Journal of Soft Computing Exploration

Homepage: https://shmpublisher.com/index.php/joscex

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Model bidirectional GRU with bayesian optimization for dry gas pressure forecasting in transmission pipeline networks

Mohammad Alvinanda Kurniawan[1], Aditya Firman Ihsan[2], I Wayan Palton Anuwiksa[3]
[1,2,3]School of Computing, Telkom University, Indonesia

Article Info

Article history:

Received June 09, 2026
Revised June 13, 2026
Accepted June 16, 2026

Keywords:

Bidirectional GRU
Bayesian optimization
Time series forecasting
Gas pipeline
Deep learning

ABSTRACT

Operational robustness and infrastructure safety require that stable dry gas pressure in transmission pipeline networks be maintained. Due to the highly nonlinear and time-varying behaviors of the dry gas pressure, conventional methods cannot predict its tendencies accurately. Deep learning models have recently gained good performance for time-series forecasting, however combination of Bidirectional Gated Recurrent Unit (BiGRU) and Bayesian Optimization based approach for dry gas pressure forecasting remains underexplored. Based on these, this study develops Bayesian optimization-enhanced forecasting framework for BiGRU to enhance the prediction performance. The framework is based on a Gaussian Process surrogate model and an acquisition function — in this case Expected Improvement, which guides the search for optimal hyperparameter configurations. We used operational SCADA time-series data consisting of hourly measurements (pressure, temperature, flowrate and gas composition) collected from a dry gas transmission pipeline across two annual periods. Pre-processing of data consists of dealing with missing data, Min–Max normalization and sliding window conversion. Here a multistep forecasting scenario for the next 20 hours was used, and thismultistep prediction was evaluated calculating the MAE, RMSE, and R^2. The optimized BiGRU performed the MAE of 11.7729 psi, RMSE of 14.7827 psi and R^2 of 0.9389, which enhanced the baseline model by 13.67%, 12.90%, and 1.60 percentage points respectively. These results indicate that Bayesian Optimization improves the forecasting performance of BiGRU and, at the same time, decreases the manual hyperparameter tuning efforts.

Corresponding Author:

Mohammad Alvinanda Kurniawan,
School of Computing,
Telkom University,
Jl. Telekomunikasi No.1, Sukapura, Kec. Dayeuhkolot, Kabupaten Bandung, Jawa Barat 40257, Indonesia.
Email: alvinnanda.kurniawan@gmail.com
https://doi.org/10.52465/joscex.v7i3.166

1. INTRODUCTION

Natural gas transmission lines are an important part of utility infrastructure in modern energy distribution systems. These corridors are designed to carry billions of cubic feet of natural gas from fields to industrial end-users and power plants [1], [2]. While in operation, dry gas pressure at strategic locations in the section must be constrained within known parameters of operation to prevent disruption of supply to downstream areas of the infrastructure. The pressure fluctuations can make the system less effective at the same time increasing the risk of pipeline ruptures, compressor malfunctions and equipment damage. Thus, accurate forecasting of the gas pressure is highly valuable to enable proactive operational decisions that guarantee a reliable gas transmission system [3], [4].

Contemporary pipeline monitoring systems (CMS) continuously collect SCADA data of operational variables such as pressure, temperature and flow rate, producing high-resolution historical time-series data which enable novel data-driven forecasting methods without the need for explicit mathematical formulation of complex gas flow equations [3], [5], [6]. However, the pressure fluctuations of dry gas largely demonstrate nonlinear behavior due to compressor performance, demand variations, and operational conditions in pipelines [1], [7], [8]. Despite the wide spectrum of forecasting methods that have been utilized in industry, it has been shown that many traditional statistical models fail to incorporate long-range temporal dependencies and nonlinear relationships present in high-dimensional data. Additionally, most of the previous work here is oriented towards flow forecasting rather than pressure prediction which only partially tackles data science applications related to operational problems in gas transmission systems. As a result, prediction performance can suffer under varying operating conditions and the importance of more robust forecast methods starts to emerge [6].

Batch Sequential Prediction where deep learning performed outstanding in a number of sequential prediction tasks. For a range of industrial applications, Recurrent Neural Network (RNN) architectures, in particular Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), exhibit more nondeterministic temporal modeling capability [9], [10]. During last few years, BiGRU has gained traction due to its ability to learn and process sequential information both forwards and backwards which allows the model a wider temporal context than traditional unidirectional architectures [10-13]. The performance of BiGRU in several recent applications has been verified, such as intelligent power load forecasting [11], stock price prediction [12], air quality forecasting [14]. The results reflect that BiGRU will be an effective candidate technique for reaching successful realizations of dry gas pressure forecasting models in transmission pipeline systems. However, even though BiGRU has a lot of advantages, the predicting performance highly relies on the choice of hyperparameter. In many past studies, hyperparameters were set with manual trial-and-error procedures or conventional search methods that are computationally onerous requiring extensive computational resources to conduct the searches and often not discovering effective settings for complex industrial datasets. Consequently, the model stability and generalization of BiGRUs depending its application in each forecasting window may be significantly different.

