Research Article
Open Access

Application of gated recurrent units for time series prediction of gas production rates

Abstract

Accurate forecasting of gas production rates is essential for effective reservoir management and early detection of flow assurance issues such as liquid loading. This study explores the application of a Gated Recurrent Unit (GRU)-based deep learning model to predict daily gas production rates in a gas condensate well using a univariate time series approach. To improve prediction accuracy and minimize the impact on the results, a dataset spanning 1,500 days was utilized. The dataset was divided into training and testing subsets, comprising 1,200 days and 300 days, respectively. Gas rate data were normalized prior to model training. The GRU model was trained using a sliding window of 10-day sequences and evaluated based on standard regression metrics. The model achieved a Mean Squared Error (MSE) of 78.58, a Mean Absolute Error (MAE) of 3.35, and a Mean Absolute Percentage Error (MAPE) of 3.98%, indicating strong predictive accuracy and generalization capability.

In addition to accurately tracking stable production behavior, the model successfully captured sudden changes in gas rate trends with a short one-day delay. This characteristic is particularly valuable for the early identification of production anomalies, including the onset of liquid loading, a condition that typically develops over a period of several days. The findings highlight the potential of GRU-based models not only for high-resolution production forecasting but also as intelligent monitoring tools for real-time anomaly detection. The study concludes by recommending the integration of additional operational parameters such as wellhead pressure and temperature to further enhance predictive performance and extend applicability in reservoir surveillance.

Keywords: gas condensate wells, liquid loading, production rate prediction, gated recurrent unit (gru), time series forecasting

Volume 8 Issue 1 - 2026

Vugar M. Fataliyev,[1] Kanan F. Aliyev,[2] Sonakhanim A. Salimova[3]

[1]ASOIU, Oil and gas engineering department, Doctor of Technical Sciences, Azerbaijan
[2]ASOIU, Oil and gas engineering, PhD candidate, Azerbaijan
[3]OGI (Oil and Gas institute), Hydrogasdynamics of Reservoir Systems, Azerbaijan

Correspondence: Dr. Vugar M. Fataliyev, ASOIU, Oil and gas engineering department, Doctor of Technical Sciences, Azerbaijan

Received: December 17, 2025 | **Published:** January 2, 2026

Introduction

Petroleum production forecasting has evolved into a central discipline within reservoir engineering, bridging operational planning, development optimization, and investment decision-making. Historically, forecasting relied on deterministic and empirical approaches such as Decline Curve Analysis (DCA), material balance models, and numerical simulation. Experimental investigations and process modeling under laboratory conditions are developing in parallel, as the tolerances and assumptions introduced by theoretical models require calibration and correction to improve model accuracy and reliability.[1,2] At the same time, it should be noted that although theoretical, laboratory, and hybrid modeling approaches are effective for conventional reservoirs, they often prove insufficient when applied to hydrocarbon fields, particularly gas-condensate systems, due to inherent nonlinearity, uncertainty, and complex multiphase flow dynamics. These challenges were highlighted by Fataliyev and Hamidov[3] during the design of experimental investigations of gas injection into reservoirs. These limitations have accelerated the shift toward machine learning and neural networks, which offer flexibility, data-driven modeling, and the ability to capture patterns beyond explicit physical laws.

Early applications of neural networks in petroleum engineering date back to Ali,[4] who emphasized their potential in handling complex process modeling tasks. Subsequent developments included multi-neural network forecasting frameworks such as those by Nguyen et al.[5] and higher-order neural networks (HONNs) presented by Chakra et al.[6] which aimed to model cumulative production while filtering noisy datasets. Further innovation came with the use of complex-valued Multi-Valued Neurons (MLMVN) by Aizenberg et al.[7] enhancing network flexibility in capturing time series dependencies.

