Advanced Energy Systems and Intelligent Geoscience (AESIG)
https://www.mospbs.com/journal/aesig

• *Article* •

Study on High-Performance Prediction of Wellbore Multiphase Flow Pressure Drop Distribution Based on Artificial Neural Networks

Xi Ouyang[1,2], Xiang Rao[1,2,3*]

[1] School of Petroleum Engineering, Yangtze University, Wuhan 430100, China.

[2] State Key Laboratory of Low Carbon Catalysis and Carbon Dioxide Utilization (Yangtze University), Wuhan 430100, China.

[3] Western Research Institute, Yangtze University, Karamay 834000, China.

[*]**Corresponding Author:** Xiang Rao **Email:** raoxiang0103@163.com

Received: 02 February 2026 **Accepted:** 07 February 2026

Abstract: This paper constructs a proxy model to predict the pressure drop distribution in wellbore multiphase flow based on artificial neural networks (ANNs). 10,000 sets of high-quality samples involving 15 parameters involving wellbore multiphase flow and covering diverse working conditions were generated based on the Beggs-Brill algorithm, which ensures the physical consistency of data and the comprehensiveness of working conditions; two ANN models were constructed to realize the bidirectional nonlinear mapping from wellhead parameters to bottom-hole pressure and from bottom-hole parameters to wellhead pressure, respectively. Case verification results show that the mean relative error of the forward ANN model is only 0.10%, that of the reverse ANN model is 0.15%, and the prediction time for a single sample is less than 0.002 seconds for both models. The research results provide an efficient and reliable technical tool for the dynamic production regulation of oil and gas fields and the rapid evaluation of multi-well development schemes, helping to improve the recovery efficiency and economic benefits of oil and gas exploitation.

Keywords: Wellbore multiphase flow; Pressure drop prediction; Neural network; Beggs-Brill algorithm; Coupled simulation; Machine learning

1 Introduction

The prediction of wellbore multiphase flow parameters (including pressure drop, flow pattern, flow rate, lithology, etc.) serves as the core technical cornerstone for the entire process of oil and gas field development. It is directly linked to the optimization of tubing selection, the improvement of artificial lift system efficiency, safety risk prevention and control, and the goal of maximizing development benefits. The prediction accuracy not only affects the estimation of single well production and the extension of tubing service life, but also plays a decisive role in operational safety and cost control, being a key factor determining the success or failure of engineering projects (Mahmoud et al., 2021, Rao et al., 2024, Rao et al., 2022). At present, oil and

DOI: https://doi.org/10.58244/aesig.263567

gas development is advancing toward complex working conditions such as deep formations (burial depth > 4000 m), deepwater environments (water depth > 1500 m), transportation in special geometric structures (annuli, microchannels), and non-Newtonian fluid transportation (polymer solutions, sand-laden fluids). Under such conditions, multiphase flow exhibits complex characteristics including strong nonlinear coupling, transient flow pattern transition, and multi-field coupling of flow field-temperature field-pressure field (Zheng et al., 2024, Santoso et al., 2016, Osgouei et al., 2015), posing severe challenges to parameter prediction technologies.

The academic community has carried out systematic research in this field, evolving three core technical approaches: traditional theoretical mechanism derivation, mechanism-data driven fusion, and pure data driven modeling. Among them, traditional mechanism models rely on experimental data and theoretical deduction to establish the fundamental framework for multiphase flow prediction: Beggs and Brill (1973) developed a pressure drop calculation model covering segregated flow, intermittent flow and dispersed flow based on 584 sets of full-inclination experimental data, quantifying the influence of inclination angle on liquid holdup for the first time, which has become a classic benchmark for inclined pipe design; Orkiszewski (1967) optimized the flow pattern classification system to include four typical flow patterns such as bubbly flow and slug flow, and field verification with 148 wells showed a mean error of only 0.8%; Petalas and Aziz (2000) proposed a generalized model applicable to all geometric structures and fluid properties, and validation with 5951 sets of data indicated that the prediction error of liquid holdup for 62% of the data was controlled within ±15%; Cioncolini and Thome (2017) introduced the momentum Weber number and Froude number based on 6291 sets of experimental data, which effectively addressed the problem of liquid film asymmetry in annular flow and horizontal pipes, with the prediction error of 70% of the data remaining within ±15%.

With the increasing complexity of working conditions, the shortcomings of traditional mechanism models in adapting to transient flow, small sample scenarios and strong coupling effects have gradually become prominent. Against this background, the innovative approach of mechanism-data driven fusion has emerged as the times require (Rao et al., 2025a, 2025b, 2025c, Liu et al., 2025), becoming the core direction to break through the bottleneck of prediction under complex working conditions: Zheng et al. (2024) proposed a Knowledge-Guided Machine Learning (KGML) framework, which deeply integrates multiphase flow physical models with the XGBoost algorithm. When applied to the prediction of bottom-hole pressure in deepwater gas fields of the South China Sea, the Mean Absolute Percentage Error (MAPE) was as low as 0.11%, effectively solving the problem of accurate prediction in small sample scenarios; Sun et al. (2018) established a transient multiphase flow model for acid gas invasion in high-temperature and high-pressure gas wells, fully considering the interphase mass and heat transfer and wellbore-formation coupling effects, clarifying the abrupt change characteristics of gas solubility at the well depth of 500-1000 m, and revealing the inherent mechanism of concealed and sudden kicks; Pure data driven models focus on algorithm optimization and in-depth adaptation to actual scenarios (Rao et al., 2024, Rao, 2024a, Rao, 2024b), further expanding the application boundary of multiphase flow parameter prediction: Ling and Hibiki (2025) systematically sorted out the application progress of machine learning in gas-liquid flow pattern identification through