Therefore, several research gaps remain. First, transmission pipeline system dry gas pressure forecasting studies are still limited compared to other prediction targets (e.g., gas flow and energy demand). And finally, to the best of our knowledge, this is one of the few explorations that combines the usage of BiGRU with Bayesian Optimization for this type of applications. Thirdly, the role of automated hyperparameter optimization for forecasting robustness and predictive performance in industrial SCADA settings is yet to be quantitatively assessed [15]–[17]. Bayesian optimization [18], has been suggested in study [19] as a more efficient approach for hyperparameter tuning that constructs a probabilistic surrogate model of the objective function and then employs an acquisition function to directly explore the search space. Yet most studies available only explore BiGRU & Bayesian Optimization separately, while limited studies connect the two approaches to predict dry gas pressure in transmission pipelines. Moreover, existing studies primarily focus on optimizing the model accuracy but fail to appraise comprehensively the impact of automated hyperparameter optimization towards forecasting robustness and performance.

In light of the above, this study proposes a Bidirectional Gated Recurrent Unit (BiGRU) framework coupled with Bayesian Optimization for devising a prediction model of pressure behavior within gas transmission pipeline systems. Bayesian Optimization is used to automatically find the best setting for some important hyper-parameters like sequence length, hidden units, learning rate, dropout rate and batch size. The proposed methodology is validated on two years of actual SCADA operating data obtained from a dry gas trans-mission pipeline. More specifically, the goal of this work is to explore the application of BiGRU for dry gas pressure forecasting and evaluate how much Bayesian Optimization (BO) improves forecasting performance over a reference configuration. The development of the method contributes to a more accurate and efficient solution for intelligent operational management from the perspective of forecasting in natural gas transmission systems.

2. METHOD

Research Framework

In this paper, we present a novel forecasting framework for dry gas pressure prediction in transmission pipeline systems based on the combination of Bidirectional Gated Recurrent Unit (BiGRU) and Bayesian Optimization (BO). The overall research workflow used in this work is shown in Figure 1, which includes data acquisition and data preparation, model development, hyperparameter tuning for time series forecasting (and prediction), and to evaluate the performance of them.

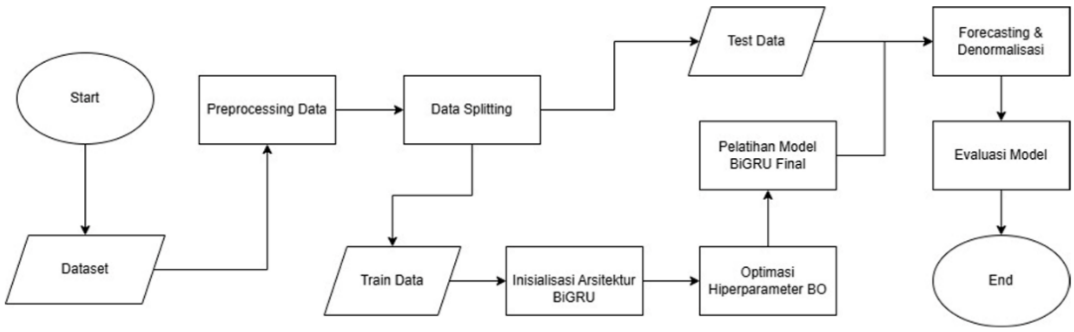

Figure 1. The proposed framework

The framework we propose is divided into four major periods. The operational SCADA data are collected and preprocessed by filling missing values, normalization and sequence transformation. Second, the transformed data are used to construct BiGRU forecasting model and Bayesian Optimization is used for determining best hyperparameter configuration automatically. Third, the optimized model is employed to predict multistep dry gas pressures. Last, prediction performance is assessed by the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and coefficient of determination (R^2).

Dataset acquisition and characteristics

The dataset used in this study contains measurements from a Supervisory Control and Data Acquisition (SCADA) system for a dry gas transmission pipeline network. SCADA constantly logs readings from field instruments, thus offering high-frequency historical data with which to exploit for data-driven forecasting applications. Measurements were done at hourly intervals throughout two annual operational cycles in order to adequately capture the temporal dynamics associated with pipeline operations.

