.635	-6.594	1.824	6.182	0.366

2nd Layer Nodes	Output
1	5.8514
2	-7.4070
3	-8.5784
4	5.5111
5	-11.770
6	11.0305
Bias	2.4979