a review study, covering data preprocessing methods, comparison of algorithm performance and analysis of generalization ability, which provided a clear technical roadmap for subsequent research; Shijo and Behera (2023) compared four mainstream algorithms such as LightGBM and MLP, and confirmed that XGBoost exhibited the best performance in pressure drop prediction for dense-phase transportation, offering a reference for algorithm selection in engineering practice; Bello and Asafa (2014) constructed a soft sensor using a functional network to realize the simultaneous prediction of bottom-hole pressure and temperature in multiphase flow wells, with the coefficient of determination (R^2) of both training and test sets exceeding 0.99, maintaining an excellent level of prediction accuracy.

Despite the phased achievements made in existing research, there are still three prominent bottlenecks: First, the scope of working condition adaptation is narrow. Domestic models are mostly developed for single working conditions of oilfields in specific regions, and the accuracy of classic foreign models declines significantly in special scenarios such as non-Newtonian fluid flow and eccentric annuli. In addition, most data driven models rely on training with a single dataset and lack the generalization ability for the transition of different flow patterns and cross-diameter adaptation. Second, the consistency between sample quality and physical laws is insufficient. The sample size in some studies is only a few hundred sets, and outliers under extreme working conditions (such as pressure abrupt change data during sudden gas invasion) are not fully processed. Lacking the constraints of physical laws such as multiphase flow momentum conservation and interphase mass transfer, pure data driven models tend to generate prediction results that contradict the mechanism under unseen extreme working conditions (e.g., predicted pressure drop lower than the actual value leading to the risk of tubing overload). Third, the functional coverage and collaborative optimization capability are inadequate. Most existing achievements focus on the prediction of a single parameter, with insufficient integration of practical engineering demands such as drilling-multiphase flow coupled regulation and high gas cut flow measurement, making it difficult to support integrated production regulation and decision-making.

This paper conducts research aiming at the rapid prediction of wellbore multiphase flow pressure drop: a high-quality sample dataset considering the physical correlation of parameters is generated based on the Beggs-Brill algorithm to ensure the authenticity and coverage of samples; two sets of models are constructed to realize the bidirectional mapping from wellhead parameters to bottom-hole pressure and from bottom-hole parameters to wellhead pressure, respectively; the accuracy and efficiency of the proposed models and the Beggs-Brill algorithm are systematically compared from three dimensions: training process monitoring, test set evaluation and new sample verification. The research results are expected to provide a new tool for the rapid pressure prediction and dynamic optimization in oil and gas field production, helping to improve the exploitation efficiency and economic benefits.

2 Methodology

2.1 Sample Generation

The calculation of pressure drop for the multiphase flow of crude oil, natural gas and water in the wellbore is the core of oil and gas exploitation. As the most widely used semi-empirical method, the Beggs-Brill algorithm is based on the law of conservation of momentum. Through the

segmented iterative logic of flow pattern classification, liquid holdup calculation and pressure gradient decomposition, it can accurately solve the pressure drop along the entire wellbore, providing a reliable basis for the generation of sample labels. To ensure the reliability of subsequent numerical simulation and model verification, sample generation follows the principles of comprehensive working condition coverage, consistent physical logic and controllable label accuracy. A complete process was established based on the Beggs-Brill algorithm, and 10,000 sets of valid samples were finally generated.

Fifteen key variables affecting multiphase flow pressure were selected as the sample input parameters, covering three categories: wellhead operating conditions, pipeline geometric characteristics and fluid physical properties. Their value ranges were determined based on field actual operating conditions and literature research, covering typical scenarios such as low, medium and high flow rates, different well depths and inclination angles. The initial samples were generated by the method of uniform distribution combined with Gaussian noise. Invalid parameter combinations were eliminated through mandatory boundary constraints, and then the parameter correlation was revised according to the physical laws of multiphase flow. The core constraint was set that the production gas-oil ratio (PGOR) is not less than the solution gas-oil ratio (SGOR), so as to ensure the rationality and diversity of the samples.

The sample labels were obtained by solving with the Beggs-Brill algorithm, and the specific process is as follows: After segmenting the wellbore, the liquid volume fraction and Froude number were calculated based on the pressure and temperature at the starting point of each segment to determine the flow pattern; the actual liquid holdup was solved by the idea of horizontal pipe benchmark combined with inclination angle correction; the mixture density, gravity gradient and friction gradient were then calculated to iteratively obtain the pressure drop of this segment; the cumulative pressure drop of all segments was taken as the total pressure drop along the entire wellbore (label value). Invalid samples with pressure less than 0 were eliminated during the iteration, and a sample dataset with comprehensive working conditions and consistent physical logic was finally formed.