The aggregated dataset contained the various univariate time-series variables that reflect the operational situation of the transmission system. Such as: target variable dry gas pressure, together with temperature measurements flow-related variables and gas composition loggers as predictive features from October 2023. The chosen variables were deemed operationally relevant since they represent the physical conditions affecting how pressure behaves within the transmission line.

In total, 17,500 sequential samples were created from the processed time series data which were used for training and validation of our models. The input–output sequences used in our work were created from the original observations based on a multistep forecasting-based sliding window mechanism. Table 1 gives an overview of the characteristics of the SCADA dataset used in this work.

Table 1. Characteristics of the SCADA dataset

Characteristic	Value
Data Source	SCADA System
Application Domain	Dry Gas Transmission Pipeline
Observation Period	Two annual operational periods
Data Type	Multivariate Time Series
Total Input Features	17
Input Variables	Historical Pressure, Temperature, Energy Rate, Volume Rate, and Gas Composition Features (C2, C3, iC4, nC4, iC5, nC5, C6, C7, C8, C9, N_2, CO_2, H_2O)
Target Variable	Pressure
Recording Interval	Hourly
Forecast Horizon	20 Hours
Data Normalization	Min-Max Scaling

.

Data preprocessing

Before developing the model, we carried out several data preprocessing steps on the SCADA dataset to enhance its quality and experiment with transforming raw observations into a form appropriate for multistep forecasts. The preprocessing comprised four sequential steps. Initially missing observations were treated by interpolation to preserve time-series continuity. Second, Min–Max normalization was used to scale the feature values into a common range to accelerate model convergence when trained. In accordance with [20], Eq. (1).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where X denote the original value, X_{min} and X_{max} represent the minimum and maximum values of each feature, respectively. To create the sequential samples for time dependencies, the normalized data are created as sequential samples with a sliding window mechanism. The model takes sequences of past observations as inputs and predicts future Dry Gas Pressure values.

Third, the sequential samples were created based on a sliding window mechanism with respect to the normalized time-series data. A sequence of past observations from the history for each sample was fed to the model and values of pressure in the forecasting horizon made up the prediction target. The generated sequences were then arranged in preparation for subsequent training and evaluation of the models.

In our study, no custom feature was created. Instead, as these variables host domain-relevant signals that impinge on dry gas pressure dynamics of transmission pipeline operations, the forecasting framework was directly fed with operational variables available from the SCADA system: pressure, temperature, flow-related variables and measurements of gas composition.

Data Splitting Strategy

Sequential samples were created by partitioning the spectrograms to maintain the temporal nature of the forecasting problem to avoid information leakage due to random shuffling. The first 80% of the observations were assigned to the model learning phase, and the last 20% of observations were saved as a validation set. As a result, the test set featured only observations that came after the training period, which allows for an accurate prediction of what would happen when deploying this type of model in real life.

Of the model development, the last 20% of sequential samples were further categorized as the validation set. It kept the temporal ordering of the data and an independent portion to assess its performance, allowed for early stopping during training, as well as guided a hyperparameter search process. Consequently, the partitioning of the created sequential samples was effective as follows: (63.96% training set; 15.99% validation set and 20.05% test set) as detailed in Table 2.

Table 2. The distribution of final sequential samples used in the study

Dataset Partition	Number of Sequential Samples	Percentage
Training	11,193	63.96%
Validation	2,798	15.99%
Testing	3,509	20.05%

The validation subset was used internally during model development to enable early stopping and evaluate candidate hyperparameter configurations within the Bayesian Optimization framework, while the testing subset remained unobserved until final performance evaluation.

Bidirectional Gated Recurrent Unit

Bidirectional Gated Recurrent Unit (BiGRU):An improvement of the standard GRU model that learns sequential information in both forward and backward direction [13], [21]. The dual-direction nature of this mechanism allows for simultaneous exploration of past and future contextual information, leading to more informative temporal representations. The initial part of GRU, the update gate outputs an amount for keeping information from previous hidden state and which number updating hidden state. The update gate is specified by Eq. (2).

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{2}$$

Here z_t is the update gate vector at time step t, x_t is the current input vector, h_{t-1} is the hidden state of the previous time step, W_z and U_z are the corresponding weight matrices, b_z is the bias term, and σ denote the sigmoid activation function. The sigmoid function produces values between 0 and 1, determining the proportion of information to be preserved.

Second part is also known as the reset gate, which decide how much of history should be forgotten when creating new candidate information. The reset gate can be defined as Eq. (3).