A significant advancement came with deep learning architectures, especially Long Short-Term Memory (LSTM) networks, which proved adept at handling sequence data and long-range temporal dependencies. Liang et al.[8] concluded that modeling multivariate hydrocarbon production time-series data using deep neural networks with functionality similar to LSTM architectures may yield more accurate and computationally efficient production forecasts. Sagheer and Kotb,[9] applied LSTMs to forecast petroleum production, showing marked improvement over both classical models and traditional ANNs. Building on this, Al-Shabandar et al.[10] proposed a deep Gated Recurrent Neural Network (GRNN), which selectively filters relevant temporal features to improve predictive robustness in fluctuating production systems. Gouda et al,[11] implemented an artificial neural network as an intelligent modeling tool due to its high capability and flexibility in capturing complex data patterns. The study provides a detailed comparison between widely used empirical correlations, the Peng–Robinson Equation of State, the Soave–Redlich–Kwong Equation of State, and the proposed ANN model. Statistical and graphical analyses demonstrated the superior performance of the ANN-based model in predicting the target properties.

These advancements are particularly relevant to gas condensate reservoirs, which exhibit complex thermodynamic behavior as pressure falls below the dew point. This results in retrograde condensation, where heavier hydrocarbons accumulate near the wellbore, forming condensate banking and ultimately leading to liquid loading. Liquid loading, a key flow assurance issue, disrupts continuous gas production, introduces backpressure, and causes nonlinear, often abrupt, production declines that challenge traditional forecasting models. Fataliyev et al.[12] extensively reviewed this phenomenon, highlighting that condensate precipitation can occur due to pressure reduction, leading to liquid accumulation, reduced gas flow rates in the bottomhole zone, and unstable well operation. The transition from annular to slug flow regimes typically marks the onset of liquid loading in a production well, posing risks to well integrity and potentially resulting in production failure. This paper reviews the conventional understanding of well liquid loading and examines various techniques used to mitigate its impact. Based on detailed analysis, it proposes a novel pipe element and an automated control system designed to maintain stable production in gas-condensate wells; however, the automated control system still requires integration with reliable gas flow rate forecasting.

Several studies have attempted to incorporate these complexities into neural network-based predictions. Zendehboudi et al.[13] used ANN optimized with particle swarm techniques to predict the condensate-to-gas ratio (CGR) in retrograde systems. Khamis and Fattah,[14] developed models to estimate oil–gas ratios in gas condensate and volatile oil samples. Ashinze et al,[15] expanded this direction by simulating gas condensate production using ANNs that reflect both production decline and flow anomalies caused by condensate accumulation. On the system dynamics front, Ali and Guo,[16] integrated neuro-adaptive models to simulate transient flow performance and pressure drop, key indicators of liquid loading onset. Fataliyev and Aliyev,[17] investigated the application of deep learning neural networks to predict the liquid loading status of wells, which is critical for cost mitigation and production efficiency. Four distinct models were developed and trained using experimental datasets compiled from multiple sources, with key input parameters including wellhead pressure, gas rate, and tubing inside diameter. The models achieved prediction accuracy ranging from 57% to 80%, demonstrating the potential of deep learning for liquid loading prediction in gas-condensate wells. Model performance improved by approximately 17% when both gas rate and wellhead pressure were used simultaneously as input features, suggesting that further optimization is possible by incorporating additional parameters such as fluid properties, well geometry, temperature, and condensate production rate. However, increasing model complexity may complicate implementation and reduce practicality. Al-Fattah and Startzman,[18] highlighted this aspect and demonstrated that dimensionality reduction techniques and sensitivity analysis of input variables can be employed to eliminate redundant and insignificant parameters, thereby simplifying the neural network architecture.

This brief summary highlights that, despite these efforts, many existing models are neither explicitly designed to identify or predict liquid loading events nor optimized to capture systems exhibiting abrupt dynamic transitions driven by condensate banking. This gap motivates the present study, which proposes the application of Gated Recurrent Units (GRUs) for simultaneous forecasting of gas condensate production and early prediction of liquid loading events. GRUs offer a computationally efficient alternative to LSTMs while maintaining the capacity to capture long-term dependencies. Their simpler gating structure reduces overfitting and training complexity, making them suitable for deployment in real-time and constrained-data environments common in field operations.

The novelty of this study lies in demonstrating the effectiveness of a univariate GRU-based deep learning model not only for accurate short-term forecasting of gas production in condensate wells but also for capturing abrupt changes in production behavior with minimal delay. Additionally, the study introduces the novel application of GRU models as a potential diagnostic tool for early detection of liquid loading onset an area that remains underexplored in existing research.

