SPE-187112-MS

Comparison of Shale Oil Production Forecasting using Empirical Methods and Artificial Neural Networks

Anuj Suhag, Rahul Ranjith, and Fred Aminzadeh, University of Southern California

This paper was prepared for presentation at the SPE Annual Technical Conference and Exhibition held in San Antonio, Texas, USA, 9-11 October 2017.

Abstract

The objective of this work is to evaluate the efficacy of empirical models in forecasting oil production in shale reservoirs, bycomparing and analyzing their fit and effectiveness to our dataset. The following three modelswere considered: A Conventional Decline Curve Analysis (CDC), an Unconventional Rate Decline (URD) Approach, and a Logistics Growth Analysis (LGA) method. A comparative study is performed to evaluate the use of Artificial Neural Networks (ANN) for production forecasts and to reinforce the thinking that it is imperative to include physical parameters in mathematical models to predict accurateforecasts.

For this project, we used non-linear regression to fit empirical models to the dataset obtained from North Dakota Industrial Commission (NDIC). We evaluated the fit of modelswith the help of coefficient of determination. Physical parameters, such as porosity, saturation, shale volume, etc., and log data from sonic logs, gamma ray logs, etc., were selected as input to the ANN model andwere aided by Analysis of Variances (ANOVA).

Amongst the empirical models for shale play, URD method is the most commonly used since it is idealfor fractured reservoirs with extremely low permeability. URD model did fit the cumulative production profiles, but could not accurately fit the monthly production profile. The CRD approach was overallunsuccessful in generating accurate future production profiles. Values forecasted from the ANN show less than 10% error in estimation. The inclusion of physical parameters has proven to be extremely promising in the forecast production from fields that do not have sufficient history for statistical fitting.

Through aselection of physical properties from different sources, we have built an ANN model that fits with the production data in wells that have adiverse production history. Our work has shown the importance of including physical parameters into a process that was heretofore seen as a time series regression problem. In general, our new ANN-based method generated the best results.

Introduction

Hydrocarbon accumulations in petroleum reservoirs around the world migrated from fine-grained, black, organic-rich sedimentary source rocks, also known as organic shales. Historically, thoseshale formations have been regarded as source rocks for hydrocarbon origination and migration to sandstone or limestone reservoirs. When massive amounts of organic debris deposits in a marine-rich environment and undergoesrapid burial (Passey et al., 2010), oil and gas shalelayersare formed. Consequently, over time, these organic constituents convert into hydrocarbons under the effect of temperature and pressure changes in the subsurface due to burial.

The Bakken Shale in North Dakota, USA, is one of the largest global shale plays, having an estimated 7.4 billion barrels of recoverable oil. The formation comprises of three layers: lower shale, middle dolomite, and upper shale. The shaleswere deposited in relatively deep anoxic marine conditions, and the dolomite was deposited as a coastal carbonate bank during a time of shallower, well-oxygenated water. The middle dolomite member is the principal oil reservoir, roughly two miles below the surface. Both the upper and lower shale members are organic-rich marine shale. To date, about a billion barrels of Bakken crude has been recovered. Despite the price crash in 2014, as of January 2017, daily production values have averaged 280,000 bpd.

Prediction methods for production performance that are reliable and repeatable have been problematic in shale reservoirs. The ultra-low permeability of shale, complex flow mechanisms, and formation heterogeneity are major contributors to this challenge. Empirical, analytical and PVT characteristics, inadequate data availability, and time can all negatively impact accuracy. Currently, a single-phase flow is assumed for analytical methods for forecasting production, thereby arguably making the application for multi-phase flow analysis redundant.

Empirical Methods

Within the currently used empirical methods for production forecasting in shale reservoirs, three have shown consistent reliability:

- Duong's Unconventional Rate Decline (URD) Method
- Arps Conventional Decline Curve (CDC) Analysis
- Logistic Growth Analysis (LGA)

Duong's Method

This method addresses the fracture-dominated flow in shale reservoirs. Traditional decline curve analysis (DCA) is based on a drainage area with relatively good matrix permeability, which subsequently establishes pseudo-radial and Boundary Dominated Flows (BDF). In a shale reservoir, since the drainage zone encompasses mainly natural fractures and the Stimulated Reservoir Volume (SRV), it can be considered as a fracture-dominated flow reservoir where no BDF is established due to micro-permeability of the matrix. Duong (2011) stated that with negligible matrix contribution, the determination of EUR based on the traditional concept of a drainage area is inaccurate. He also outlined a detailed procedure for evaluating and forecasting cumulative production using his model, which involves two diagnostic plots using Eq. 1 and Eq. 2.

$$q/Np = \alpha^2 \qquad [1]$$

Where*a* is the intercept constant, *q* is the oil production rate in STB/day, and *Np* is the cumulative production in STB. A log-log plot of q/Np vs. t gives a straight line with a negative slope, *-m*, and an intercept, *a*, which are two of the four unknown parameters in this method. This can be used in,

$$q = q1 * \mathrm{t(a, m)} + q_\infty \qquad [2]$$

$$t(a,m) = t^{-m(e^{\left(\frac{a}{1-m}\right)*(t^{1-m})-1})} \qquad [3]$$

Arps Conventional Decline Curve Analysis (CDC)

Arps (1945) developed the mathematical relations for three types of graph representations of production decline for conventional reservoirs. They are exponential, hyperbolic and harmonic decline. Traditionally, CDC involves fitting a trendline through the historical performance of a well on a semi-log plot and extrapolating that line to estimate future production performance, under the assumption that the past trend will not change under operational conditions over field life. It can be expressed as follows,

$$q = qi/(1 + (bDi * t))^{1/b} \quad [4]$$

Where D is the decline constant, b is the decline exponent, and $qi$is the initial production rate. Typically, in shale reservoirs, high initial production rates are followed by rapid decline ruled by transient flow regimes (Medeiros et al., 2008). The benefit of using the hyperbolic model for shales is that a satisfactory match is obtained for wellbore performance characterized by long transient flow regimes (Duong, 2010).

Logistics Growth Analysis (LGA)

This is an empirical mathematical technique used to analyze numerous physical trends (Clark, 2011). LGA can analyze well production performance and predict future reserves, especially in shale oil reservoirs. Clark (2011) explained the use of LGA with original concepts and modifications necessary for its application to tight permeability formations.

The recommended approach begins with the least square regression of cumulative production versus time using the following mathematical expression,

$$Q(t) = \frac{Kt^n}{a+t^n} \quad [5]$$

Where K is the carrying capacity (maximum physically recoverable oil), a is the constant of t^n when half the oil has been recovered and n is the exponential parameter. Subsequently, the corresponding equation for production rate versus time is given by,

$$q(t) = \frac{Kna\, t^{n-1}}{(a+t^n)^2} \quad [6]$$

These models consider time-series analysis. Thus, they have a high percentage of error as they do not consider the physical uniqueness of the field. From the above, Duong's method is the most widely used empirical method for shale production forecasting.