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{3}$$

Where r_t represent the reset gate vector, W_r and U_r are trainable weight matrices, and b_r is the corresponding bias vector. A lower reset gate value causes the model to forget irrelevant historical information, allowing it to adapt to changing temporal patterns.
The reset gate is subsequently combined with the previous hidden state to generate the candidate hidden state, which is computed as Eq. (4).

$$\tilde{h}_t = tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \tag{4}$$

Where $\tilde{h}_t$ denotes the candidate hidden state, W_h and U_h are learnable weight matrices, b_h is the bias term, $\odot$ denotes element-wise multiplication, and $tanh$ is the hyperbolic tangent activation function that maps the output into the range $[-1,1]$.

Finally, the current hidden state is obtained by combining the previous hidden state and the candidate hidden state through the update gate as follows Eq. (5).

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{5}$$

Where h_t represents the hidden state at the current time step. The update gate dynamically balances the contribution of past memory and newly learned information, enabling GRU to effectively model long-term temporal dependencies while avoiding the vanishing gradient problem.

In the proposed BiGRU architecture, two GRU layers operate simultaneously in opposite directions. The forward GRU processes the input sequence from the beginning to the end, while the backward GRU processes the sequence in reverse order. The outputs of both directions are then concatenated to form a comprehensive feature representation in Eq. (6).

$$h_t^{BiGRU} = [\overrightarrow{h_t}; \overleftarrow{h_t}] \tag{6}$$

Where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ denote the hidden states generated by the forward and backward GRU networks, respectively, and $[\cdot;\cdot]$ represents the concatenation operation. The bidirectional learning mechanism enables the proposed model to exploit contextual information from both temporal directions, thereby improving its capability to capture complex nonlinear patterns in dry gas pressure fluctuations.

Bayesian Optimization

Hyperparameter selection is pivotal in the predictive performance of deep learning models. In this paper, we used Bayesian Optimization to automatically learn the best hyperparameter combination of BiGRU model.

Bayesian Optimization builds a probabilistic surrogate model from the observations using Gaussian Process (GP), with an acquisition function leading to explore promising regions in hyperparameter space, e.g.: Expected Improvement [22]. Bayesian optimization highly reduces the computational cost than exhaustive ones while keeping a competitive performance [23]. The optimization process focused on the following hyperparameters which are Sequence Length, Hidden Units, Learning Rate, Dropout Rate, Batch Size. A total of 50 optimization iterations were performed to identify the optimal parameter combination.

Experimental Setup

We devised two experimental settings in order to examine the efficacy of Bayesian Optimization in improving the forecasting performance of BiGRU. The first scenario (S1) used a baseline BiGRU model with hyperparameter values manually selected based on empirical practices from deep learning literature for time-series forecasting. Scenario 2 (S2) used Bayesian Optimization to find best hyperparameter automatically.

.

For a fair comparison, both the scenarios were trained and evaluated with same dataset, forecasting horizon, optimizer and loss function. The initial configuration was deliberately maintained to be representative of a realistic modeling situation where hyperparameters are chosen based upon standard trial-and-error methods. The comparison permitted the quantification of contribution by Bayesian Optimization relative to other techniques under similar experiments. The hyperparameter configurations of both experimental scenarios are presented in Table 3. As shown in Table 4, hyperparameter tuning was performed considering a predefined search space, so the optimized hyperparameters reported in Table 3 were evaluated by selecting a random subset from that search space. For reproducibility and transparency, the search range used for each optimized hyperparameter was chosen from standard practices in deep learning-based applications as well as the characteristics of the forecasting problem addressed.

Table 3. Comparison of baseline and optimized hyperparameter configurations

Hyperparameter	Baseline BiGRU (S1)	Optimized BiGRU (S2)
Sequence Length	24	24
Number of Hidden Units	64	197
Dropout Rate	0.20	0.2574
Learning Rate	0.001	0.00350
Batch Size	64	56
Number of Layers	1	1

Table 4. Hyperparameter search space for Bayesian Optimization

Hyperparameter	Search Range
Sequence Length	12–48
Number of Hidden Units	64–256
Dropout Rate	0.1–0.4
Learning Rate	$5\times10^{-4} - 1\times10^{-2}$
Batch Size	16–64

We performed 50 iterations of a Bayesian optimization process, using a Gaussian Process surrogate model and an Expected Improvement acquisition function that are 5 random initializations, and 45 directed searches. This optimization minimized the Mean Squared Error (MSE) of the validation of the two-fold time series cross-validation scheme.

