SPE-223264-MS

Gas Wells Production Forecast Using Artificial and Recurrent Neural Networks Models and Comparison with Traditional Decline Curve Analysis Techniques

Ibrahim M. El-Ghobashy, Belayim Petroleum Company, Cairo, Egypt; Mohamed H. M. Sayyouh and M. Abu El Ela, Department of Mining, Petroleum and Metallurgical Engineering, Cairo University, Cairo, Egypt; Ahmed H. El-Banbi, Department of Petroleum and Energy Engineering, The American University in Cairo, Cairo, Egypt

This paper was prepared for presentation at the Mediterranean Offshore Conference held in Alexandria, Egypt, 20 - 22 October 2024.

Abstract

Accurate production forecasting is crucial for effective reservoir management and field development planning. Traditional methods; such as numerical simulation, material balance equations coupled with well models, and advanced decline curve analysis (DCA) require extensive data and involve tradeoffs between accuracy and complexity. Recently, deep learning techniques, specifically artificial neural network (ANN) and recurrent neural network (RNN) models have demonstrated promising performance in this domain. This study applies ANN and RNN for production forecasting and compares their effectiveness with traditional DCA models.

An actual production dataset for a gas field containing 13 wells is used to compare ANN and RNN models to traditional decline curve analysis models. After thorough data cleaning and preparation, the dataset was divided into training (historical) and test data sets. The training data was used to build ANN and RNN models in addition to estimating the model parameters for traditional DCA models. Traditional DCA models that were investigated include Arps (exponential, hyperbolic, and harmonic), Duong, Power Law Exponential Decline (PLE), and Stretched Exponential Decline (SEPD). Finally, the prediction of different models was evaluated and compared using the testing dataset. The comparison also included long, medium, and short-term prediction.

Based on the comparison of daily gas production rate forecast, the ANN model demonstrated notable improvements over Arps models for the majority of wells, with a mean absolute percent error (MAPE) of 3.9% compared to 9.1%. However, across all wells, the average MAPE was slightly lower at 6.3% versus 8.1%. Conversely, the RNN model exhibited varied performance compared to Arps models. It showed poorer long-term performance (13% versus 4.5%), similar performance in the medium term (4.4% versus 4.2%), and superior performance in the short term (2.7% versus 4.2%). Furthermore, among DCA models specifically developed for unconventional wells, SEPD exhibited nearly identical performance to Arps (8.3% compared to 8.1%). In contrast, Duong model demonstrated the poorest performance with MAPE of 21.9%.

This study provides insights into model analysis and introduces guidelines in using artificial and recurrent neural networks for predicting production rates. The ANN and RNN models were developed through careful testing with different designs and techniques. All the models in this research were created using Python scripts and powerful open-source libraries like SciPy and TensorFlow.

Introduction

Effective well production forecasting is critical in the oil and gas industry as it enables accurate prediction of future production rates, aiding in optimal resource allocation and informed decision-making for field development. This forecasting supports the scheduling of maintenance, assessment of economic viability, and estimation of reserves, all of which are essential for both immediate operational needs and long-term strategic planning. Furthermore, reliable forecasting reduces uncertainties, mitigates financial risks, and improves overall project profitability by ensuring efficient and effective use of resources.

Numerical simulators, while comprehensive, often require substantial computational resources and detailed reservoir data. This makes them less practical for quick decision-making in dynamic environments. Decline curve analysis, although useful for predicting future production, assume that past trends will continue unchanged, which can lead to inaccuracies if reservoir conditions and operating strategies change (Arps 1945). The material balance equation, though effective in estimating original hydrocarbons in place, relies heavily on the accuracy of reservoir pressure data and may not adequately capture complex reservoir behaviors, such as those involving multiple phases or heterogeneous formations (Craft and Hawkins 1959). To use material balance models in forecasting production rates, they need to be coupled with flow equations or well models (Sallam and El-Banbi, 2018; and Farid et al. 2013).

While Arps' decline curve analysis model is still commonly used in the industry, its effectiveness is confined to conventional reservoirs due to its oversimplified assumptions about production behavior. To address these limitations, several advanced models have been developed to better handle the complexities of unconventional resources. The Power Law Exponential (PLE) model accommodates the rapid initial decline and subsequent stabilization typical of unconventional wells by incorporating variable decline rates (Ilk, et al. 2009). The Stretched Exponential Production Decline (SEPD) model offers a refined approach to capture the prolonged and intricate decline patterns in tight gas reservoirs (Valkó and Lee 2010). Additionally, the Duong model is designed specifically for tight sand/shale gas wells, integrating both early linear flow and extended boundary-dominated decline (Duong 2010).

The application of deep learning techniques, particularly Artificial Neural Network (ANN) and Recurrent Neural Network (RNN), has gained significant traction in the field of oil and gas well forecasting (Cao, et al. 2016, Suhag, Rahul and Aminzadeh 2017). These models are favored for their advanced ability to identify and interpret complex, nonlinear patterns especially for unconventional resources. Several studies have underscored the potential of ANN and RNN models to provide highly accurate forecasts (Sun and Kazi 2018, Zhan et al 2019, Almohammadi et al 2020). However, the effectiveness of these techniques remains a topic of debate. While some research has reported outstanding results, highlighting the models' capability to deliver precise predictions, other studies reported inconsistent or suboptimal performance, raising questions about their reliability (Khan and Louis 2021, Kocoglu, Gorell and McElroy 2021, Jayeola and Olusola 2022). This highlights the considerable potential of deep learning techniques in overcoming forecasting challenges within the oil and gas sector. Nonetheless, it also points to the importance of further research to fully grasp the limitations and the specific situations where these techniques excel. Achieving this understanding will require thorough comparisons with traditional forecasting methods, which can reveal both the strengths and weaknesses of deep learning approaches and help determine their optimal areas of application (Rahmanifard, Gates and Shabib 2022).