Methodology

To model the long transient linear behavior of shale wells, Arp's hyperbolic decline model fits with b>1. Thus, the extrapolation yields overestimated reserves. Additionally, as shown by Lee and Siddle 2010, with b>1, the hyperbolic equation never goes to zero and consequentially the reserves are unbounded. Logistic Growth Model addresses these problems and this method produces good results for cumulative production. However, LGA method fails to produce satisfactory results for monthly production forecast because of high variability in the data. Amongst the chosen empirical models for shale play, URD method is the most commonly used since it is ideal for fractured reservoirs with extremely low permeability. For this paper, we have compared the results generated using ANN model against the forecasted production from the Duong's URD method.

Duong's method is highly sensitive to datapre-processing and is unreliable for forecasting in wells with high production variability. Diagnostic plots were used to exclude data samples that do not fit a visible decline curve. This was done to achieve a minimum coefficient of determination (R^2) of 0.95.

By using a non-linear regression model wherein error variance is minimized, three cases (wells) are fit to the processed dataset. Excel Multivariate Solver was used to fit Duong's cumulative production equation for each instance.

Production Forecasts using Artificial Neural Networks (ANN)

Geological data such as porosity, permeability, shale volume, oil, water, and gas saturation is the right source for information to assess the hydrocarbon potential of an oil-bearing rock formation; all can be obtained from well log data. Also, production data can be used to derive these properties inversely. Historical production data enlightens usabout the productivity of a well. Also, completion strategy and data give insight to the productivity of a well completion.

Availability of historical production data calls for uncertainty analysis approaches, whereby different models are compared to the historical data. The model that best fits the historical data can then be used for well production performance forecasting (Ani et al., 2016).

In unconventional reservoirs, geological data and completion data are not statistically correlated to each other. With the addition of production data, empirical methods cannot map the complicated relationship between them. ANN, designed akin to the neural network of a human brain, can accurately map these complex relationships and can be used for forecasting the production rates of a well.

Machine learning techniques such as ANN are designed to exhibit super intelligence by recognizing patterns from existing training data set, gather the information and infer knowledge that could be used in real world problems such as Factor (Principal Component) Analysis in conjunction with Neural Networks (Aminzadeh et al,2000), uncertainty analysis (Shahkarami et al, 2015), reservoir characterization (Nikravesh et al, 2003 and Anifowose, 2011) and Generation of Pseudologs (Tamhane et al, 2002 and Long et al, 2017). For many other applications of ANN and other soft computing techniques in the oil industry see Aminzadeh and de Groot (2006).

Methodology

The process of developing an ANN model for production forecast begins withthe selection of statistically relevant data. To gauge the strength of the relationship between input features and the test parameter, we used Multivariate Analysis of Variances;only the parameters that showed significant contribution ($p<0.05$)were selectedas input for training the ANN. Following this, we performed Pearson's Product Moment correlation to evaluate the correlationamongst the variable of entry themselves. We observed no significant correlation.

ANN model architecture development was a heuristic process. After paring down to just the relevant features with output as oil production rate, different combinations of available input variables, hidden layer neurons, learning algorithms and transfer/activation functions were used to find the optimal network architecture that gave a good fit. Training function was selected for ANN per the dataset. For this paper, we have primarily used LevenbergMarquardt algorithm for training the network. When required, Bayesian Regularization has also been usedforwellswith noisy and small datasets.

The final network architecture is shown in Figure 1, with six variables in the input layer, ten neurons in the hidden layers and output as predicted monthly oil production rates for periods of 3 and six months. To avoid over-fitting or under-fitting of the network, itwas runon the pristine portion of the dataset for testing. The network was trainedtillacceptable regression resultswere achieved.

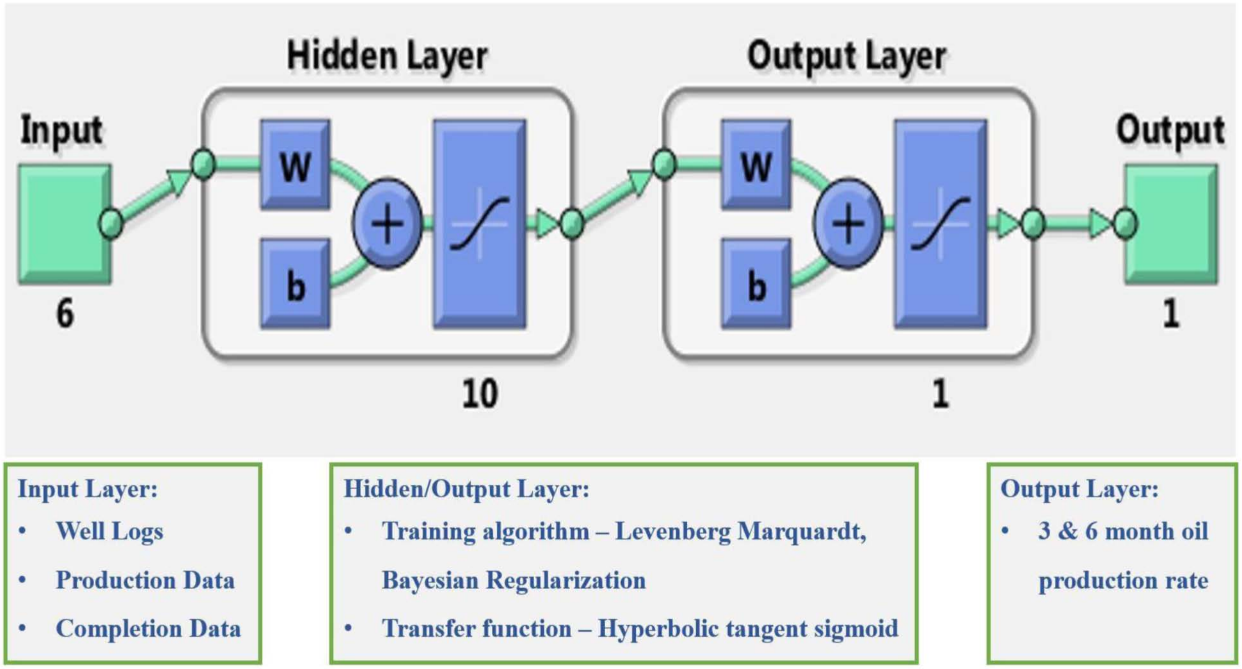

Figure 1—ANN Model Architecture

Results

Three wells from Stanley field producing from the Bakken shale play were selected for comparing the production forecasts between Duong's method and the ANN model. As a mandate, Duong's method requires a minimum of five years of historical production data to make an effective forecast; therefore, these three wells were selectedwith that in mind. Comparisons have been presented as two cases to highlight the efficacy of ANN in forecastingcumulative production as well as monthly productionrates and underscores the challenges in using Duong's method for forecasting monthly production.

For comparison purposes, Well#1 has been selected and its output detailed below.

Case 1: Comparison of Cumulative Production Forecast

ANN model forecast

Well #1 had 91 months of historical production data, from which 68 months of data were used for training, 17 months for validation, and six months as a test set. For this well, the training algorithm used was Levenberg Marquardt algorithm.