This contribution is particularly important as the petroleum industry increasingly relies on data-driven, automated solutions for production optimization. Accurate prediction of both production and liquid loading onset can reduce the frequency of unplanned shut-ins, minimize deferred production, and support proactive lift and drainage strategies. The proposed GRU-based approach thus represents a step forward in intelligent reservoir management, particularly under the challenging conditions of gas condensate systems.

Methodology

Overview of gated recurrent unit (GRU)

The Gated Recurrent Unit (GRU) is a recurrent neural network (RNN) architecture designed to capture temporal dependencies in sequential data while mitigating the vanishing gradient problem commonly encountered in traditional RNNs. Originally introduced as a simplified alternative to the LSTM unit, the GRU employs gating mechanisms to regulate information flow without the use of separate memory cells. For example, Zainuddin et al.[19] proposed an RNN-GRU–based deep learning approach to forecast the operational states of equipment generating time-series data in the oil and gas sector. The following sections outline the computational mechanics of the GRU and describe its implementation within the experimental framework.

GRU Architecture

Unlike standard RNNs, GRUs incorporate gating mechanisms that regulate information flow through the network, thereby enhancing the model's ability to capture long-term dependencies. Each GRU cell maintains a hidden state h_t at time step t, which is updated based on the current input x_t and the previous hidden state h_{t-1}. Figure 1 presents the structural configuration of the GRU cell, illustrating the general architecture of the unit. The cell utilizes two primary gates: the update gate z_t and the reset gate r_t. As demonstrated by Tan et al,[20] who proposed a novel end-to-end Dual-Attention Time-Aware Gated Recurrent Unit (DATA-GRU) for irregular multivariate time series to predict patient mortality risk, these gates play distinct yet complementary roles in regulating memory retention and temporal dynamics.

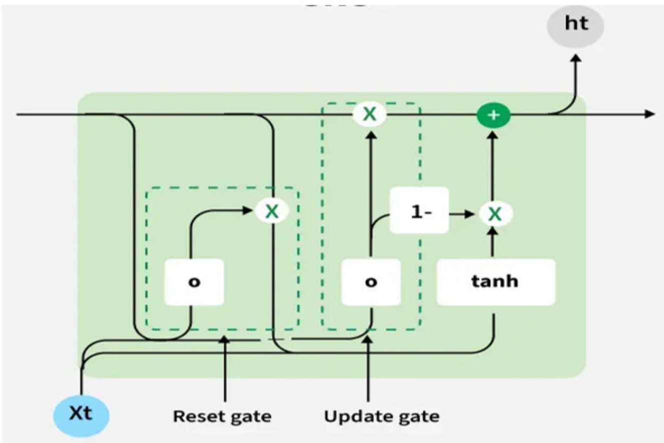

Figure 1 Structure of the sample GRU cell.

Update Gate (z_t)

The update gate determines the extent to which the hidden state from the previous time step (h_{t-1}) should be carried forward to the current hidden state (h_t). It effectively decides how much past information needs to be retained. Mathematically, it is defined as:

$$z_t = \sigma\left(W_z x_t + U_Z h_{t-1} + b_z\right) \quad (1)$$

Where:

$x_t \in R^n$ is the input vector at time step t,

$h_{t-1} \in R^m$ is the previous hidden state,

W_z and U_z are the input and recurrent weight matrices respectively,

b_z is the bias vector,

σ is the sigmoid activation function.

The value of $z_t \in [0, 1]^m$ acts as a gate, with values close to 1 indicating strong retention of the previous state.

Reset gate (r_t)

The reset gate controls how much of the previous hidden state should be forgotten or reset before computing the candidate hidden state. This mechanism allows the GRU to discard irrelevant past information, especially useful for modeling short-term dependencies:

$$r_t = \sigma\left(W_r x_t + U_r h_{t-1} + b_r\right) \quad (2)$$

A smaller value of r_t encourages the unit to ignore previous hidden activations, focusing more on the current input.

Candidate hidden state ($\tilde{h}_t$)

The candidate activation h_t, representing the potential new state, is computed using the reset gate:

$$\tilde{h}_t = tanh\left(W_h x_t + U_h\left(r_t \odot h_{t-1}\right) + b_h\right) \quad (3)$$

Here, $\odot$ denotes the Hadamard (elementwise) product, and tanh is the hyperbolic tangent activation function. By incorporating $r_t \odot h_{t-1}$, the reset gate selectively forgets components of the previous state that are not relevant for the current computation.