Performance Evaluation

To evaluate the forecasting performance of the developed model, three widely used evaluation metrics were used: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination R^2. These metrics were chosen because they offer complementary perspectives on prediction accuracy, error magnitude, and the amount of variance in the observed data explained by the forecasting model [24], [25]. MAE stands for absolute error, which measures the average magnitude of the error in a set of predictions, regardless of its direction Eq. (7).

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_t - \hat{y_t}| \tag{7}$$

The Root Mean Square Error (RMSE) is the square root of mean squared errors and thus greater punishment for large distances between predicted value compared to actual. RMSE is formulated as Eq. (8).

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y_t})^2} \tag{8}$$

In the meantime, coefficient of determination R^2 evaluates how much variation in observed data is accounted for by forecasting model and can be expressed as Eq. (9).

$$R^2 = 1 - \frac{\sum_{t=1}^{N}(y_t - \hat{y_t})^2}{\sum_{t=1}^{N}(y_t - \bar{y})^2} \tag{9}$$

Where N denote number of data in the test set, y_t represents actual value at time (t), $\hat{y}_t$ represents model prediction value at time (t), and $\bar{y}$ represents the average of the actual values across all test data. Lower values of RMSE and MAE, together with a higher R^2, indicate better forecasting performance.

Because this study uses a straightforward multi-step forecasting strategy with a 20-hour prediction horizon, evaluation metrics are initially calculated for each step and then averaged across all steps to represent overall forecasting performance. Consequently, the reported performance in terms of MAE, RMSE, and R^2 is averaged across all forecasting horizons, rather than reflecting the error observed at a single prediction step.

3. RESULTS AND DISCUSSIONS

This section discusses the experimental and analytical results of a Bidirectional Gated Recurrent Unit (BiGRU) network using Bayesian Optimization for dry gas pressure prediction. Experiments comparing the ability of the proposed model to automatically optimize hyperparameters versus the baseline BiGRU configuration are conducted across datasets. Finally, further analysis is performed to evaluate the convergence of Bayesian Optimization and its applicability as a forecasting method based on this research.

Forecasting Performance Analysis

Figure 2 shows the convergence process in Bayesian optimization when tuning hyperparameters. During the initial iterations, the optimization algorithm rapidly increases the objective function value and slowly converges around a stable optimum. Figure 2 shows that, based on the mathematical approximation of the Gaussian Process surrogate model, it was possible to avoid over-exploring the search space and thus obtained the best hyperparameter configuration at the 17^{th} iteration.

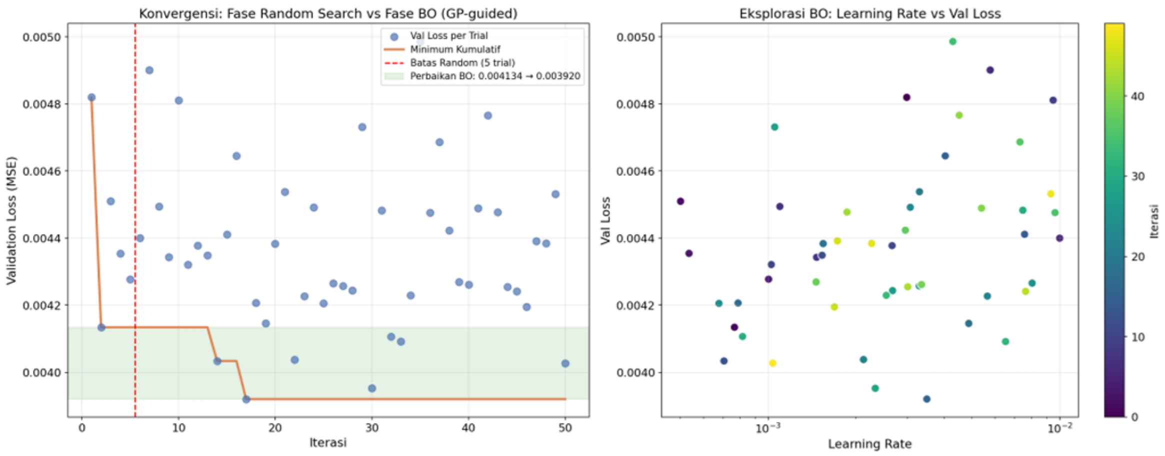

Figure 2. Convergence of bayesian optimization during 50 search iterations

The optimal hyperparameter configuration identified by Bayesian Optimization is summarized in Table 5.