Figure 2 illustrates the regression plots of the network trained for the data from Well #1. High R-values of training, validation and test sets are indicative of the network's good fit.

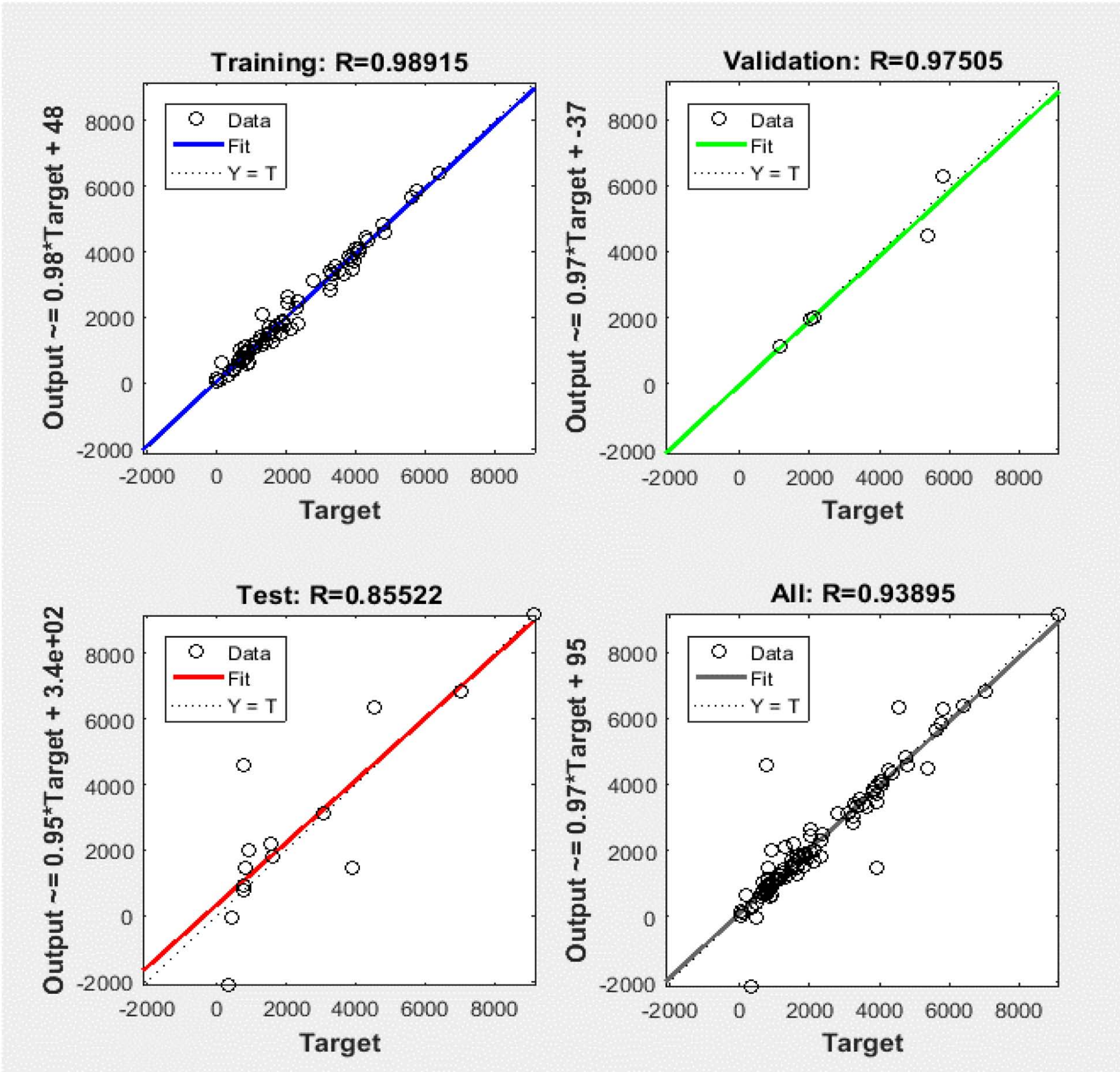

Figure 2—Regression plots for Well#1

Figure 3 is a scattered plot of the actual monthly production on the X axis and the output of the ANN for Well #1. It is evident from the linear trend-line with zero intercepts that predicted values and actual values are very close to each other. The error of estimation can be seen in Table 1.