Deep Learning

Machine learning, a key area of artificial intelligence, is centered on developing algorithms and statistical models that allow computers to learn from data and make predictions autonomously. When applied to time-series analysis and forecasting, machine-learning methods are used to detect patterns, trends, and relationships in sequential data to forecast future values. This process involves training models on historical data to uncover temporal patterns and use these insights for predicting future events (Géron 2019). An advanced subset of this field is deep learning, which utilizes complex neural network with multiple layers to capture and model intricate patterns in data. This approach significantly enhances the ability to handle complex, nonlinear relationships and improves the accuracy and performance of predictive models, making it a powerful tool for tackling challenging forecasting problems (Brownlee 2018). This study examines two primary deep learning techniques: standard artificial neural network (ANN) and recurrent neural network (RNN). These methods have gained prominence in recent research for their application in production forecasting.

Artificial Neural Network

Artificial Neural Network (ANN) are sophisticated computational models designed to replicate the brain's neural processing capabilities, structured into layers of interconnected nodes, or neurons. As depicted in (Figure 1), ANNs typically consist of three main types of layers: the input layer, which processes raw data; one or more hidden layers, where data is transformed through weighted connections and activation functions; and the output layer, which generates final predictions (LeCun, Bengio and Hinton 2015). The training process of an Artificial Neural Network (ANN) primarily consists of two fundamental steps: forward propagation and backpropagation. In forward propagation, data passes through the network's layers sequentially, with each layer applying transformations based on the weights and biases it has acquired, ultimately producing the final output. Then, during backpropagation, the network computes the gradients of the loss function with respect to each weight, making adjustments to minimize errors and enhance the overall accuracy of the model. The performance of ANNs is significantly influenced by hyper-parameters such as the learning rate (η), the number of hidden layers, and the number of neurons per layer (Goodfellow, Bengio and Courville 2016, Rumelhart and Hinton 1986). These hyper-parameters are crucial in controlling the model's learning process and convergence. For instance, the learning rate affects the size of weight updates during training, affecting the speed and stability of convergence (Kingma and Ba 2015). The output of each layer in an ANN is generally described by:

$$h^l = f\left(W^l h^{l-1} + b^l\right) \tag{1}$$

Where h^l represents the vector of activations in layer l, W^l denotes the weight matrix connecting layer $l - 1$ to layer l, h^{l-1} is the vector of activations from the previous layer $l - 1$ and b^l is the bias vector for layer l. The activation function f introduces non-linearity to the model. Proper tuning of these hyper-parameters is essential for optimizing model performance and enhancing training efficiency, which in turn affects the accuracy and applicability of the network in various real-world tasks (Hutter, Kotthoff and Vanschoren 2012).

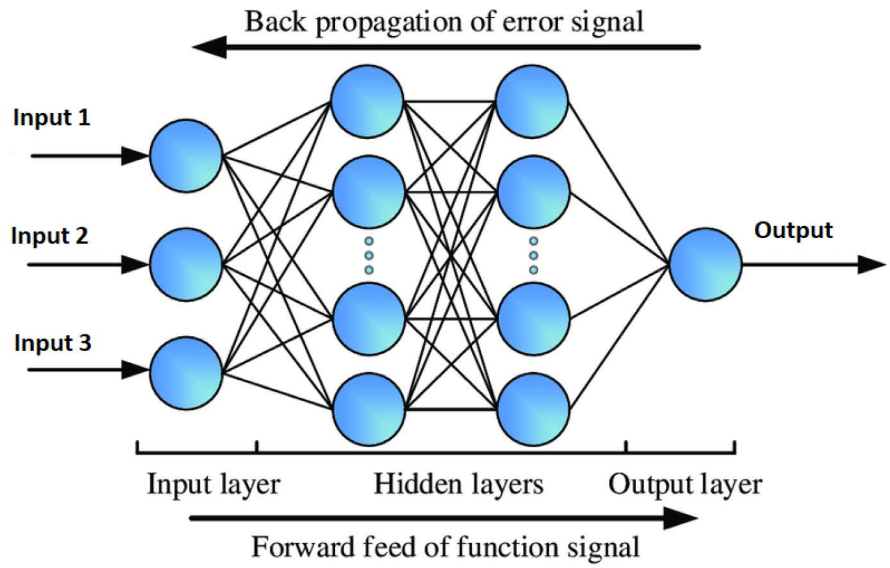

Figure 1—Graphical representation of artificial neural network

Recurrent Neural Network

Recurrent Neural Networks (RNNs) are designed to handle sequential data by utilizing feedback loops that enable the retention of temporal dependencies. Unlike traditional feedforward neural networks, RNNs incorporate recurrent connections to preserve contextual information over time. As (Figure 2) illustrates, the hidden state at each time step t is updated based on the previous hidden state h_{t-1} and the current input x_t, described by the general equation for each hidden layer:

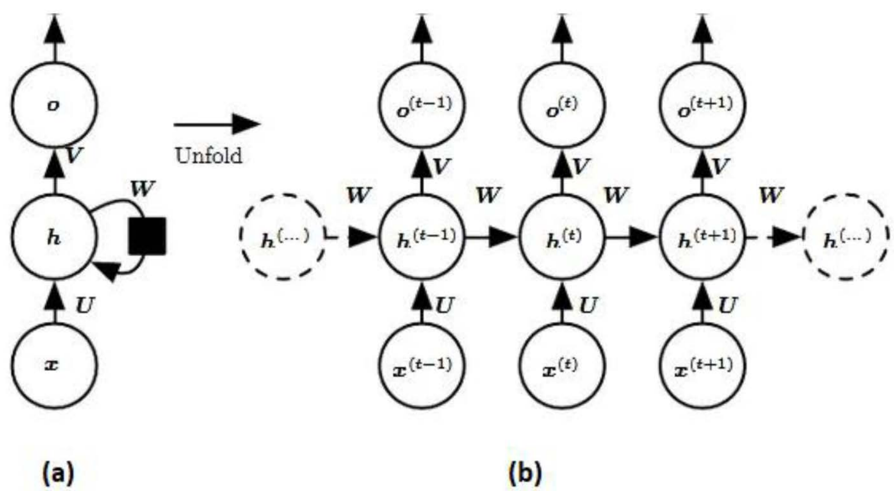

Figure 2—Graphical representation of recurrent neural networks. Fig. 2a shows the RNN schematic folded and Fig. 2b shows the unfolded version. The figure illustrates the hidden state layer between input and output layers.

$$h_t = f\left(W_h h_{t-1} + W_x x_t + b_h\right) \tag{2}$$

Where h_t represents the hidden state at time t, W_h and W_x are weight matrices for the hidden state and the input, respectively, and b_h is the bias term. For the output at each time step, the general equation applied is:

$$y_t = W_y h_t + b_y \tag{3}$$

Here, y_t is the output, W_y is the weight matrix for the output layer, and b_y is the bias term. These equations facilitate RNNs in learning and making predictions based on sequential input data.

These fundamental structures enable RNNs to process and predict sequential data effectively. However, they are often hindered by issues like vanishing and exploding gradients, which necessitated the development of advanced architectures such as Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) networks.

Long Short-Term Memory

Long Short-Term Memory (LSTM) networks, presented with (Figure 3), are an advanced form of Recurrent Neural Networks (RNNs) designed to address the limitations of conventional RNNs, especially their difficulties in managing long-term dependencies. LSTMs overcome these challenges through a specialized architecture that includes memory cells and gating mechanisms. Each LSTM unit comprises three key components: the forget gate, which determines what information should be discarded from the memory cell; the input gate, which regulates the integration of new data into the memory cell; and the output gate, which generates the final output based on the cell's current state. This advanced design allows LSTMs to effectively capture and utilize information across long sequences, making them particularly advantageous for tasks such as time-series forecasting and sequence-based applications. Their ability to model complex temporal patterns has led to widespread use in various fields, including finance for stock price prediction, speech processing for speech-to-text systems, and natural language processing for machine translation (Schmidhuber 2015, Hochreiter and schmidhuber 1997, Graves, Mohamed and Hinton 2013, Shenfield and Howarth 2020).

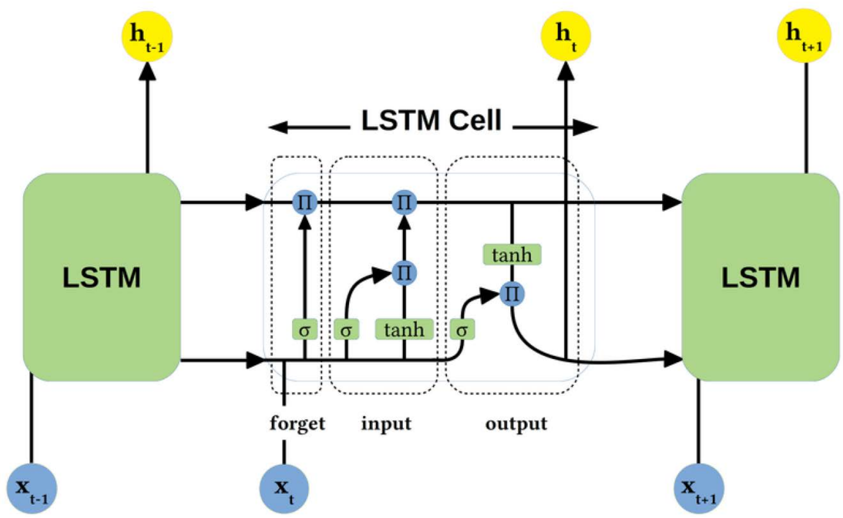

Figure 3—Graphical representation of an LSTM Cell shows the main three gates

Methodology

In this study, data from 13 gas wells is used. The data is collected and prepared before the development of neural network models. Two types of neural networks—Artificial Neural Networks (ANN) and Recurrent Neural Networks (RNN)—are employed. Following their development, the models are validated and tested by comparing their predictions with the actual data, and their performance is assessed against other decline curve analysis (DCA) methods. Most of the models in this research are implemented using Python, utilizing popular libraries such as NumPy, SciPy, and TensorFlow. The study methodology is illustrated in (Figure 4).