2.2 Model 1

Model 1 takes 15 core parameters(Table 1) at the wellhead side as inputs and the bottom-hole pressure as the single output. Its core objective is to establish a strong nonlinear mapping relationship between multi-parameter inputs and bottom-hole pressure, replacing traditional iterative algorithms to realize engineering-level rapid and accurate prediction. The model design strictly follows the principle of data adaptability, structural optimization, efficient training and comprehensive evaluation, and encompasses four core modules: data preprocessing, network structure design, training parameter configuration and evaluation system construction, which ensures the dual achievement of prediction accuracy and engineering practicability.

Model 1 adopts the architecture of the Artificial Neural Network (ANN). Combining with the characteristics of strong nonlinearity and multi-parameter coupling in wellbore multiphase flow pressure prediction, a three-layer structure consisting of an input layer, hidden layers and an output layer is optimized and determined through multiple rounds of cross-validation. The network structure is shown in Figure 1.

Table 1 Setting Range of Core Sample

Table 1 Setting Range of Core Sample Parameters

Parameter Name	Setting Range	Parameter Name	Setting Range
Wellhead Pressure	2.5~15 MPa	Relative Density of Natural Gas	0.65~0.85
Inner Diameter of Tubing	0.05~0.08 m	Crude Oil Volume Factor	1.0~1.05
Wellbore Length	500~3000 m	Crude Oil Surface Tension	0.01~0.02
Wellbore Inclination Angle	-90~90°	Crude Oil Viscosity	10~100mPa·s
Starting Point Temperature	20~90℃	Natural Gas Viscosity	0.01~0.15 mPa·s
Crude Oil Volumetric Flow Rate	100~500 m^3/d	Gas Compressibility Factor	0.8~1.0
Production Gas-Oil Ratio	30~70 m^3/m^3	Solution Gas-Oil Ratio	10~100 m^3/m^3
Crude Oil Density	800~850 kg/m^3		

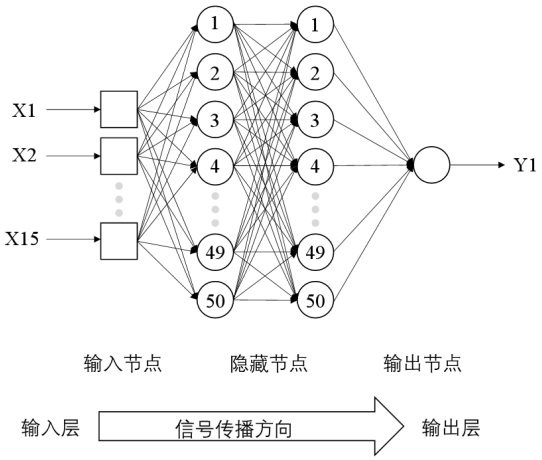

Fig. 1 Schematic Diagram of the Network Structure

Parameters

The input layer is configured with 15 neurons, corresponding one-to-one with the 15 input parameters. Each neuron directly receives the normalized value of the corresponding feature, realizing lossless data transmission. The number of neurons is fully matched with the dimension of input parameters to ensure the complete input of feature information. The hidden layers adopt a bottleneck structure design, with the number of neurons set to 50 and 50 in sequence. The core advantage of this structure lies in the reduction in the number of neurons in the middle layer, which forces the model to discard redundant information and focus on extracting the low-dimensional core features that affect bottom-hole pressure. In contrast, the expansion of the number of neurons in the two side layers can enhance the deep mapping capability of features, achieving an optimal balance between fitting complex nonlinear relationships and controlling the computational complexity of the model. The output layer is equipped with 1 neuron corresponding to the predicted value of bottom-hole pressure, and a linear activation function is adopted. This ensures that the denormalized output values can cover the pressure range in actual engineering (2.5~15 MPa), avoiding the value truncation problem caused by nonlinear activation functions.

Signal transmission between layers is realized through a fully connected approach, and the Tanh function is adopted as the activation function for the hidden layers, with its expression given as follows:

$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{1}$$

This function can map neuron outputs to the interval [−1,1], which not only enhances the model's fitting ability for the coupled nonlinear relationships among multiple parameters but

also effectively alleviates the vanishing gradient problem during training, thus adapting to the complex distribution characteristics of industrial data. The mean squared error (MSE) is adopted as the loss function of the model, with its expression given as follows:

$$\textbf{Loss} = \frac{1}{N}\sum_{k=1}^{N}(y_k - \hat{y}_k)^2 \tag{2}$$

where N denotes the number of samples; y_k represents the actual bottom-hole pressure of the k-th sample; and $\hat{y}_k$ is the corresponding predicted value. MSE has high sensitivity to samples with large errors, which aligns with the stringent requirements for pressure prediction accuracy in engineering scenarios and can effectively constrain prediction deviations under extreme working conditions. Parameter statistics show that the total number of parameters of the model is 3401, which effectively controls the model size while ensuring fitting capability, thus facilitating the rapid deployment and real-time prediction in engineering applications.

Table 2 presents the parameter settings for the model training in this paper.