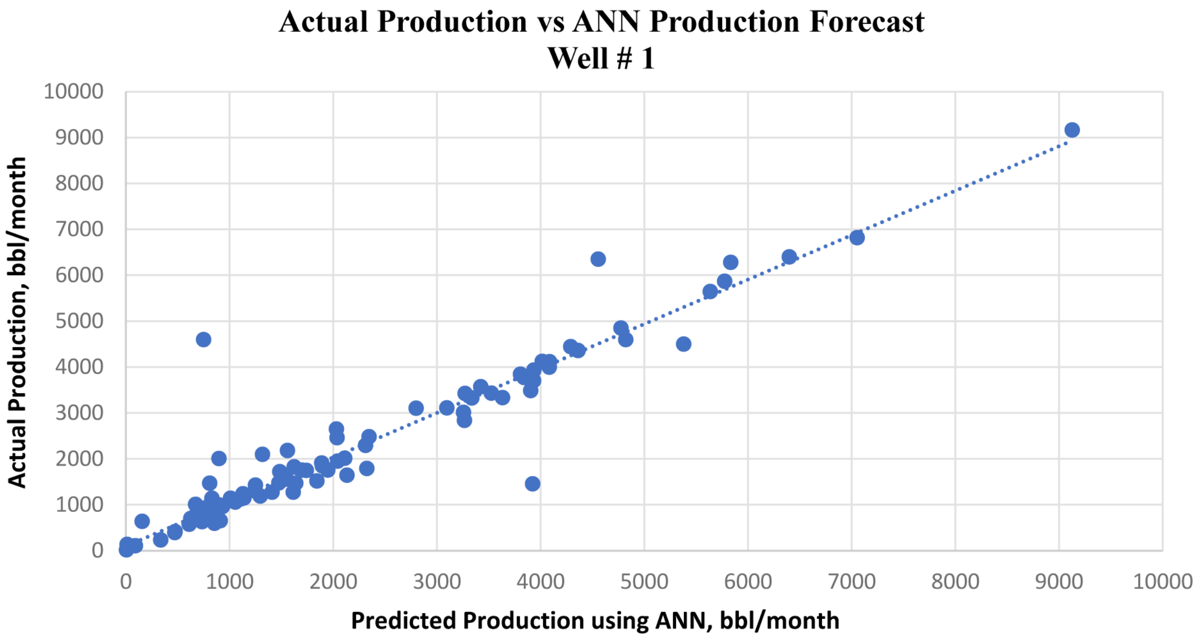

Figure 3—Actual production vs. ANN forecast

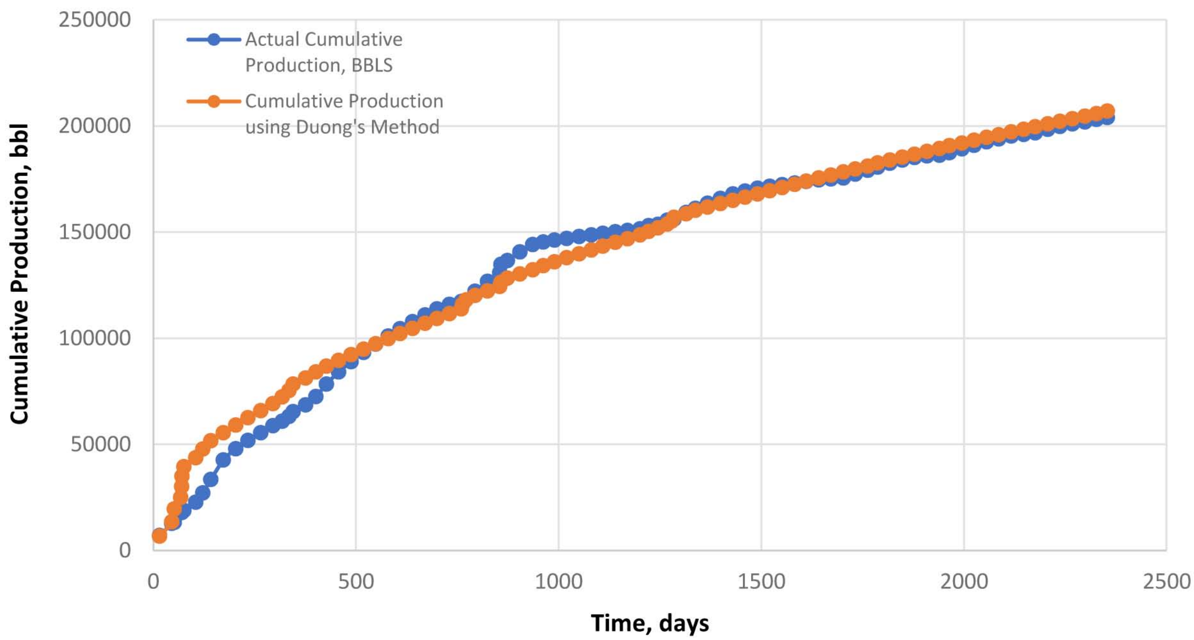

Figure 4—Actual production vs. Cumulative production using Duong's method

Table 1—Cumulative Production Forecast error estimation comparison

Month	Actual Cumulative Production	ANN Model Forecast	Duong's Method Forecast	ANN Forecast Error	Duong's Forecast Error
	bbl	bbl	bbl	%	%
85	186176	187225.61	199636.46	-0.56	-7.23
86	186182	187241.29	200887.60	-0.57	-7.90
87	186275	187345.47	202129.78	-0.57	-8.51
88	190831	193695.92	203363.14	-1.50	-6.57
89	191286	193651.61	204587.85	-1.24	-6.95
90	196925	199298.7	205804.05	-1.21	-4.51
91	203980	206116.63	207011.87	-1.05	-1.49

Case 2: Comparison of MonthlyProductionRateForecast

The same network architecture was used in this case as in the well with six months of monthly production values predicted using the ANN model. As seen in Figure 5, ANN has a good fit with the actual production curve.