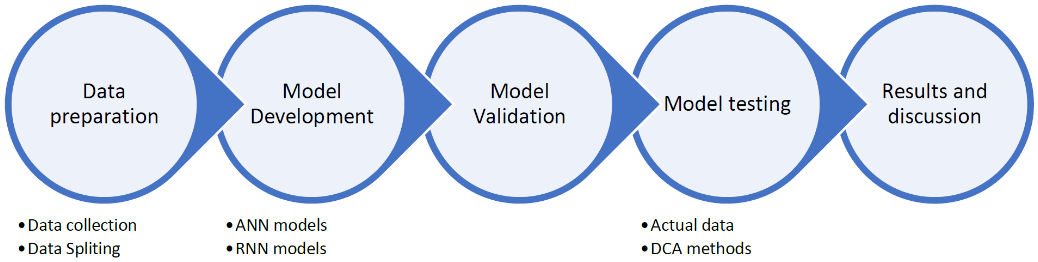

Figure 4—Methodology for ANN and RNN models construction and comparison with DCA models

Data Preparation

This study includes production performance data from 13 gas wells located in a dry gas field in the Mediterranean, each with an average production duration of approximately 1,800 days. Following data collection, the dataset undergoes rigorous quality checks and validation to remove outliers before being divided into (training + validation) and testing subsets. The validation dataset is randomly selected from the first dataset to ensure an unbiased assessment of model performance. It is important to note that the final portion of the (training + validation) dataset holds significant value, as all decline curve analysis (DCA) techniques rely exclusively on this segment for regression. Consequently, excluding this part from training could negatively impact the model's performance. The data splitting process is depicted in (Figure 5).

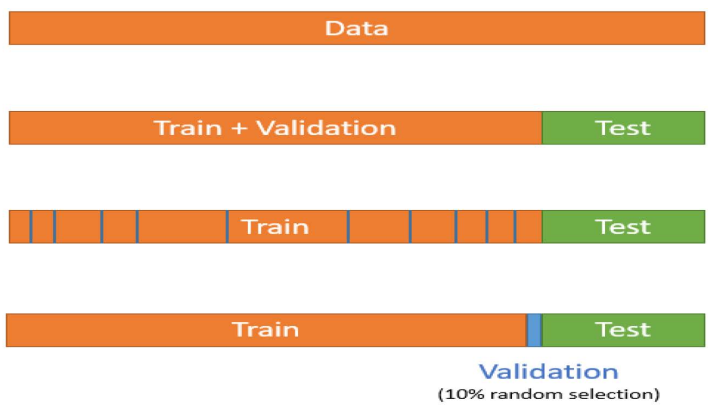

Figure 5—Data splitting process

Model Inputs and Outputs

The input matrix configuration varies according to the deep learning algorithm applied. For ANN models, extensive testing identified that the most effective structure involves using time as the only input node and gas rate as the output node. Introducing additional parameters like WHP and BHP was observed to decrease the model's accuracy. On the other hand, the RNN model takes a 100-day sequence of gas rates as input to forecast the subsequent sequence. The output sequence length is adjusted based on the forecast horizon, with 30 days for short-term, 100 days for medium-term, and 200 days for long-term predictions.

Model Development and Validation

The development and validation process for both ANN and RNN models follows a similar approach, with specific adaptations depending on the model type. After dividing the data into training, validation, and testing datasets, the models are constructed by exploring a range of hyper-parameters. The model's performance on the validation dataset is assessed using the Root Mean Square Error RMSE (Eq. 4) to determine the optimal set of hyper-parameters, including learning rate, number of nodes per layer, and dropout ratio. For ANN models, a single layer is used, with the number of nodes per layer varying between 10 and 90, the dropout ratio set to 0.1 or 0.2, the learning rate chosen from 1×10^{-5}or 1×10^{-6} and RELU as activation function. In contrast, the RNN model may consist of one or two layers, with the number of nodes per layer ranging between 32, 64, and 128, the dropout ratio between 0.1 and 0.2 and the activation function used is LSTM.

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2\right)^{0.5} \tag{4}$$

Model Testing

After the model development and validation phases, the model is tested using actual data and evaluated with the Mean Absolute Percentage Error (MAPE) (Eq. 5). DCA methods are then employed to predict performance over the testing period, and their accuracy is assessed using MAPE. For the RNN model, long-term predictions (~200 days) proved inadequate, so the testing period is segmented into two distinct ranges: the first 100 days are used for medium-term predictions, and the first 30 days are chosen for short-term predictions. The model's performance is then evaluated for these specific periods.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \bar{y}}{y_i}\right| \times 100 \tag{5}$$

The comparison involves seven decline curve analysis (DCA) models. Four of these models are based on Arps' methodology: the first model represents exponential decline, the second model is for hyperbolic decline with b = 0.5, the third model addresses harmonic decline, and the fourth model adjusts the b value to optimally fit the regression period. Additionally, the comparison includes models designed for unconventional resources, such as the Duong model, the Stretched Exponential Production Decline (SEPD) model, and the Power Law Exponential (PLE) model.