Furthermore, to quantify the predictive performance of the model in a comprehensive and objective manner, a multi-dimensional evaluation index system covering absolute error, relative error and goodness of fit is established, which not only reflects the overall fitting accuracy of the model but also takes into account the actual deviation requirements in engineering applications. The calculation formulas for each index are given as follows:

$$\mathrm{MSE} = \frac{1}{N_{\text{test}}}\sum_{k=1}^{N_{\text{test}}}(y_{\text{pred},k} - y_{\text{actual},k})^2 \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} \tag{4}$$

$$\mathrm{MAE} = \frac{1}{N_{\text{test}}}\sum_{k=1}^{N_{\text{test}}}|y_{\text{pred},k} - y_{\text{actual},k}| \tag{5}$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N_{\text{test}}}(y_{\text{actual},k} - y_{\text{pred},k})^2}{\sum_{k=1}^{N_{\text{test}}}(y_{\text{actual},k} - \overline{y}_{\text{actual}})^2} \tag{6}$$

Where N_{test} is the number of samples in the test set; $y_{\text{pred},k}$ is the predicted bottom-hole pressure of the k-th sample; $y_{\text{actual},k}$ is the actual bottom-hole pressure of the k-th sample; and

Table 2 Key Training Parameters of the Model

Training Parameters	Parameter Values/Settings
Optimizer	Adam Adaptive Optimizer
Initial Learning Rate	0.0005
Maximum Number of Iterations	10,000 iterations
Learning Rate Scheduling Strategy	ReduceLROnPlateau (factor=0.9,patience=1000)
Gradient Clipping	max_norm=1.0
Loss Function	Mean Squared Error (MSE)
Hidden Layer Structure	50 neurons in Layer 1, 50 neurons in Layer 2
Activation Function	Tanh for hidden layers, Linear for output layer
Random Seed	1

$\overline{y}_{\text{actual}}$ is the mean value of the actual bottom-hole pressure in the test set.

2.3 Model 2

The core difference between Model 2 and Model 1 lies in the reverse mapping of inputs and outputs: the former takes the bottom-hole pressure and 14 auxiliary parameters as inputs and the wellhead pressure as the single output, with its core objective being to establish a reverse nonlinear mapping relationship between bottom-hole operating conditions and wellhead pressure. This model is adapted to the engineering scenario of rapidly inverting wellhead pressure based on known formation and bottom-hole parameters (e.g., the indirect prediction of wellhead parameters for remote well sites where direct measurement is unfeasible). On the basis of inheriting the basic framework of Model 1, the design of Model 2 is adaptively optimized for the reverse adjustment of inputs and outputs, so as to ensure the accuracy and engineering practicability of reverse pressure prediction.

Model 2 adopts the identical architecture as Model 1. The core reasons for this design are as follows: the input and output dimensions remain unchanged, the task type is still continuous value regression, and the nonlinear characteristics of multiphase flow pressure transmission are consistent in both forward and reverse prediction tasks. Thus, the network topology requires no adjustment to adapt to the reverse mapping task. The network structure is shown in Figure 1.

The core training parameters (e.g., optimizer, learning rate, epochs) are kept consistent with those of Model 1; only the log recording and model saving strategies are adjusted in accordance with the output characteristics of the reverse task. This adjustment is made to ensure the fairness of experimental comparison between the forward and reverse prediction tasks, and the specific parameters are listed in Table 2.

3 Case Studies

3.1 Performance Analysis of Model 1

After the model was trained, performance analysis was conducted from three dimensions: training process stability, test set generalization accuracy, and new working condition sample verification, to systematically evaluate its accuracy and engineering applicability.

Model 1 completed 10,000 rounds of full-batch iteration with a total time consumption of approximately 2 minutes, and no anomalies such as gradient explosion, numerical oscillation or convergence stagnation occurred. As can be seen from the training and test loss curves (Figure 2), its convergence process exhibits distinct characteristics of high efficiency and stability: the training loss and test loss decreased rapidly and synchronously in the initial stage, quickly capturing the coupling relationships between core parameters (e.g., wellhead pressure, tubing inner diameter, pipeline inclination angle) and bottom-hole pressure; with the progress of itera-

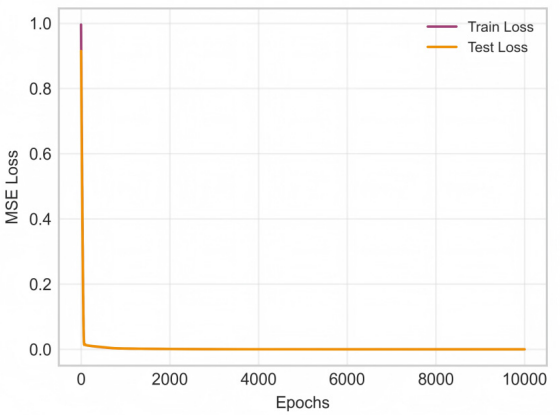

Fig. 2 Training and Test Loss Curves of Model 1

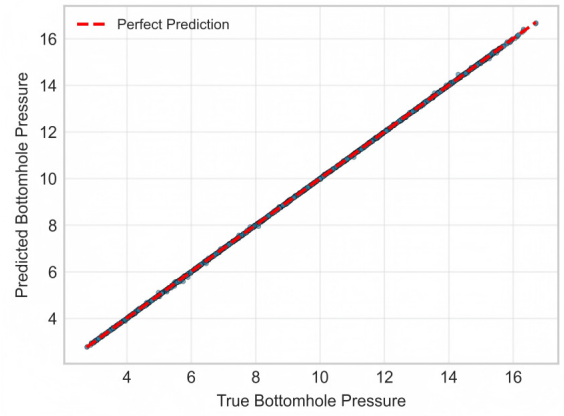

Fig. 3 Scatter Plot of True and Predicted Bottom-Hole Pressure Values for the Model 1 Test Set