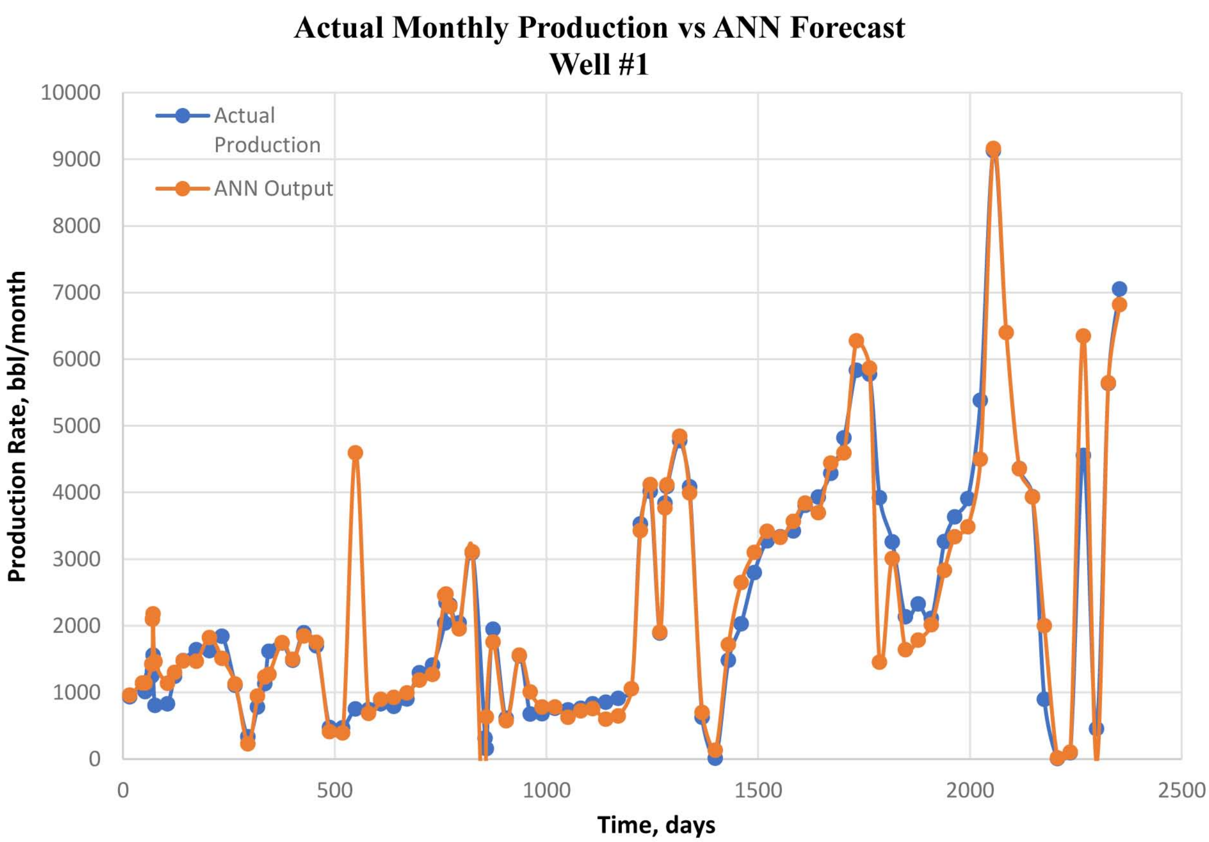

Figure 5—Actual monthly production vs. ANN forecast

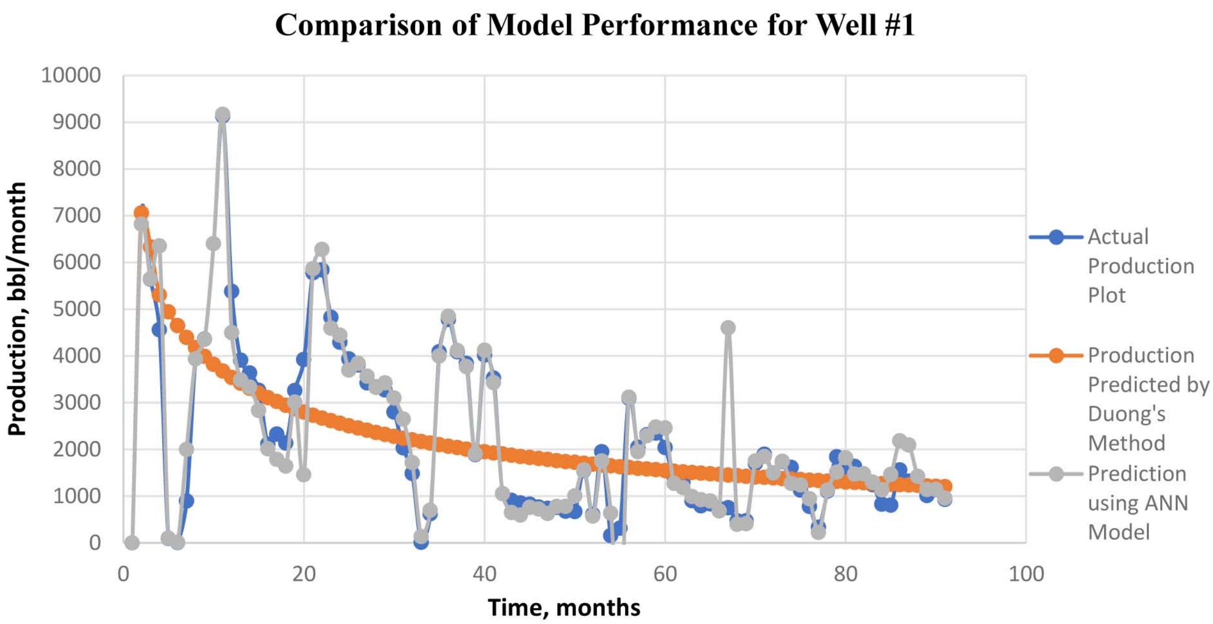

Figure 6—Comparison of ANN Forecast and Duong's Forecast to actual production

Table 2—Overview of Monthly Production Forecast comparison between ANN Model and Duong's Method

Month	Actual Monthly Production	ANN Model Forecast	Duong's Method Forecast	ANN Forecast Error	Duong's Forecast Error
	BBL	BBL	BBL	%	%
88	1249	1424.212	1229.014599	14.028143	1.600112
89	1011	1142.04	1220.428947	12.961388	20.71503
90	1139	1143.269	1211.985943	0.374792	6.407897
91	933	956.9854	1203.681791	2.5707808	29.01198
			Mean of Error (%)	7.4837761	14.43376
			SD of Error (%)	6.0726468	10.96695

Conclusions

- The focus of this study was to compare the existing practices and provide an alternate model for forecasting oil production from Bakken shale play.
- This forecast model can be applied to wells producing oil from any shale formation.
- Duong's method is unproven for wells with thetransition from thelinear or bi-linear flow. Duong's method is also extremely sensitive to data pre-processing.
- The ANN-based model used in this study addresses the physical uniqueness of each well.
- ANN based forecast model has an average error of 0.95% (underestimated) and a standard deviation of 0.35%. This is a better result than the reported error for Duong's method (mean error = 6.16% underestimated and 2.23 % standard deviation).
- ANN based forecast model shows promising results for monthly production rates.

References

1. Ani, M., Oluyemi, G., Petrovski, A., & Rezaei-Gomari, S. (2016,). Reservoir Uncertainty Analysis: The Trends from Probability to Algorithms and Machine Learning. Society of Petroleum Engineers. doi:10.2118/181049-MS
2. Anifowose, F. A. (2011,). Artificial Intelligence Application in Reservoir Characterization and Modeling: Whitening the Black Box. Society of Petroleum Engineers. doi:10.2118/155413-MS
3. Holdaway, K. R. (2012,). Oilfield Data Mining Workflows for Robust Reservoir Characterization: Part 2. Society of Petroleum Engineers. doi:10.2118/149785-MS
4. Bansal, Y., Ertekin, T., Karpyn, Z., Ayala, L. F., Nejad, A., Suleen, F., Sun, Q. (2013,). Forecasting Well Performance in a Discontinuous Tight Oil Reservoir Using Artificial Neural Networks. Society of Petroleum Engineers. doi:10.2118/164542-MS
5. Al-Otaibi, B. Z., Schechter, D. S., & Wattenbarger, R. A. (2015,). Production Forecast, Analysis, and Simulation of Eagle Ford Shale Oil Wells. Society of Petroleum Engineers. doi:10.2118/172929-MS
6. Maity, D., Aminzadeh, F., 2015. Novel Fracture Zone Identifier Attribute using Geophysical and Well Log Data for Unconventional Reservoirs, *Interpretation Journal*, Vol.**3**, No. 3, p.T155–T167
7. Zhao, B., Brouwer, F.G.C., Aminzadeh, F., Morris, S., and Harris, R., 2008, Multiple predictions and reservoir characterization of a tight sand reservoir, *SEG Las Vegas 2008 Annual Meeting*.
8. Shahkarami, A., Mohaghegh, S. D., & Hajizadeh, Y. (2015, Assisted History Matching Using Pattern Recognition Technology. Society of Petroleum Engineers. doi:10.2118/173405-MS
9. Arps, J.J. (1945) Analysis of Decline Curves. *Transactions of the AIME*, **160**, 228–247. http://dx.doi.org/10.2118/945228-G
10. Clark, A. J. (2011). *Decline Curve Analysis in Unconventional Resource Play using Logistic Growth Models*. MS Thesis, University Texas at Austin, Austin, Texas (August 2011).
11. Passey, Q., R., Bohacs, K., M., Esch, W., L., Klimentidis, R., Sinha, S., 2010, From Oil- Prone Source Rock to Gas-Producing Shale Reservoir – Geologic and Petrophysical Characterization of Unconventional Shale-Gas Reservoir, SPE International Oil & Gas Conference and Exhibition, Beijing, China, 8-10 June 2010, SPE 11350.
12. Duong, A. N. (2010,). An Unconventional Rate Decline Approach for Tight and Fracture-Dominated Gas Wells. Society of Petroleum Engineers. doi:10.2118/137748-MS
13. Ilk, D., Rushing, J. A., Perego, A. D., & Blasingame, T. A. (2008,). Exponential vs. Hyperbolic Decline in Tight Gas Sands: Understanding the Origin and Implications for Reserve Estimates Using Arps; Decline Curves. Society of Petroleum Engineers. doi:10.2118/116731-MS
14. Mattar, L. (2008,). Production Analysis and Forecasting of Shale Gas Reservoirs: Case History-Based Approach. Society of Petroleum Engineers. doi:10.2118/119897-MS
15. Delaihdem, D. K., (2013), *Decline Curve Analysis and Enhanced Shale Oil Recovery Based on Eagle Ford Shale Data*. MS Thesis. The University of Alaska, Fairbanks, Alaska.
16. Joshi, K., (2012), *Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs with Emphasis on The Duong Method*. MS Thesis, Texas A&M University, College Station, Texas.
17. Long, W., Chai, D., and Aminzadeh, F. (2017), Pseudo Density Log Generation Using Artificial Neural Network" Journal of Petroleum Technology, *Special Issue on Intelligent Field Technology*, Vol. **69**, No. 5.
18. Tamhane, D., Wong, P. M., Aminzadeh, F., 2002, Integrating Linguistic Descriptions and Digital Signals in Petroleum Reservoirs, *International Journal of Fuzzy Systems*, Vol **4**, No. 1, pp585–591.