Results and Discussion

Figures 6 through 8 present a comparative analysis of the various models utilized in the study for a representative well. (Figure 6) displays a comparison between the ANN model and the four Arps models. (Figure 7) illustrates the performance of the RNN model across long, medium, and short-term forecasting periods. (Figure 8) provides a comparison between the Arps model (best-fit b) and other DCA models.

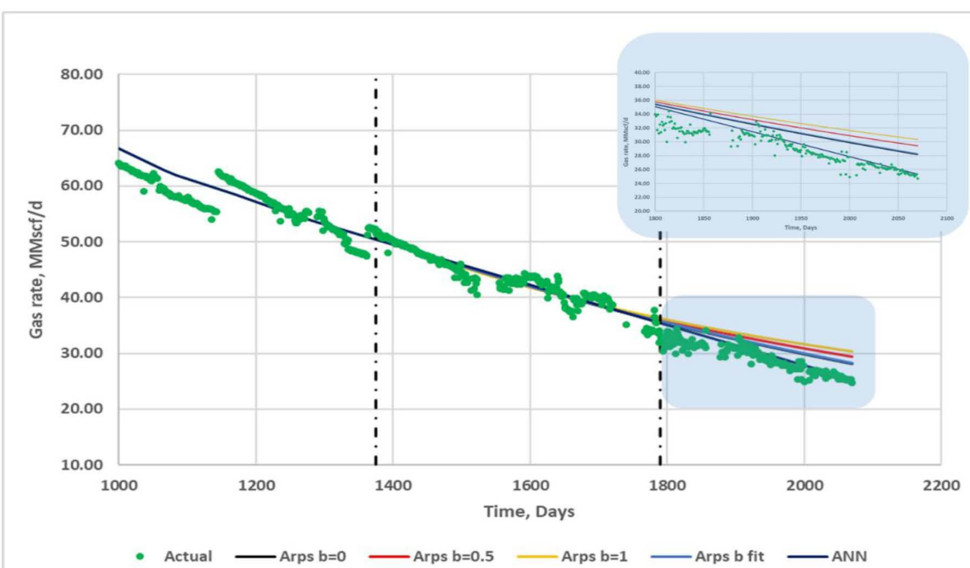

Figure 6—Production forecast with ANN and Arps models

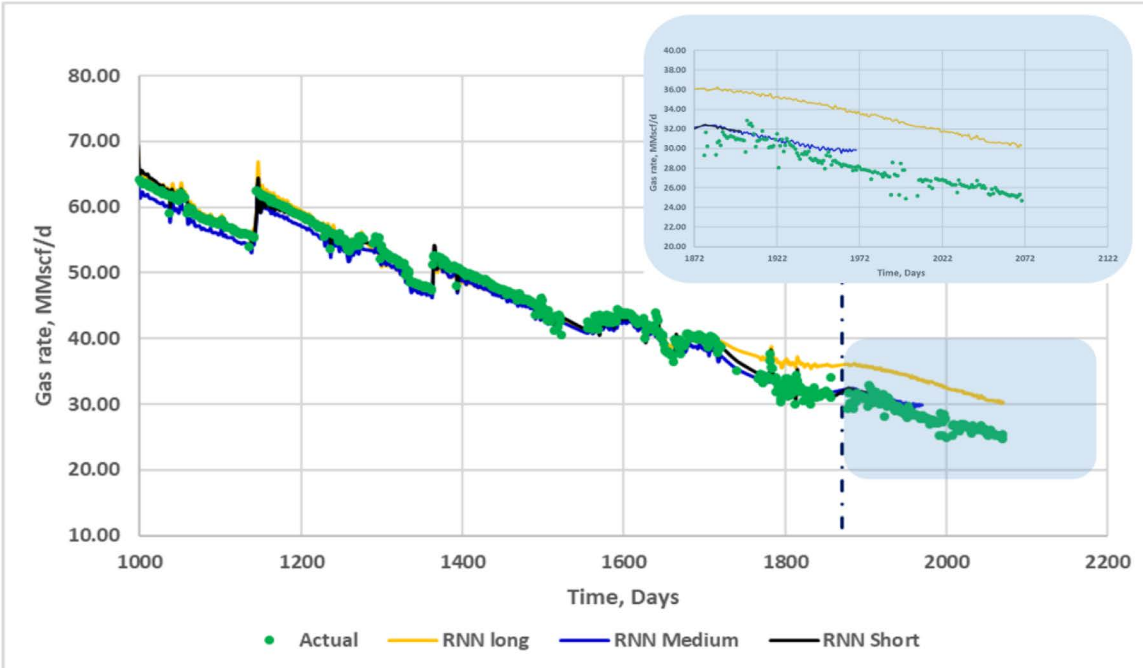

Figure 7—Production forecast with RNN for long, medium and short term

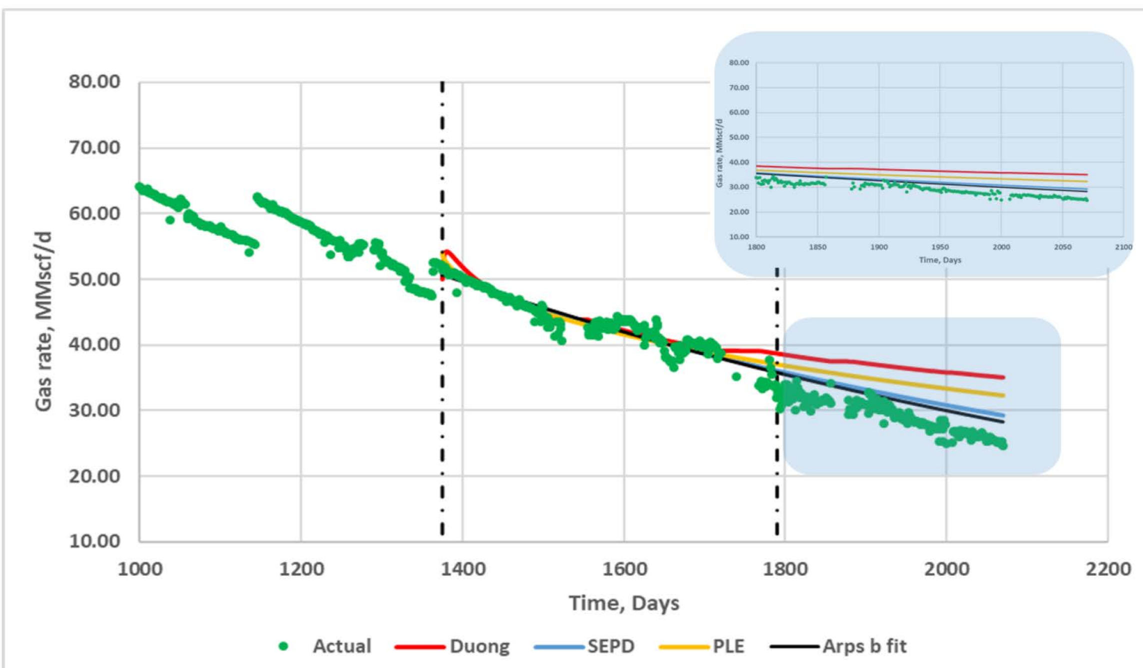

Figure 8—Production forecast with DCA models Duong, SEPD, PLE and Arps (best-fit b)

Based on the figures presented, Figure 6 demonstrates that the ANN model outperforms the Arps models. Figure 7 reveals that the RNN model performs poorly in long-term forecasts, is acceptable for medium-term predictions, and shows strong performance for short-term forecasts. In Figure 8, the SEPD model exhibits performance comparable to the Arps model, while the Duong model shows inferior performance. Table 1 provides a summary of the MAPE comparisons between the ANN model and all DCA models, focusing on daily production rates for all wells. Table 2 extends this comparison to cumulative production of the forecasted part of data. Table 3 details the MAPE comparison between the RNN model and the Arps model (best-fit b) across long, medium, and short-term production rates, while Table 4 presents the same comparison for cumulative production of forecasted part of data.