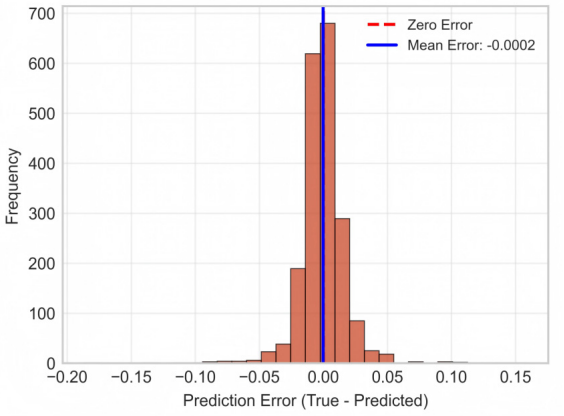

Fig. 4 Histogram of Prediction Error Distribution for the Model 1 Test Set

tion, the difference between the two remained in an extremely small range without any overfitting tendency; in the final stage, both the training loss and test loss converged stably to 0.0000, a result attributed to the high reliability of the sample data and the adaptability of the model architecture.

Based on 2000 test samples, the generalization ability of Model 1 was comprehensively verified using dual indicators of goodness of fit and error distribution. As shown in Figure 3, the predicted values of all test samples are highly consistent with the true values calculated by the Beggs-Brill algorithm and closely align with the y=x ideal fitting line, indicating that the model has accurately reproduced the physical laws governing multiphase flow pressure transmission. Figure 4 shows that the prediction errors of the test set follow a normal distribution centered at 0.00, with the error range concentrated in -0.2~0.15 MPa.

Table 3 Prediction Results and Error Statistics of Model 1 for New Samples

Sample Number	Core Input Parameters: Wellhead Pressure (MPa)/Wellbore Length (m)/ Wellbore Inclination Angle (°)	Calculated Value (MPa)	Predicted Value (MPa)	Absolute Error (MPa)	Relative Error (MPa)
1	12.6292/1842.3531/88.78	13.0777	13.0922	0.0145	0.11
2	5.7821/2857.6352/27.46	6.1302	6.1297	0.0005	0.01
3	5.5445/2837.7148/-37.76	6.2100	6.2087	0.0013	0.02
4	11.1562/1282.0933/45.59	13.2397	13.1769	0.0628	0.47
5	7.9828/2923.8266/-37.24	8.2538	8.2499	0.0049	0.06
6	9.7987/1974.9808/-85.53	10.1530	10.1369	0.0161	0.16
7	9.4293/2928.5495/79.68	10.0760	10.0843	0.0083	0.08
8	14.8952/1711.7856/-27.61	15.3257	15.3210	0.0047	0.03
9	11.2719/1289.4075/-22.25	11.6638	11.6586	0.0052	0.04
10	11.8369/1587.3410/-75.76	12.2193	12.2110	0.0083	0.07

There are no systematic deviations or extreme outliers in the results, which fully demonstrates the high reliability and authenticity of the model's predictions.

To verify the generalization ability and prediction reliability of Model 1 under unknown working conditions, 10 sets of new samples covering diverse and complex working conditions were selected for validation in this study. The sample parameters span a full range, including wellhead pressure of 5.54–14.90 MPa, pipeline length of 1282–2929 m, and pipeline inclination angle of -85.53°–79.68°. These samples cover not only conventional working conditions with gentle inclination angles, but also edge working conditions that are prone to accuracy fluctuations in engineering practice, such as high inclination angle (79.68°), large negative inclination angle (-85.53°), and long pipeline (2928.55 m). Thus, they can fully reflect the adaptability of the model in actual complex engineering scenarios.

As can be seen from Table 3, Model 1 exhibits extremely high prediction accuracy: the mean relative error of the 10 samples is merely 0.10%, with the minimum relative error being only 0.01% and the corresponding absolute error just 0.0005 MPa, achieving almost perfect consistency with the calculated values from the Beggs-Brill algorithm. The error distribution shows distinct uniformity and controllability, with the relative errors generally concentrated in the range of 0.01%–0.47% and no extreme outliers observed. The ratio of RMSE to MAE is close to 1, indicating the absence of systematic offset in the errors, which are all acceptable minor random deviations in the process of multiphase flow transmission. Even under the extreme working conditions of a wellbore inclination angle close to -90° (Sample 6) and a high inclination angle exceeding 70° (Sample 7), the relative errors are only 0.16% and 0.08% respectively, with no accuracy degradation caused by extreme parameter values. This fully verifies the model's strong adaptability to complex working conditions. Notably, the maximum relative error of 0.47% (Sample 4) corresponds to an absolute error of just 0.0628 MPa, and this deviation magnitude has no impact whatsoever on pressure judgment and decision-making in practical engineering scenarios such as well control operations and gas lift design. Meanwhile, on the premise of maintaining high accuracy, the model has achieved a breakthrough improvement in prediction efficiency, with the prediction time for a single sample being only 0.0017 seconds. It successfully resolving the core contradiction of traditional algorithms of high accuracy but low efficiency, and achieves the dual advantages of high accuracy and high speed. The model can meet the timeliness requirements of scenarios such as real-time regulation and dynamic optimization in oil and gas field development, thus providing a technical solution with both reliability and practicability for the forward prediction of wellbore pressure.