19. Aminzadeh, F., Barhen, J., Glover, C.W. and Toomanian, N. B. 2000, Reservoir Parameter Estimation Using a Hybrid Neural Network, *Computers, and Geosciences*, Vol **26**, pp 869–875,
20. Aminzadeh, F., and de Groot, P., 2006, Neural Networks and Soft Computing Techniques, with applications in the oil industry EAGE Publishing.
21. Nikravesh, M., Aminzadeh, F., Zadeh, L. A., 2003, Soft Computing and Intelligent Data Analysis, Vol. **51**, Developments in Petroleum Science Series, Elsevier, 724p.

Appendix

Additional figures and tables for all wells

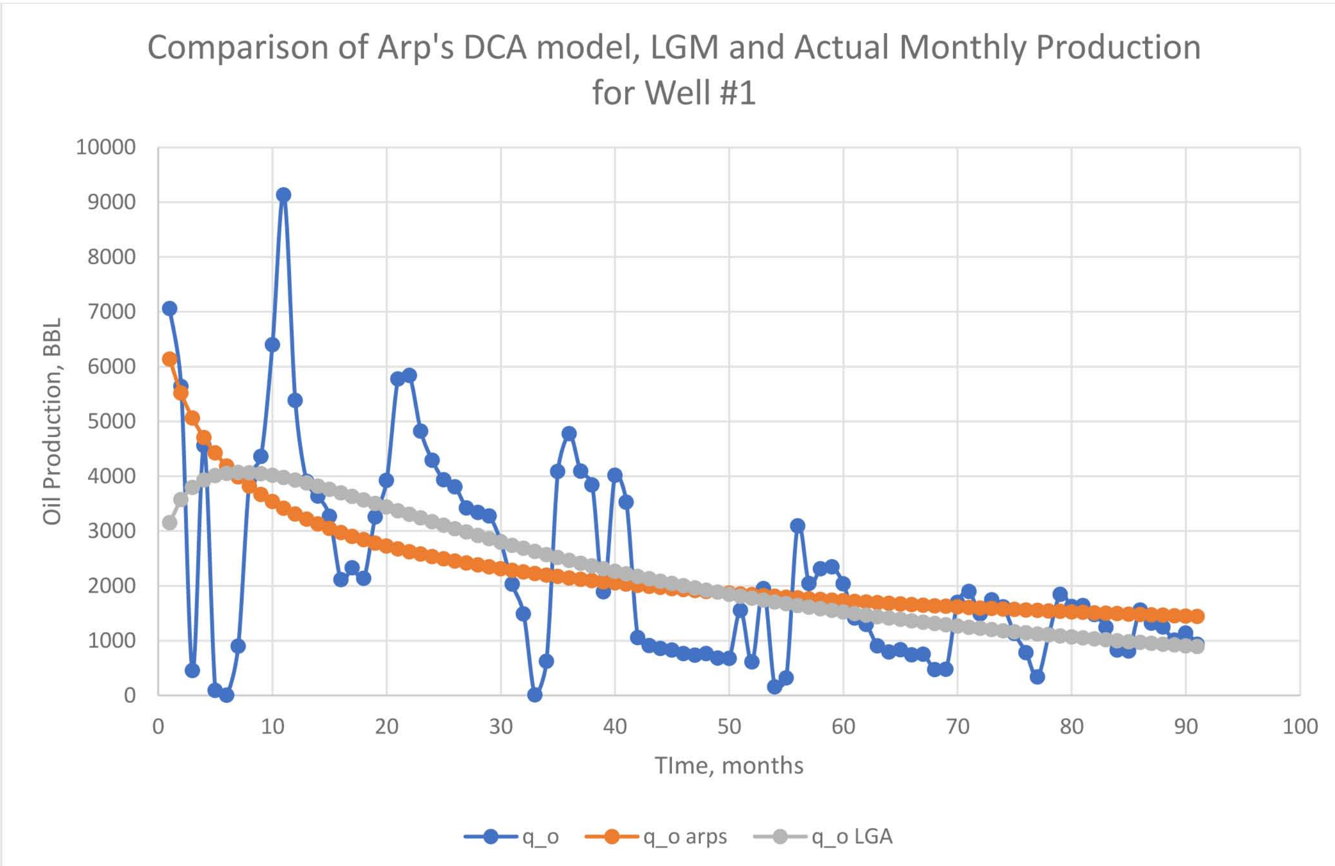

Figure 7—Actual production vs. Arps's DCA and Logistic Growth Methodfor Well #1