Table 1—Comparison of the predicted daily production rate MAPE for DCA and ANN models

Well	Arps b=0	Arps b=0.5	Arps b=1	Arps b fit	Duong	SEPD	PLE	ANN
Well-01*	32.3	39.9	46.3	33.9	65.7	34.3	48.9	39.5
Well-02	8.7	11.5	13.5	9.1	26.2	11.3	18.5	3.9
Well-03	3.8	2.8	4.8	4.8	20.4	4.2	8.4	3.6
Well-04	3.7	3.5	6.7	6.7	29.1	5.6	10.9	4.3
Well-05	11.0	5.8	6.6	9.7	27.5	7.0	6.8	6.8
Well-06	8.7	11.3	13.7	13.7	20.3	14.7	23.5	9.3
Well-07	7.8	9.4	10.7	8.3	17.6	8.0	15.2	5.6
Well-08	10.0	6.9	4.4	4.4	11.2	5.3	3.4	3.8
Well-09	16.6	18.9	20.5	16.9	37.4	17.0	31.4	14.5
Well-10	8.1	4.3	4.1	8.1	29.4	7.2	13.8	5.1
Well-11	11.4	5.8	3.3	3.8	16.9	6.7	3.5	3.9
Well-12	3.5	6.0	8.6	8.6	23.8	8.5	13.7	4.5
Well-13	1.8	2.4	3.4	3.4	3.2	3.7	5.6	9.8
Average	**7.9**	**7.4**	**8.4**	**8.1**	**21.9**	**8.3**	**12.9**	**6.3**

Table 2—Comparison of the predicted cumulative production of the forecasted part MAPE for DCA and ANN models

Well	Arps b=0	Arps b=0.5	Arps b=1	Arps b fit	Duong	SEPD	PLE	ANN
Well-01*	30.7	37.8	43.7	32.2	61.9	32.5	46.2	37.9
Well-02	8.5	11.2	13.2	8.9	25.6	11.1	18.0	3.6
Well-03	1.2	1.8	4.4	4.4	19.9	3.7	8.1	1.0
Well-04	0.8	3.2	6.7	6.7	28.8	5.6	10.8	0.9
Well-05	6.6	0.6	4.7	5.2	26.3	2.1	5.2	0.9
Well-06	7.4	10.9	13.3	13.3	20.0	14.3	23.1	8.6
Well-07	7.6	9.2	10.4	8.1	17.2	7.8	14.8	5.4
Well-08	10.1	7.1	4.3	4.3	10.8	5.4	2.0	3.8
Well-09	15.2	17.4	18.9	15.5	35.2	15.6	29.4	4.9
Well-10	6.6	2.7	1.9	6.6	27.6	5.9	12.5	4.5
Well-11	11.1	5.2	0.0	2.0	16.7	6.1	0.1	1.8
Well-12	1.2	4.4	7.1	7.1	22.4	7.1	12.4	0.8
Well-13	0.8	1.8	2.9	2.9	1.8	3.4	5.3	9.6
Average	**6.4**	**6.3**	**7.3**	**7.1**	**21.0**	**7.3**	**11.8**	**3.8**

Table 3—Comparison of the predicted daily production rate MAPE for RNN and Arps models for long, medium and short term.