Table 5. Best Hyperparameter configuration

Parameter	**Optimal Value**
Sequence Length	24
Num units	197
Num layers	1
Dropout Rate	0.2574
Learning Rate	0.00350
Batch Size	56

Next, the forecasting performance of the baseline and optimized models was evaluated using the Mean Absolute Error (MAE), Root Means Squared Error (RMSE), and coefficient of determination (R^2). Table 6 shows the comparison results.

Table 6. Performance comparison between baseline and proposed models

Model	MAE (psi)	RMSE (psi)	R^2
Baseline BiGRU (S1)	13.64	16.97	0.923
Proposed BiGRU + BO (S2)	11.7729	14.7827	0.9389

Table 6 shows the forecasting performance of the baseline BiGRU model (S1) and the Bayesian-optimized BiGRU model (S2). Experimental results show that the optimized model consistently outperforms the baseline model for all evaluation metrics. In fact, the optimized model achieved a lower MAE (11.7729 psi) and a decreased RMSE (14.7827 psi), while increasing the coefficient of determination R^2 from 0.923 to 0.9389.

Bayesian optimization reduced MAE by 13.67% and RMSE by 12.90%, while increasing R^2 (1.60 percentage points) compared to the baseline configuration on average. This improvement reflects the BiGRU model's ability to better extract temporal features of dry gas pressure data.

This improved forecasting performance stems from the use of Bayesian optimization to automatically find an effective model hyperparameter mix without extensive manual trial and error. The proposed framework improves the BiGRU architecture's learning capability and generalization performance in nonlinear gas pressure dynamics modeling by optimizing key parameters, including sequence length, hidden units, learning rate, dropout rate, and batch size.

Figure 3 shows a plot of actual dry gas pressure values versus predictions from the developed model with optimized hyperparameters. As can be seen, the prediction curve closely follows the actual fluctuations throughout the testing period, indicating that the proposed model is capable of accurately approximating local variations and long-term dependencies.

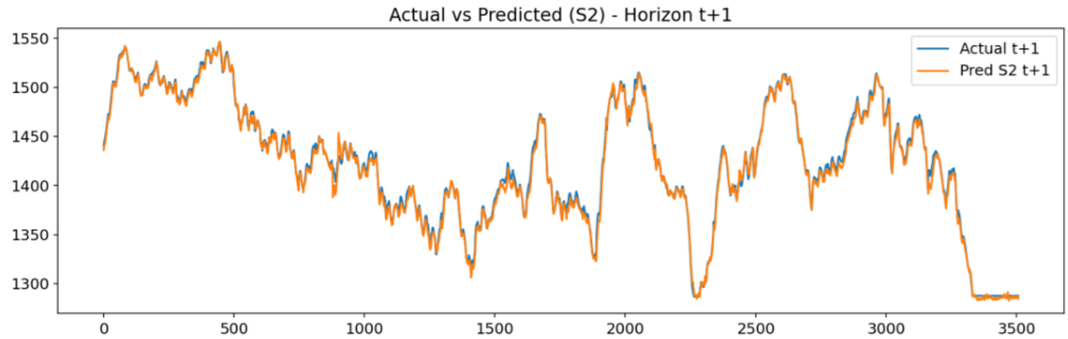

Figure 3. Comparison between actual and predicted dry gas pressure values using the optimized BiGRU model

Moreover, prediction results on the entire observation period show that over time series trend and noise beans in consideration, the forecasting performance of proposed model is pretty stable among many operating conditions implying an excellent generalization ability for practical usage covering real transmission pipeline systems.

Discussion

The better performance obtained by applying the optimized BiGRU model confirms the importance of hyperparameter selection on such exercises as multistep forecasting tasks. Bayesian Optimization allowed for the discovery of a better trade-off between model complexity, regularization and learning dynamics as opposed to manual selection. As a result, the optimized model demonstrated improved generalization ability and lower prediction error on unseen test data.

The recommended framework might also prove effective due to the properties of SCADA data which provide hourly operational dynamics and form nonlinear correlations between process variables. The bidirectional learning mechanism of BiGRU is helpful to capture temporal patterns in these industrial datasets, which helps accurate pressure forecasting.

What's more, the optimization process found a 24-hour sequence length to be optimal: this seems reasonable since having on the order of one day worth of past operational information gave enough context for pressure behavior relevant over roughly twenty future hours. This result is possibly a reflection of the diurnal operational characteristics associated with gas transmission systems.

Although direct numerical comparisons across prior studies should be interpreted with caution because differences in datasets, application domains, forecasting horizons and measurement scales influence forecast performance. Thus To this end, Table 7 provides a qualitative comparison designed merely to illustrate methodological distinctions and the unique contributions of our proposal rather than establish strict performance superiority.