3.2 Performance Analysis of Model 2

Model 2 focuses on the reverse nonlinear mapping task, which is mainly adapted to the engineering scenario where direct measurement of wellhead parameters at remote well sites is unfeasible, enabling the rapid and accurate acquisition of wellhead pressure through AI inversion. This model adopts the identical architecture as Model 1 and completed full-batch iterative training in a GPU computing environment with a total time consumption of approximately 2 minutes, with no anomalies such as gradient oscillation, numerical overflow or convergence stagnation occurring. As can be seen from the training and test loss curves (Figure 5), the initial loss was 0.88; the loss de-

creased rapidly during training and the difference from the test loss remained consistently below 0.0001. The learning rate dropped to 0.00044 at last, and both the training loss and test loss converged to 0.0000 and remained stable consistently, with the gradient norm stabilizing below 0.00009. This fully verifies the model's adaptability to the reverse mapping task.

Based on 2000 test samples, the generalization ability of Model 2 for reverse mapping was comprehensively verified using dual indicators of goodness of fit and error distribution. As shown in Figure 6, the predicted values of all test samples are closely clustered around the y=x ideal fitting line with the true wellhead pressure values calculated by the Beggs-Brill algorithm, with no obvious dispersion or deviation observed. This indicates that the model has accurately reproduced the physical laws governing the transmission of formation pressure to the wellhead through wellbore multiphase flow. Figure 7 shows that the prediction errors of the test set follow a symmetric bell-shaped normal distribution centered at 0.00, with the error range concentrated in -0.15–0.15 MPa. There are no extreme outliers exceeding this range, the frequency distribution of errors on both sides is balanced with no systematic deviations, and the overall mean error is close to 0. This fully confirms the high reliability and authenticity of the model in reverse pressure calculation.

To verify the generalization performance and prediction reliability of Model 2 in the reverse mapping task, the same 10 sets of new samples were used for prediction and validation, with detailed comparisons presented in Table 4.

As can be seen from Table 4, Model 2 demonstrates a prediction accuracy far exceeding the industry level, with a mean relative error of merely 0.15% for the 10 samples. The minimum

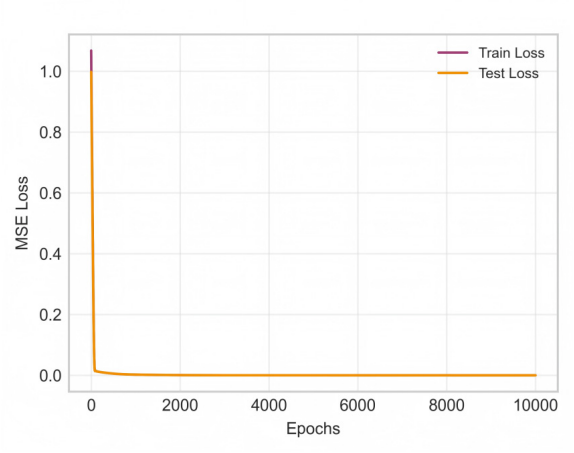

Fig. 5 Training and Test Loss Curves of Model 2

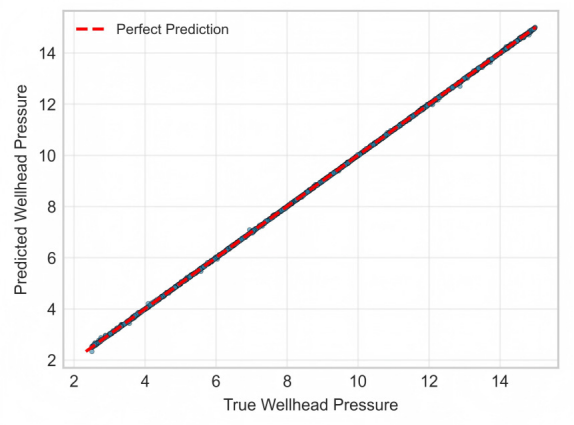

Fig. 6 Scatter Plot of True and Predicted Bottom-Hole Pressure Values for the Model 2 Test Set