Well	Long		Medium		Short	
	RNN	Arps	RNN	Arps	RNN	Arps
Well-01*	53.3	25.2	11.0	16.0	4.1	15.0
Well-02	18.8	6.9	3.6	4.6	3.9	4.2
Well-03	6.1	2.8	5.0	2.8	3.6	2.9
Well-04	7.8	2.5	1.9	2.8	2.5	3.7
Well-05*	2.6	6.7	5.0	8.2	10.0	7.9
Well-06*	12.7	21.1	3.1	19.2	3.6	17.6
Well-07	13.2	4.2	2.3	3.0	1.9	2.2
Well-08	14.0	6.3	3.4	7.6	1.8	8.1
Well-09*	4.5	14.7	9.9	17.0	25.5	26.5
Well-10	14.5	2.5	1.8	2.5	2.3	2.1
Well-11	18.5	7.8	10.1	7.4	3.1	8.7
Well-12	11.1	2.7	7.5	2.6	2.4	2.1
Well-13*		7.6	12.8	3.6	1.4	1.0
Average	**13.0**	**4.5**	**4.4**	**4.2**	**2.7**	**4.2**

Table 4—Comparison of the predicted cumulative production of the forecasted part MAPE for RNN and Arps models for long, medium and short term.

Well	Long		Short		Medium	
	RNN	Arps	RNN	Arps	RNN	Arps
Well-01*	53.3	25.2	11.0	16.0	4.1	15.0
Well-02	18.8	6.9	3.6	4.6	3.9	4.2
Well-03	6.1	2.8	5.0	2.8	3.6	2.9
Well-04	7.8	2.5	1.9	2.8	2.5	3.7
Well-05*	2.6	6.7	5.0	8.2	10.0	7.9
Well-06*	12.7	21.1	3.1	19.2	3.6	17.6
Well-07	13.2	4.2	2.3	3.0	1.9	2.2
Well-08	14.0	6.3	3.4	7.6	1.8	8.1
Well-09*	4.5	14.7	9.9	17.0	25.5	26.5
Well-10	14.5	2.5	1.8	2.5	2.3	2.1
Well-11	18.5	7.8	10.1	7.4	3.1	8.7
Well-12	11.1	2.7	7.5	2.6	2.4	2.1
Well-13*		7.6	12.8	3.6	1.4	1.0
Average	**12.9**	**1.1**	**4.1**	**0.9**	**2.0**	**4.8**

Conclusions

Based on a comparison study between old and modern production forecast methods for a volumetric conventional gas reservoir, several insights were derived. It is shown in this study that the comparative analysis of forecasting methods for gas well production underscores the continued relevance of Arps models alongside more advanced techniques. Although Arps models, despite their long-standing use, remain effective for predictions, the study demonstrates that ANN models provide superior accuracy for long-term forecasts, extending up to 300 days. Conversely, RNN models excel in medium-term (~100 days) and short-term (~30 days) predictions. Modified DCA methods, including the Stretched Exponential Production Decline (SEPD) model, offer performance comparable to Arps models for conventional resources. However, the Duong method is less reliable due to higher error rates. Thus, while Arps models continue to be a valuable tool, incorporating ANN and RNN can significantly enhance forecasting capabilities across various time frames.

Nomenclature

ANN	Artificial Neural Network
DCA	Decline Curve Analysis
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
PLE	Power Law Exponential
RELU	Rectified Linear Unit
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SEPD	Stretched Exponential Production Decline

Symbols

h^l	Vector of activations in layer l
W^l	Weight matrix connecting layer $l-1$ to layer l
h^{l-1}	Vector of activations from the previous layer $l-1$
b^l	Bias vector for layer l
h_t	Hidden state at time t,
W_h	Weight matrices for the hidden state
W_x	Weight matrices for the input

References

Alimohammadi, Hamzeh, Hamid Rahmanifard, and Nancy Chen. 2020. "Multivariate Time Series Modelling Approach for Production Forecasting inUnconventional Resources." SPE Annual Technical Conference & Exhibition. Denver, Colorado, USA. SPE-201571-MS. doi:10.2118/201571-MS.

Arps, J.J. 1945. "Analysis of Decline Curves." *Transactions of the AIME* (SPE-945228-G) 228–247. doi:10.2118/945228-G.

Brownlee, Jason. 2018. *Better Deep Learning. Machine Learning Mastery*.

Cao, Q, R Banerjee, S Gupta, J Li, W Zhou, and B Jeyachandra. 2016. "Data Driven Production Forecasting Using Machine Learning." SPE Argentina Exploration and Production of Unconventional Resources Symposium. Buenos Aires, Argentina: SPE-180984-MS. doi:10.2118/180984-MS.