Table 7. Qualitative comparison of the proposed approach with previous studies

Study	Application Domain	Forecasting Model	Hyperparameter Optimization Strategy	Main Contribution
Liu et al. (2022) [6]	Natural gas flow forecasting	LSTM	Manual tuning	Demonstrated the capability of LSTM to model nonlinear temporal patterns in gas transmission systems.
Wang & Jiang (2025) [18]	Energy demand forecasting	GRU	Grid Search	Improved forecasting accuracy through systematic hyperparameter selection.
Wang et al. (2025) [11]	Industrial multivariate forecasting	BiGRU	Manual tuning	Utilized bidirectional learning to capture complex temporal dependencies.
Proposed study	Dry gas pressure forecasting in transmission pipelines	BiGRU	Bayesian Optimization	Integrated BiGRU with automated hyperparameter optimization for direct multistep dry gas pressure forecasting using SCADA data.

As demonstrated in Table 7, existing studies have experimented with different deep learning architectures and hyperparameter tuning techniques for applications on time-series forecasting. As a result, the potential of dry gas pressure forecasting for transmission pipeline systems remains resolutely absent from existing studies more so if BiGRU is combined with Bayesian Optimization. We address this gap by proposing a framework to achieve automatic temporal hyperparameter tuning directly in an industrial SCADA operational environment.

4. CONCLUSION

This study explores the accuracy of an out-of-sample multistep forecasting model for a dry gas pipeline system that combines Bidirectional Gated Recurrent Units (Bi-GRU) and Bayesian Optimization (BO). The results demonstrate that Bayesian Optimization is capable of selecting better hyperparameter configuration settings, resulting in superior forecasting performance compared to the baseline BiGRU model, using manually selected configurations. The proposed framework successfully captures the nonlinear temporal dynamics of dry gas pressure and achieves good predictions up to 20 hours into the future. In terms of scientific utility, this study adds to the exploration of dry gas pressure forecasting by validating the novel combination of BiGRU and automated hyperparameter optimization in an industrial SCADA setting. Finally, the analysis emphasizes systematic hyperparameter tuning to improve the robustness and generalizability of deep learning models in industrial time series forecasting problems.

However, there are several important limitations to consider. Data from a single dry gas transmission system using a fixed forecast horizon of 20 hours was used to evaluate the proposed framework. Therefore, the applicability of the results to alternative pipeline layouts, in-service conditions, or prediction scopes should be further assessed. As a next step, future studies could extend the current work by testing the proposed dual-objective framework on additional datasets generated from different transmission systems, and assessing the scope of adaptive forecasting or optimized solutions using alternative optimization strategies or hybrid deep learning architectures to further improve model performance while enhancing operational applicability.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Author 1 was responsible for the main research activities, including conceptualization, data collection, system implementation, experimentation, analysis, and manuscript preparation. **Authors 2 and 3** contributed as research supervisors by providing academic guidance, methodological direction, technical evaluation, and critical review throughout the research and manuscript development process.

REFERENCES

[1] G. Fambri, C. Diaz-Londono, A. Mazza, M. Badami, T. Sihvonen, and R. Weiss, "Techno-economic analysis of Power-to-Gas plants in a gas and electricity distribution network system with high renewable energy penetration," *Appl. Energy*, vol. 312, p. 118743, 2022.

[2] T. Grudniewski, M. Dobrzyńska-Dąbska, Z. Ciekanowski, V. Kuznetsov, A. Rojek, and S. Żurawski, "Management of Transmission and Distribution Infrastructure as Part of the National Energy Security System," *Eur. Res. Stud. J.*, vol. 28, no. 4, pp. 855–872, 2025.

[3] A. F. Ihsan, S. Uttunggadewa, S. D. Rahmawati, I. Giovanni, and S. N. Himawan, "Performance Analysis of Deep Learning Implementation in Operational Condition Forecasting of a Gas Transmission Pipeline Network.," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 13, no. 4, 2023.

[4] Z. Yang *et al.*, "Resilience of natural gas pipeline system: a review and outlook," *Energies*, vol. 16, no. 17, p. 6237, 2023.

[5] K. Raghunandan, "Supervisory control and data acquisition (SCADA)," in *Introduction to Wireless Communications and Networks: A Practical Perspective*, Springer, 2022, pp. 321–337.

[6] L. Liu, J. Liang, L. Ma, H. Zhang, Z. Li, and S. Liang, "Gas pipeline flow prediction model based on LSTM with grid search parameter optimization," *Processes*, vol. 11, no. 1, p. 63, 2022.