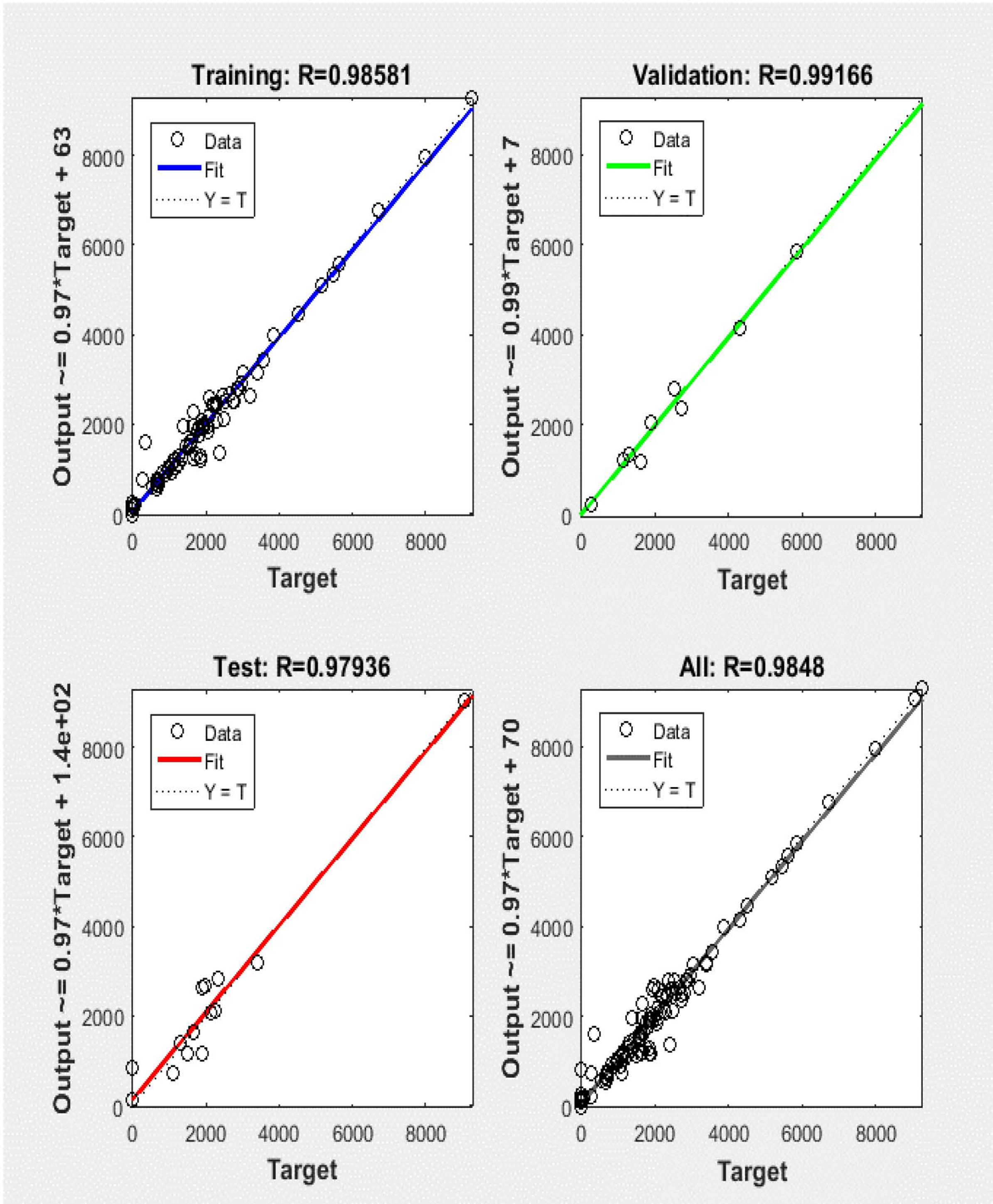

Figure 8—Regression Plot for Well #2

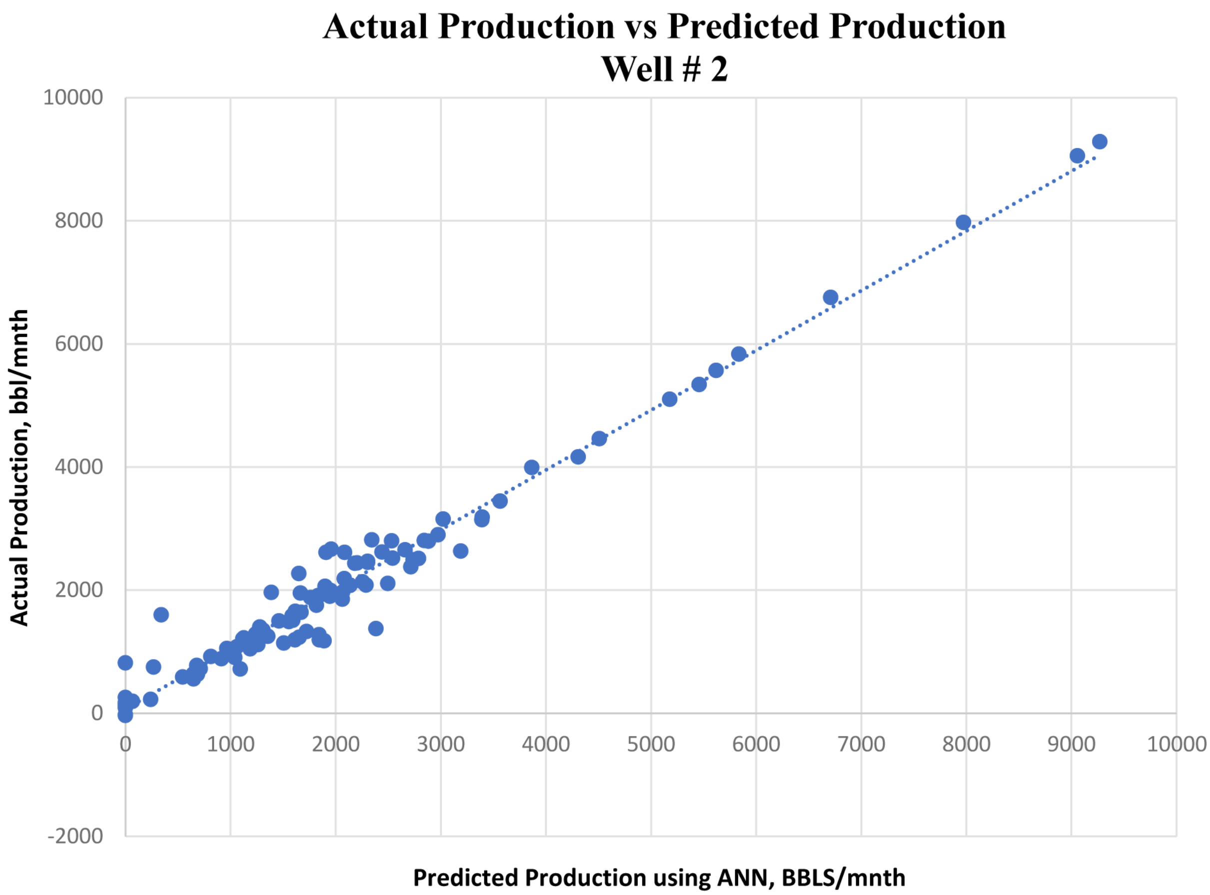

Figure 9—Actual production vs. ANN forecast for Well #2

Table 3—Cumulative Production Forecast error estimation comparison for Well#2

Month	Actual Cumulative Production	ANN Model Forecast	Duong's Method Forecast	ANN Forecast Error	Duong's Forecast Error
	bbl	bbl	bbl	%	%
98	209116	193046	203482	7.68	2.69
99	209802	195100	204308	7.01	2.62
100	210451	197753	205125	6.03	2.53
101	210995	201200	205934	4.64	2.40
102	211673	210483	206735	0.56	2.33
103	212366	213288	207528	-0.43	2.28