Craft, Benjamin Cole, and Murray Franklin Hawkins. 1959. *Applied Petroleum Reservoir Engineering*. Englewood Cliffs: Prentice-Hall.

Duong, A N. 2010. "An Unconventional Rate Decline Approach for Tight and Fracture-Dominated Gas Wells." Canadian Unconventional Resources and International Petroleum Conference. Calgary, Alberta, Canada: SPE-137748-MS. doi:10.2118/137748-MS.

Farid, A.M., El-Banbi, A.H., and Abdelwaly, A.A. (2013). "An Integrated Model for History Matching and Predicting Reservoir Performance of Gas/Condensate Wells." *SPE Res Eval & Eng* **16** (4): 412–422. SPE-151869-PA. http://dx.doi.org/10.2118/151869-PA.

Géron, Aurélien. 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press.

Graves, Alex, Abdel-rahman Mohamed, and Geoffrey Hinton. 2013. "Speech recognition with deep recurrent neural networks." IEEE International Conference on Acoustics, Speech and Signal Processing. 6645–6649.

Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation*.

Hutter, Frank, Lars Kotthoff, and Joaquin Vanschoren. 2012. *Practical Recommendations for Gradient-Based Training of Deep Architectures*. Springer.

Ilk, D, N L Johnson, S M Currie, and T A Blasingame. 2009. "A Simple Methodology for Direct Estimation of Gas-in-Place and Reserves Using Rate-Time Data." SPE Rocky Mountain Petroleum Technology Conference. Denver: SPE-123298-MS. doi:10.2118/123298-MS.

Jayeola, Ifeoluwa, and Bukola Olusola. 2022. "Machine Learning Prediction Versus Decline Curve Prediction: A Niger DeltaCase Study." SPE Nigeria Annual International Conference and Exhibition. Lagos, Nigeria: SPE-211956-MS. doi:10.2118/211956-MS.

Khan, Hassan, and Clifford Louis. 2021. "An Artificial Intelligence Neural Networks Driven Approach to Forecast Production in Unconventional Reservoirs – Comparative Analysis withDecline Curve." International Petroleum Technology Conference. IPTC-21350-MS. doi:10.2523/IPTC-21350-MS.

Kingma, Diederik, and Jimmy Ba. 2015. "Adam: A Method for Stochastic Optimization." International Conference on Learning Representations.

Kocoglu, Y, S Gorell, and P McElroy. 2021. "Application of Bayesian Optimized Deep Bi-LSTM Neural Networks for Production Forecasting of Gas Wells in Unconventional Shale Gas Reservoirs." Unconventional Resources Technology Conference. Houston, Texas, USA: URTEC-2021-5418-MS. doi:10.15530/urtec-2021-5418.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep learning." *Nature* 436–444. doi:10.1038/nature14539.

Rahmanifard, Hamid, Ian Gates, and Abdolmohsen Shabib. 2022. "Comparison of Machine Learning and Statistical Predictive Models for Production Time Series Forecasting in Tight Oil Reservoirs." Unconventional Resources Technology Conference. Houston, Texas, USA: URTEC-3703284-MS. doi:10.15530/urtec-2022-3703284.

Rumelhart, David E, and Geoffrey E Hinton. 1986. "Learning representations by back-propagating errors." *Nature* 533–536. doi:10.1038/323533a0.

Sallam, M. and El-Banbi, Ahmed H. 2018. "Analysis of Multi-Layered Commingled and Compartmentalized Gas Reservoirs." *J. Petrol Explor Prod Technol*, Springer, (March 2018), Vol. **8**, pp. 1573–1586. https://doi.org/10.1007/s13202-018-0454-3.

Schmidhuber, Juergen. 2015. "Deep Learning in Neural Networks: An Overview." *Neural Networks* 85–117. doi:10.1016/j.neunet.2014.09.003.

Shenfield, Alex, and Martin Howarth. 2020. "A Novel Deep Learning Model for the Detection and Identification of Rolling Element-Bearing Faults." *Sensors*. doi:10.3390/s20185112.

Suhag, Anuj, Rahul Rahul, and Fred Aminzadeh. 2017. "Comparison of Shale Oil Production Forecasting using Empirical Methods." SPE Annual Technical Conference and Exhibition. San Antonio, Texas, USA: SPE-187112-MS. doi:10.2118/187112-MS.

Sun, J, X Ma, and M Kazi. 2018. "Comparison of Decline Curve Analysis DCA with Recursive Neural Networks RNN for Production Forecast of Multiple Wells." SPE Western Regional Meeting. Garden Grove, California, USA: SPE-190104-MS. doi:10.2118/190104-MS.

Valkó, Peter P, and W John Lee. 2010. "A Better Way to Forecast Production from Unconventional Gas Wells." SPE Annual Technical Conference and Exhibition. Florence: SPE-134231-MS. doi:10.2118/134231-MS.

Zhan, Cheng, Sathish Sankaran, Vincent LeMoine, Jeremy Graybill, and Didi-Ooi Sher Mey. 2019. "Resources, Application of Machine Learning for Production Forecasting for Unconventional." Unconventional Resources Technology Conference. Denver, Colorado, USA: URTEC-2019-47-MS. doi:10.15530/urtec-2019-47.