Final hidden state update

The final output of the GRU cell, the updated hidden state is then derived by interpolating between the previous between the previous hidden state and the candidate activation, weighted by the update gate:

$$\tilde{h}_t = \left(1 - z_t\right) \odot h_{t-1} + z_t \odot \tilde{h}_t) \quad (4)$$

This equation enables the unit to either preserve historical context $(if\ z_t \approx 0)$ or adopt new information $(if\ z_t \approx 1)$ This formulation allows the model to dynamically retain or update information across time steps, enabling efficient modelling of long-range dependencies with fewer parameters compared to LSTM units.[21]

Training procedure of the Gated Recurrent Unit (GRU) network

The GRU model is trained using sequential input data to predict future values based on past observations. The training process comprises the following steps:

Input preparation

Sequential data are organized into fixed-length overlapping windows, where each input sequence consists of T consecutive time steps (e.g., 10 days of gas rates). Each sequence is paired with the corresponding target output, which is the value immediately following the sequence (e.g., the 11th day's gas rate).

Initialization of hidden states

The GRU model was implemented in a stateless configuration, which is well suited for short-horizon sliding-window forecasting of gas production rates. Using a fixed input window of 10 consecutive days, the hidden state was reset at the beginning of each sequence, enabling the model to focus on localized temporal dynamics relevant to short-term production behavior while mitigating the accumulation of long-term noise and non-stationary effects commonly observed in field data. This design ensures independent processing of training samples, prevents information leakage between windows, and improves model generalization.

For testing and continuous forecasting, a rolling-window prediction strategy was adopted. At each time step, the model was provided with the most recent 10-day production history to generate a one-step-ahead forecast. Within each window, hidden states were updated sequentially to capture intra-window temporal dependencies but were not propagated across consecutive windows. Temporal continuity was implicitly maintained through the overlapping nature of the sliding windows, allowing the model to learn consistent production trends while preserving robustness and stability in the presence of abrupt production rate changes.

Forward propagation

The input sequence is fed into the GRU cell one time step at a time. At each time step t, the hidden state h_t is updated based on the current input x_t and the previous hidden state h_{t-1}, effectively capturing temporal dependencies within the sequence. After processing all time steps, the final hidden state h_T encapsulates the learned representation of the sequence.

Prediction

The final hidden state h_T is passed through an output layer to generate the predicted value $\hat{y}$, representing the forecast for the next time step.

Loss computation

The prediction $\hat{y}$ is compared against the true target value-y using a suitable loss function, such as Mean Squared Error (MSE), to quantify the prediction error.

Backpropagation through time (BPTT)

The computed loss gradients are propagated backward through the GRU's time steps to update the model's parameters (weights and biases). This gradient calculation accounts for temporal dependencies and parameter sharing across time steps.

Parameter update

An optimization algorithm (e.g., Adam or SGD) uses the gradients to adjust the GRU's parameters, aiming to minimize the loss over the training data.

Epoch iteration

Steps 2–7 are repeated for multiple epochs, each epoch involving a complete pass over the training dataset. At the beginning of each epoch and for each input sequence, the hidden states are reinitialized to zero to ensure independent processing. Over successive epochs, the model parameters converge to values that improve prediction accuracy.

Remarks:

i. The hidden state is reset at the start of each input sequence and epoch, preventing carryover of information across sequences during training unless explicitly modeled (e.g., stateful RNNs).

ii. Parameter sharing across time steps allows the GRU to generalize temporal patterns efficiently.

Results

In this study, the gas production rate data of the well X was first normalized using the Minimax Scaler technique to ensure that all input values fell within the [0, 1] range, thereby facilitating more efficient convergence during model training. The time series dataset consists of daily gas production rates spanning approximately 1,500 days. For the purpose of model development and evaluation, the dataset was divided chronologically into two subsets: the first 1,200 days were allocated for training, while the remaining 300 days were reserved for testing. This split was chosen to preserve the temporal continuity of the data, which is crucial for recurrent neural network (RNN)-based architectures such as the Gated Recurrent Unit (GRU).