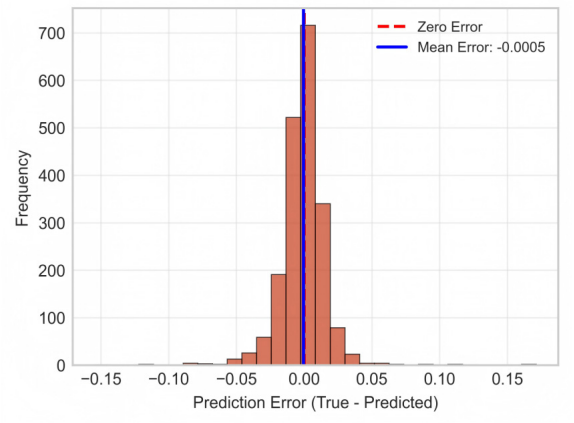

Fig. 7 Histogram of Prediction Error Distribution for the Model 2 Test Set

Table 4 Prediction Results and Error Statistics of Model 2 for New Samples

Sample Number	Core Input Parameters: Bottom-Hole Pressure (MPa)/Wellbore Length (m)/ Wellbore Inclination Angle (°)	True Well-head Pressure Value (MPa)	Predicted Well-head Pressure Value (MPa)	Absolute Error (MPa)	Relative Error (%)
1	13.0777/1842.3531/88.78	12.6292	12.6190	0.0102	0.08
2	6.1302/2857.6352/27.46	5.7821	5.7811	0.001	0.02
3	6.2099/2837.7148/-37.76	5.5445	5.5450	0.0005	0.01
4	13.2397/1282.0933/45.59	11.1562	11.2331	0.0769	0.69
5	8.2538/2923.8266/-37.24	7.9828	7.9850	0.0022	0.03
6	10.1529/1974.9808/-85.53	9.7987	9.8234	0.0247	0.25
7	10.0760/2928.5495/79.68	9.4293	9.4218	0.0075	0.08
8	15.3257/1711.7856/-27.61	14.8952	14.8825	0.0127	0.09
9	11.6638/1289.4075/-22.25	11.2719	11.2685	0.0034	0.03
10	12.2193/1587.3410/-75.76	11.8369	11.8242	0.0127	0.11

relative error is only 0.01%, corresponding to an absolute error of just 0.0005 MPa, achieving a high degree of consistency with the calculated values from the Beggs-Brill algorithm. The error distribution exhibits distinct uniformity and controllability: the relative errors are generally concentrated in the range of 0.01%–0.69% with no extreme outliers observed, and the ratio of RMSE to MAE is close to 1. This indicates the absence of systematic offset in the errors, which are all acceptable minor random deviations in the process of multiphase flow reverse solution, and also confirms that the model has no serious prediction errors for individual samples. Even under the extreme working conditions of a wellbore inclination angle close to -90° (Sample 6) and a high inclination angle exceeding 70° (Sample 7), the relative errors are only 0.25% and 0.08% respectively, with no accuracy degradation caused by extreme parameter values. This fully verifies the model's strong adaptability to complex working conditions. The maximum relative error of 0.69% (Sample 4) corresponds to an absolute error of just 0.0769 MPa, and this deviation magnitude has no impact whatsoever on the judgment and decision-making of wellhead parameters in practical engineering scenarios such as well control operations and gas lift design. Meanwhile, on the premise of maintaining high accuracy, the model has achieved a breakthrough improvement in prediction efficiency, with the prediction time for a single sample being only 0.0019 seconds, thus providing a technical solution with both reliability and practicability for the reverse prediction of wellbore pressure.

4 Conclusions

Focusing on the problem of forward and reverse prediction of wellbore multiphase flow pressure, this paper takes the Beggs-Brill algorithm as the core support to complete sample generation, model construction and performance verification, and develops a technical solution for wellbore pressure prediction with both high accuracy and high efficiency. The main conclusions are as follows:

The sample generation process constructed

based on the Beggs-Brill algorithm strictly adheres to the principles of comprehensive working condition coverage, consistent physical logic and controllable label accuracy. Fifteen core influencing parameters were selected to generate initial samples by combining uniform distribution with Gaussian noise, and the rationality of samples was guaranteed through boundary constraints and physical law correction, finally generating 10,000 sets of valid samples. This sample set covers typical and edge working conditions such as low, medium and high flow rates, different well depths and inclination angles, and takes the total wellbore pressure drop accurately solved by the Beggs-Brill algorithm as the label. It provides a highly reliable data foundation for the subsequent training and verification of neural network models, and addresses the issue of insufficient sample quality for multiphase flow prediction models.

Two ANN models were constructed to adapt to the requirements of forward and reverse pressure prediction for wellbore multiphase flow, both adopting a three-layer architecture of input layer- hidden layer-output layer. By optimizing the bottleneck structure of the hidden layer and the configuration of activation functions, an optimal balance was achieved between fitting nonlinear relationships and controlling model complexity. Model 1 (with wellhead parameters as input and bottom-hole pressure as output) and Model 2 (with bottom-hole parameters and auxiliary parameters as input and wellhead pressure as output) featured stable training processes without anomalies such as gradient explosion and convergence stagnation, and both their training and test losses converged to extremely small values, demonstrating excellent model adaptability.

Performance verification results show that both models exhibit extremely high prediction accuracy and generalization ability. Model 1 achieved a mean relative error of merely 0.10% with a single-sample prediction time of 0.0017 s, and Model 2 had a mean relative error of 0.15% with a single-sample prediction time of 0.0019 s. Both models can be accurately adapted to engineering edge working conditions such as high inclination angles and extreme negative inclination angles without any accuracy degradation. Compared with the traditional Beggs-Brill mechanism-based algorithm, the proposed models have greatly improved prediction efficiency while maintaining the same high accuracy, successfully resolving the core contradiction of traditional algorithms of high accuracy but low efficiency.

In summary, the sample generation process constructed in this paper is reliable and feasible. The two ANN models realize fast forward prediction and reverse inversion of wellbore pressure respectively, which can meet various engineering requirements in oil and gas field development such as real-time regulation, dynamic optimization and indirect measurement of parameters at remote well sites. This study provides an efficient and accurate new approach for the prediction of wellbore multiphase flow pressure, and demonstrates important engineering application value and popularization prospects.