[7] C. Ji and W. Sun, "A review on data-driven process monitoring methods: Characterization and mining of industrial data," *Processes*, vol. 10, no. 2, p. 335, 2022.

[8] F. Wu, J. Jiang, X. Peng, L. Teng, X. Meng, and J. Li, "Influence of natural gas composition and operating conditions on the steady-state performance of dry gas seals for pipeline compressors," *Lubricants*, vol. 12, no. 6, p. 217, 2024.

[9] N. I. A. Kader, U. K. Yusof, M. N. A. Khalid, and N. R. N. Husain, "A review of long short-term memory approach for time series analysis and forecasting," in *International Conference on Emerging Technologies and Intelligent Systems*, 2022, pp. 12–21.

[10] I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent neural networks: A comprehensive review of architectures, variants, and applications," *Information*, vol. 15, no. 9, p. 517, 2024.

[11] Z. Wang, M. Xu, J. Hao, W. Li, and S. Cheng, "Research on explainable BiGRU deep learning framework for short term load forecasting in smart power systems," *Discov. Artif. Intell.*, vol. 5, no. 1, p. 407, 2025.

[12] M. Louisa, G. Darmawan, and B. Tantular, "Enhancing Stock Price Forecasting with CNN-BiGRU-Attention: A Case Study on INDY," *Mathematics*, vol. 13, no. 13, p. 2148, 2025.

[13] M. Alam, "Multi-directional gated recurrent unit and convolutional neural network for load and energy forecasting: A novel hybridization," *AIMS Math.*, vol. 8, no. 9, pp. 19993–20017, 2023.

[14] N. Zaini, L. W. Ean, A. N. Ahmed, and M. A. Malek, "A systematic literature review of deep learning neural network for time series air quality forecasting," *Environ. Sci. Pollut. Res.*, vol. 29, no. 4, pp. 4958–4990, 2022.

[15] J. M. Tan *et al.*, "Hyperparameter optimization: Classics, acceleration, online, multi-objective, and tools," *Math. Biosci. Eng.*, vol. 21, no. 6, pp. 6289–6335, 2024.

[16] W. Deng, X. Xin, R. Song, X. Yang, W. Wang, and G. Yu, "A time series forecasting method for oil production based on Informer optimized by Bayesian optimization and the hyperband algorithm (BOHB)," *Comput. Chem. Eng.*, vol. 197, p. 109068, 2025.

[17] D. V. Trivedi and S. Patel, "An analysis of GRU-LSTM hybrid deep learning models for stock price prediction," *Int. J. Sci. Res. Sci. Eng. Technol*, vol. 9, no. 3, pp. 47–52, 2022.

[18] T. Wang and X. Jiang, "ARIMA-BIGRU Stock Forecast Model Based on Bayesian Optimization," in *Proceedings of the 2025 3rd International Conference on Communication Networks and Machine Learning*, 2025, pp. 131–136.

[19] T. A. Nguyen and C. K. Ha, "Long-Term Peak Load Forecasting with CNN Via Two-Stage Random–Bayesian Hyperparameter Optimization," *J. Robot. Control*, vol. 6, no. 6, pp. 2774–2784, 2025.

[20] M. Shantal, Z. Othman, and A. A. Bakar, "A novel approach for data feature weighting using correlation coefficients and min–max normalization," *Symmetry (Basel).*, vol. 15, no. 12, p. 2185, 2023.

[21] M. A. M. Fargalla, W. Yan, and T. Wu, "Attention-based bi-directional gated recurrent unit (Bi-GRU) for sequence shale gas production forecasting," in *International Petroleum Technology Conference*, 2024, p. D021S040R001.

[22] H. Wang and K. Yang, "Bayesian optimization," in *Many-Criteria Optimization and Decision Analysis: State-of-the-Art, Present Challenges, and Future Perspectives*, Springer, 2023, pp. 271–297.

[23] M. Binois and N. Wycoff, "A survey on high-dimensional Gaussian process modeling with application to Bayesian optimization," *ACM Trans. Evol. Learn. Optim.*, vol. 2, no. 2, pp. 1–26, 2022.

[24] O. E. Obisesan, "Machine learning models for prediction of meteorological variables for weather forecasting," *Int J Env. Clim Chang.*, vol. 14, no. 1, pp. 234–252, 2024.

[25] A. Mystakidis, P. Koukaras, N. Tsalikidis, D. Ioannidis, and C. Tjortjis, "Energy forecasting: A comprehensive review of techniques and technologies," *Energies*, vol. 17, no. 7, p. 1662, 2024.