Figure 2 illustrates the normalized gas rate profile over time. A distinct vertical dashed line at day 1,200 separates the training and testing intervals. As shown, the time series exhibits non-stationary behavior with intermittent fluctuations, underlining the need for models capable of capturing temporal dependencies.

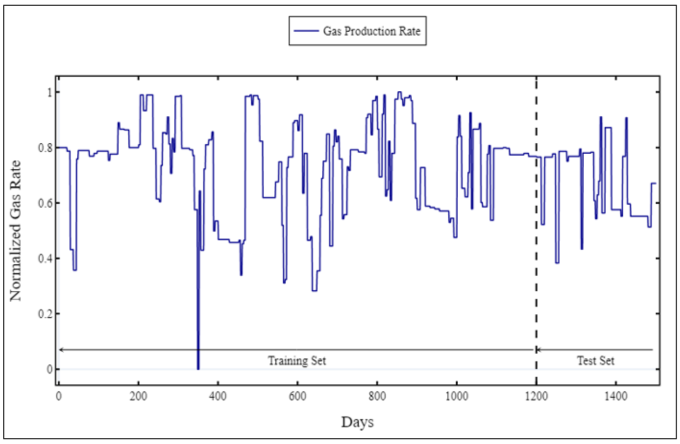

Figure 2 Time series of normalized gas production rate with the training and testing data split indicated.

To prepare the data for time-series forecasting, a sliding-window approach was employed, whereby input sequences of 10 consecutive days were constructed to predict the production rate of the subsequent day. This strategy enables the model to learn temporal patterns by analyzing historical trends over a fixed prediction horizon.

The prediction model was developed using a GRU-based deep neural network. The choice of a deep GRU architecture comprising four stacked layers with 50, 30, 30, and 20 units was motivated by the need to capture complex temporal dependencies and non-stationary fluctuations in daily gas production data, while maintaining higher computational efficiency compared with traditional LSTM-based models. A larger number of units in the initial layers allows the network to learn a diverse set of short- and medium-term temporal features, whereas the progressive reduction in the number of units in deeper layers facilitates feature compression and helps mitigate overfitting. This hierarchical structure aligns with established practices in deep recurrent network design and is well suited for modeling the nonlinear dynamics and transient flow behavior characteristic of gas-condensate reservoirs.

To further enhance generalization performance, dropout layers were interspersed between the GRU layers, promoting robust feature learning by reducing sensitivity to noise and data-specific fluctuations. The Adam optimizer was employed due to its stable and efficient convergence properties in deep recurrent architectures, and an early stopping criterion was applied to automatically terminate training after 10 consecutive epochs without improvement in validation loss. The effectiveness of these architectural and hyperparameter choices is supported by the learning curves, which exhibit a stable validation loss plateau, indicating strong generalization capability and reliable predictive performance on unseen production data.

The GRU model learning curve shown in the Figure 3 provides insights into the model's training process by displaying both training loss and validation loss over epochs.

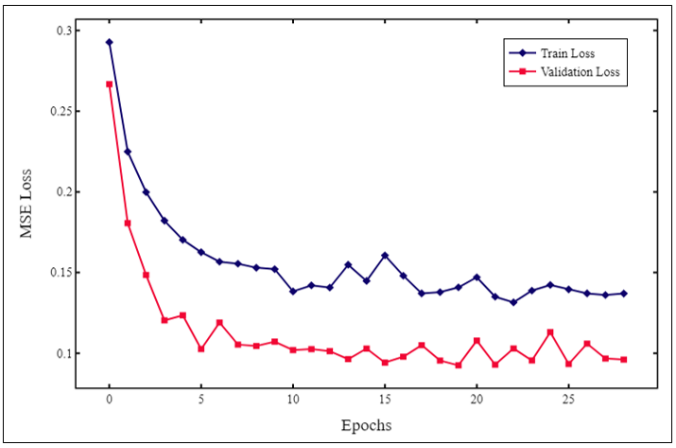

Figure 3 GRU model learning curve showing training and validation loss over epochs.