Acknowledgement

The author sincerely thanks all the financial support.

Funding Statement

The General Program of the National Natural Science Foundation of China (Grant No. 52574028)

The Youth Science Fund Program (Category A) of the National Natural Science Foundation of

China (Grant No. 52525403)

The "Science and Technology Innovation Team" Program of the Xinjiang Uygur Autonomous Region (Grant No. 2024TSYCTD0018).

Author Contributions

The authors confirm their responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Availability of Data and Materials

None.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Beggs, D. H., & Brill, J. P. (1973). A study of two-phase flow in inclined pipes. Journal of Petroleum technology, 25(05), 607-617.

[2] Bello, O., & Asafa, T. (2014, April). A functional networks softsensor for flowing bottomhole pressures and temperatures in multiphase flow production wells. In SPE Intelligent Energy International Conference and Exhibition (pp. SPE-167881). SPE.

[3] Cioncolini, A., & Thome, J. R. (2017). Pressure drop prediction in annular two-phase flow in macroscale tubes and channels. International Journal of Multiphase Flow, 89, 321-330.

[4] Ling, C., & Hibiki, T. (2025). Advances in machine learning for data-driven classification of gas–liquid flow regimes. International Communications in Heat and Mass Transfer, 169, 109582.

[5] Liu, Y., Rao, X., He, X., Fu, Q., & Hoteit, H. (2025, September). Boundary-Integral Type Neural Network (BINN) for Flow Problems in Homogeneous Reservoirs. In SPE Middle East Oil and Gas Show and Conference (p. D011S024R002). SPE.

[6] Mahmoud, A. A., Elkatatny, S., & Al-AbdulJabbar, A. (2021). Application of machine learning models for real-time prediction of the formation lithology and tops from the drilling parameters. Journal of Petroleum Science and Engineering, 203, 108574.

[7] Osgouei, R. E., Ozbayoglu, A. M., Ozbayoglu, E. M., Yuksel, E., & Eresen, A. (2015). Pressure drop estimation in horizontal annuli for liquid–gas 2 phase flow:

[8] Orkiszewski, J. (1967). Predicting two-phase pressure drops in vertical pipe. Journal of Petroleum technology, 19(06), 829-838.

[9] Petalas, N., & Aziz, K. (2000). A mechanistic model for multiphase flow in pipes. Journal of Canadian petroleum technology, 39(06).

[10] Rao, X., Liu, Y., & Shen, Y. (2025). Quantum-Classical Physics-Informed Neural Networks for Solving Reservoir Seepage Equations. arXiv preprint arXiv:2512.03923.

[11] Rao, X., Liu, Y., He, X., & Hoteit, H. (2025). Physics-informed Kolmogorov–Arnold networks to model flow in heterogeneous porous media with a mixed pressure-velocity formulation. Physics of Fluids, 37(7).

[12] Rao, X., Liu, Y., Fu, Q., He, X., Kwak, H., Zhao, H., & Hoteit, H. (2025, September). Boundary-Integral Type Neural Network (BINN) for Flow Problems in Anisotropic Reservoirs. In SPE Middle East Oil and Gas Show and Conference (p. D021S048R001). SPE.

[13] Rao, X. (2024). Performance study of variational quantum linear solver with an improved ansatz for reservoir flow equations. Physics of Fluids, 36(4).

[14] Rao, X., Luo, C., He, X., & Hyung, K. (2024, November). An Efficient Quantum Neural Network Model for Prediction of Carbon Dioxide CO2 Sequestration in Saline Aquifers. In Abu Dhabi International Petroleum Exhibition and Conference (p. D021S061R005). SPE.

[15] Rao, X. (2024, November). The first application of quantum computing algorithm in streamline-based simulation of water-flooding reservoirs. In Abu Dhabi International Petroleum Exhibition and Conference (p. D011S003R006). SPE.

[16] Rao, X., He, X., Du, K., Kwak, H., Yousef, A., & Hoteit, H. (2024). A novel projection-based embedded discrete fracture model (pEDFM) for anisotropic two-phase flow simulation using hybrid of two-point flux approximation and mimetic finite difference (TPFA-MFD) methods. Journal of Computational Physics, 499, 112736.

[17] Rao, X., Zhao, H., & Liu, Y. (2022). A meshless numerical modeling method for fractured reservoirs based on extended finite volume method. SPE Journal, 27(6), 3525–3564.

[18] Santoso, A., Goto, D., Takehira, T., Aslam, A., Kawahara, A., & Sadatomi, M. (2016). Non-Newtonian Two-Phase Flow Characteristics across Sudden Expansion in Horizontal Rectangular Minichannel. World Journal of Mechanics, 6(8), 257-272.

[19] Shijo, J. S., & Behera, N. (2023). Pressure drop prediction in fluidized dense phase pneumatic conveying using machine learning algorithms. Journal of Applied Fluid Mechanics, 16(10), 1951-1961.

[20] Zheng, H., Lin, B., Jiang, J., Jin, Y., & Peng, L. (2024, February). Knowledge-Guided Machine Learning Method for Downhole Gauge Record Prediction in Deep Water Gas Field. In Offshore Technology Conference Asia (p. D031S020R004). OTC.