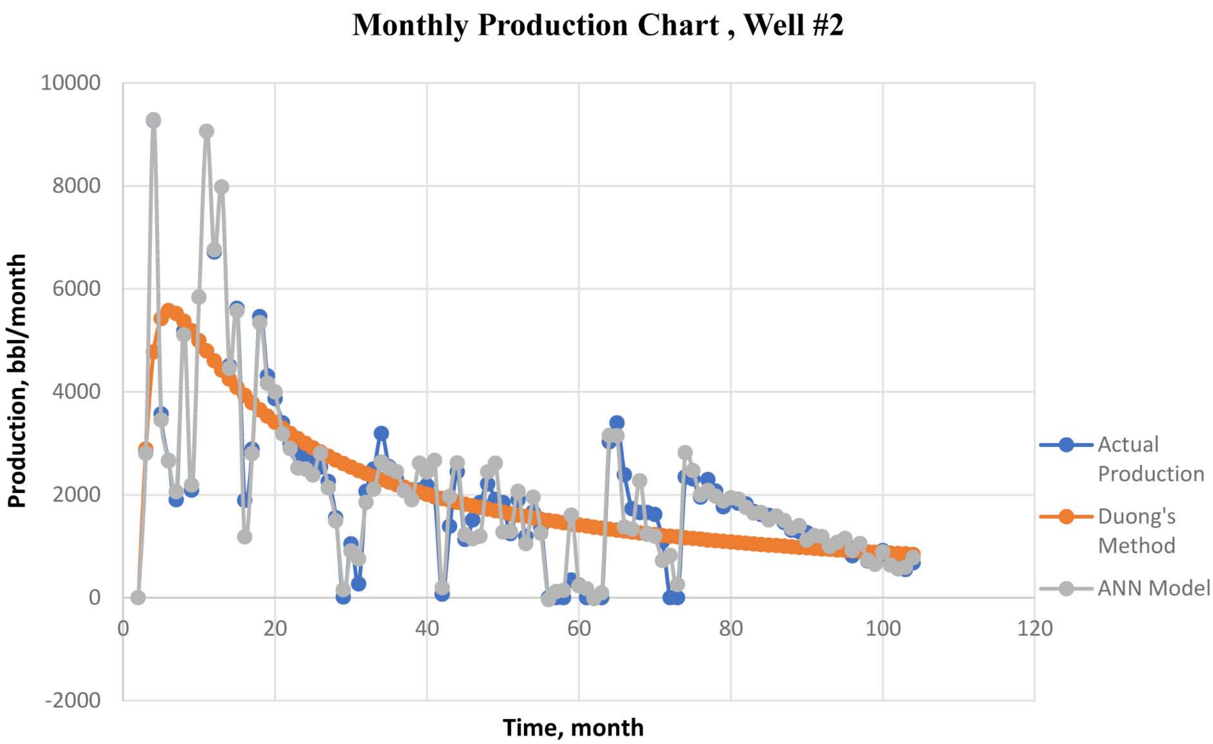

Figure 10—Comparison of ANN Forecast and Duong's Forecast to actual production

Table 4—Overview of Monthly Production Forecast comparison between ANN Model and Duong's Method for Well #2

Month	Actual Monthly Production	ANN Model Forecast	Duong's Method Forecast	ANN Forecast Error	Duong's Forecast Error
	bbl	bbl	bbl	%	%
96	713	725	890	1.69	24.79
97	650	642	881	1.31	35.55
98	914	887	873	2.93	4.53
99	686	631	864	8.03	25.98
100	649	557	856	14.21	31.90
101	544	590	848	8.48	55.89
102	678	776	840	14.48	23.91
			Mean of Error (%)	7.30	28.94
			SD of Error (%)	5.18	14.27

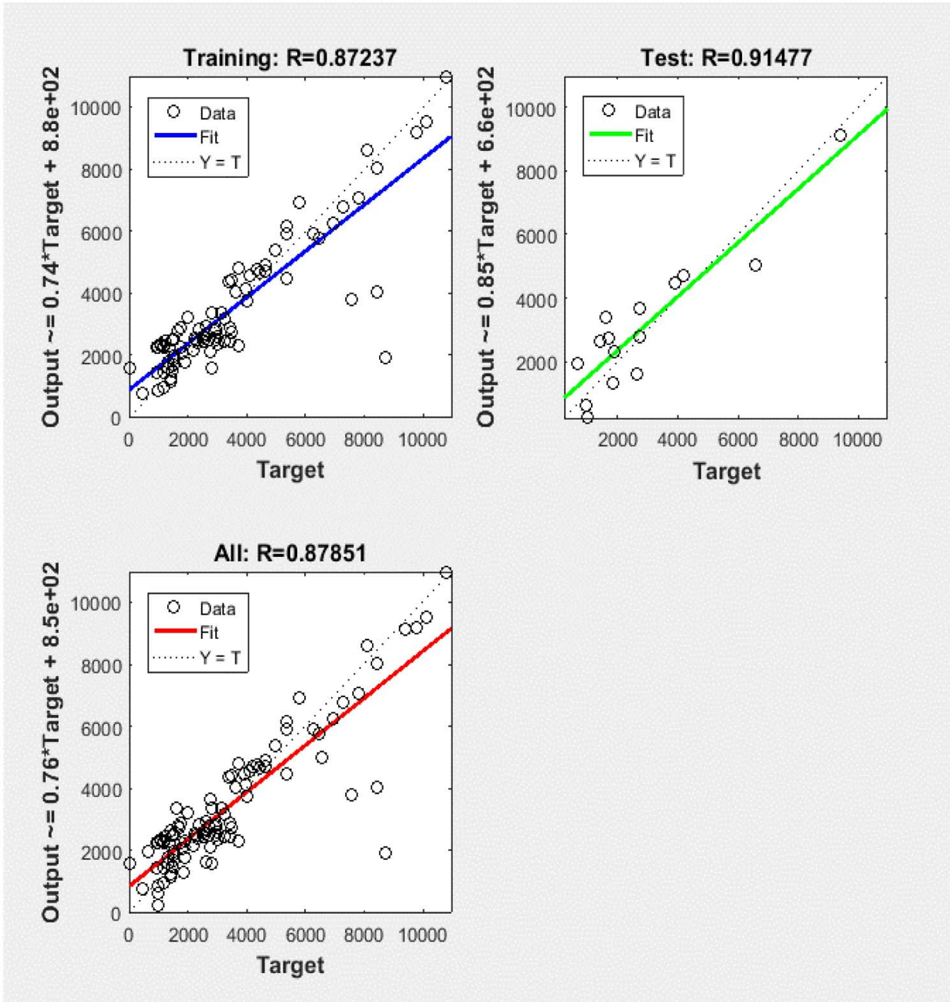

Figure 11—Regression Plot for Well #3

Table 5—Cumulative Production Forecast error estimation comparison for Well #3

Month	Actual Cumulative Production	ANN Model Forecast	Duong's Method Forecast	ANN Forecast Error	Duong's Forecast Error
	bbl	bbl	bbl	%	%
94	326181	311224.5	331316	4.032538	-1.5498
95	327184	313537.5	332816	2.890256	-1.6921
96	327650	317993.2	334302	2.109514	-1.98968
97	329056	320881	335774	0.58246	-2.00069
98	330466	327150.5	337233	-1.57125	-2.00653
99	331912	335741.3	338678	-1.15372	-1.99788

Table 6—Overview of Monthly Production Forecast comparison between ANN Model and Duong's Method for Well #3

Month	Actual Monthly Production	ANN Model Forecast	Duong's Method Forecast	ANN Forecast Error	Duong's Forecast Error
	bbl	bbl	bbl	%	%
94	1333	1565	1507	17.39	13.04
95	1003	883	1493	11.97	48.84
96	466	766	1479	64.32	217.40
97	1406	1135	1466	19.28	4.23
98	1410	1236	1452	12.31	3.00
99	1446	1479	1439	2.28	0.47
			Mean of Error (%)	21.26	47.83
			SD of Error (%)	21.91	84.99