Initially, both losses start high but decrease rapidly as the model begins to learn the underlying patterns in the data. The training loss (blue line) consistently decreases with each epoch, reflecting the model's improved ability to predict the training set. On the other hand, the validation loss (red line) decreases at a similar rate but begins to plateau after around the 10th epoch, indicating that the model's performance on unseen data has stabilized. This plateau in validation loss suggests that the model has successfully learned the general trends in the dataset without overfitting, as the validation loss remains low while the training loss continues to decrease. During training, the model's mean squared error (MSE) loss stabilized at approximately 0.15 for the training set and 0.10 for the validation set. These values, obtained from standardized inputs, indicate successful learning and good generalization. The slightly lower validation loss suggests the model benefitted from regularization and did not overfit the training data.

The fact that the model's validation loss does not show significant overfitting behavior (e.g., increasing validation loss as the training loss decreases) further supports the model's generalization ability. The use of early stopping (based on validation loss) is evident in this learning curve, as the model's performance on the test data stops improving after a certain point, preventing unnecessary computation and potential overfitting. This learning curve indicates that the GRU model is effective at learning from the data and exhibits strong generalization capabilities for time-series forecasting.

After training, predictions on the test dataset were inverse-transformed to the original scale for performance evaluation. The GRU model achieved a Mean Squared Error (MSE) of 78.58, a Mean Absolute Error (MAE) of 3.35, and a Mean Absolute Percentage Error (MAPE) of 3.98%, indicating strong predictive accuracy. The MSE reflects the average magnitude of squared prediction errors, the MAE provides an interpretable measure of the average absolute deviation in daily gas production rates, and the MAPE expresses the error as a percentage of the actual values, facilitating scale-independent comparison. A MAPE below 5% is generally considered highly accurate in forecasting applications, confirming the suitability of the proposed GRU model for short-term gas production prediction. Since all evaluation metrics were computed over the entire test interval, they represent the model's average performance across the full prediction horizon. As illustrated in Figure 5, the gas production rate during the test period varies approximately between 60 and 120 Mcf/day, placing the observed prediction errors in a relatively small range compared to the overall magnitude of production. Specifically, an MAE of 3.35 indicates that the predicted daily gas production rates deviate from measured values by approximately 3.35 Mcf/day on average. It is important to note that prediction errors are not uniformly distributed over time and are strongly influenced by production dynamics. During periods of stable production, the model closely follows the measured rates with minimal error, whereas during abrupt production changes the error temporarily increases and may reach approximately 6–7 Mcf/day. These deviations are primarily attributable to the model's short response delay to rapid transitions rather than systematic bias, and they do not significantly impact the overall predictive accuracy when averaged over the full test period.

Figure 4 presents a scatter plot comparing the GRU model's predicted gas production rates against the actual values for both the training and test sets. Ideally, points lying close to the 45° reference line indicate highly accurate predictions. As shown, the majority of points for both training (blue) and test (red) sets cluster closely around the reference line, confirming the model's strong predictive capability across both seen and unseen data.

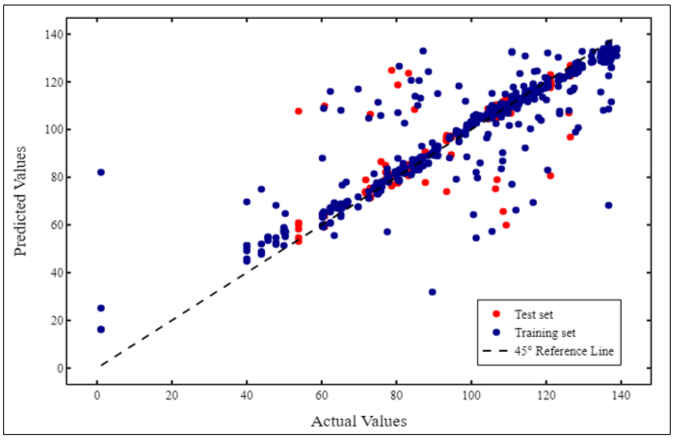

Figure 4 Actual vs. GRU-predicted gas production rates for training and test sets.

While a few outliers exist, particularly at lower and higher production values, the overall alignment indicates that the model successfully captures the underlying production trends (figure 4). The test set predictions show a slightly wider spread, which is expected due to the absence of direct exposure during training, yet the deviation remains within acceptable bounds. This plot further validates the GRU model's effectiveness in learning temporal patterns and generalizing well to future production behavior.

Figure 5 presents the comparison between the GRU model's predicted gas production rates and the actual measured values for the test period. The production profile demonstrates a typical pattern observed in gas wells, where long periods of stable flow are occasionally disrupted by sudden drops or increases in production rates. The GRU model exhibits a strong ability to closely replicate the actual rates during the stable intervals, maintaining high accuracy with minimal deviation. This consistency reinforces the model's effectiveness in learning and predicting steady-state behaviors that are commonly encountered in production operations.

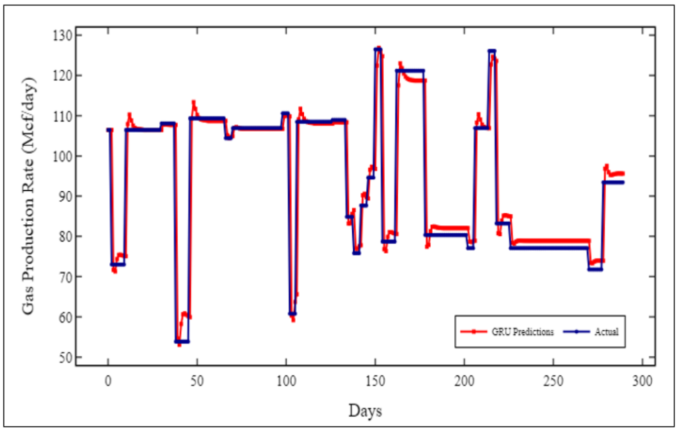

Figure 5 GRU model predictions versus actual gas production rates for the test period.

More importantly, the model also demonstrates the capability to detect abrupt changes in gas rate trends, although with a slight lag of approximately one day. This short delay is understandable, given that recurrent neural networks rely on sequential input and require temporal context to update their internal memory state. The GRU model demonstrates a strong ability to closely replicate actual gas production rates during stable intervals, maintaining high predictive accuracy with minimal deviations on the order of 1–3 Mcf/day, despite production rates over the test period varying approximately between 60 and 120 Mcf/day. Despite this minor lag, the model is able to rapidly adapt its predictions following sudden transitions in the production profile, which is a critical strength in real-world monitoring systems.

This predictive behavior has important practical implications. One such application is the early detection of liquid loading, a common flow assurance issue in gas condensate wells where accumulated liquids begin to restrict gas flow. Since liquid loading does not occur instantaneously but develops over several days (typically one week to ten days), a model that can recognize early anomalies in the production trend—such as a sudden decline in gas rate—offers significant value. The GRU model's sensitivity to these early shifts, even with a short delay, suggests its potential use not only for accurate rate forecasting but also as a diagnostic tool for identifying the onset of liquid loading.

Conclusion

This study investigated the application of a Gated Recurrent Unit (GRU) based deep learning architecture for forecasting daily gas production rates in a gas condensate well. The primary objective was to assess the model's ability to accurately predict both steady-state production behavior and sudden rate fluctuations, which are commonly observed in field operations. The proposed GRU model demonstrated strong performance, achieving a Mean Squared Error (MSE) of 78.58, a Mean Absolute Error (MAE) of 3.35, and a Mean Absolute Percentage Error (MAPE) of 3.98% on the test set. These results confirm the model's high predictive accuracy and ability to generalize well on unseen data.

It should be noted that using only univariate data may impose certain limitations when applying this approach to multivariate processes. However, the model is designed to be connected to live well data and continuously updated through ongoing real-time history matching, which inherently accounts for the influence of other factors. Consequently, these effects are reflected in the trend dynamics. For this reason, the model has demonstrated effectiveness not only in capturing stable production trends with minimal deviation but also in detecting abrupt changes in the production profile with only a short, one-day delay. This predictive behavior suggests that GRU-based models are not only suitable for production forecasting but may also serve as valuable tools for early anomaly detection. One important implication is the potential use of GRU models for the early identification of liquid loading onset, which typically manifests gradually over a one-to-two-week period. By recognizing the early signs of rate decline, the model could assist in proactive flow assurance management.

The novelty of this study lies in demonstrating that a purely data-driven GRU model can replicate both gradual and transient flow dynamics in gas wells with high reliability, without relying on mechanistic models or explicit reservoir parameters. These findings provide a strong foundation for integrating GRU-based architectures into intelligent well monitoring systems. Future research may further enhance model performance by incorporating additional operational parameters such as wellhead pressure, bottomhole pressure, and temperature. This would enable a more comprehensive understanding of well behavior and support the development of more robust and interpretable prediction frameworks for use in reservoir surveillance and production optimization.

Acknowledgements

None.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Abbasov ZY, Fataliyev VM. The effect of gas–condensate reservoir depletion stages on gas injection and the importance of the aerosol state of fluids in this process. *J Nat Gas Sci Eng*. 2016;31:779–790.
2. Hamidov NN, Fataliyev VM. Experimental study into the effectiveness of the partial gas cycling process in the gas–condensate reservoir development. *J Pet Sci Technol*. 2016;34(7):677–684.
3. Fataliyev VM, Hamidov NN. Effective "vaporizer" for recovering retrograde hydrocarbon condensate from a gas–condensate reservoir. *Int J Petrochem Sci Eng*. 2017;2(6):302–308.
4. Ali JK. Neural networks: a new tool for the petroleum industry? *Proceedings of the SPE European Petroleum Computer Conference*. 1994.
5. Nguyen HH, Chan CW, Wilson M. Prediction of oil well production: a multiple–neural–network approach. *Intell Data Anal.* 2004;8(2):151–169.
6. Chakra NC, Song KY, Gupta MM, et al. An innovative neural forecast of cumulative oil production from a petroleum reservoir employing higher–order neural networks. *J Pet Sci Eng*. 2013;106:18–33.
7. Aizenberg I, Sheremetov L, Villa Vargas L. Multilayer neural network with multi–valued neurons in time series forecasting of oil production. *Neurocomputing*. 2016;175:980–989.
8. Liang B, Liu J, You J, et al. Hydrocarbon production dynamics forecasting using machine learning: a state–of–the–art review. *Fuel*. 2023;337:127067.
9. Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*. 2019;323:203–213.
10. Al–Shabandar R, Jaddoa A, Hussain AJ, et al. A deep gated recurrent neural network for petroleum production forecasting. *Mach Learn Appl*. 2021;3:100013.
11. Gouda A, Gomaa S, Attia A, et al. Development of an artificial neural network model for predicting the dew point pressure of retrograde gas condensate. *J Pet Sci Eng*. 2022;208:109390.
12. Fataliyev VM, Hamidov NN, Aliyev KF. Advances in understanding and controlling liquid loading in gas–condensate production well. *SOCAR Proc***.** 2025;(2):92–103.
13. Zendehboudi S, Ahmadi MA, James L, et al. Prediction of condensate–to–gas ratio for retrograde gas condensate reservoirs using artificial neural network with particle swarm optimization. *Energy Fuels*. 2012;26(6):3432–3445.
14. Khamis MA, Fattah KA. Estimating oil–gas ratio for volatile oil and gas condensate reservoirs: artificial neural network, support vector machines and functional network approach. *J Pet Explor Prod Technol*. 2019;9:573–582.
15. Ashinze AS, Adeniyi AT, Giwa A. Modelling and simulation of natural gas condensate production using artificial neural network. *Proceedings of the SPE Nigeria Annual International Conference and Exhibition Society of Petroleum Engineers*. 2023.
16. Ali A, Guo L. Neuro–adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir. *IFAC–PapersOnLine*. 2019;52(29):122–127.
17. Fataliyev VM, Aliyev KF. Predictive modeling of liquid loading in gas condensate wells using deep neural networks. *Azerb Oil Ind***.** 2025;(9):13–19.
18. Al–Fattah SM, Startzman RA. Predicting natural gas production using artificial neural network. *Proceedings of the SPE Hydrocarbon Economics and Evaluation Symposium. Society of Petroleum Engineers*. 2001.
19. Zainuddin Z, Hasan MH. Predicting machine failure using recurrent neural network–gated recurrent unit (RNN–GRU) through time series data. *Bull Electr Eng Inform*. 2021;10(2):1011–1019.
20. Tan Q, Ye M, Yang B, et al. Data–GRU: dual–attention time–aware gated recurrent unit for irregular multivariate time series. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020;34(1):5956–5963.
21. Wang Y, Feng S, Wang B, et al. Deep transition network with gating mechanism for multivariate time series forecasting. *Appl Intell*. 2023;53:24346–24359.