processes

MDPI

Article

Prediction of Leakage Pressure during a Drilling Process Based on SSA-LSTM

Dong Chen [1], Baolun He [2,*], Yanshu Wang [1], Chao Han [3], Yucong Wang [2] and Yuqiang Xu [2,*]

[1] Sinopec Matrix Corporation, Qingdao 266071, China
[2] National Key Laboratory of Deep Oil and Gas, China University of Petroleum (East China), Qingdao 266580, China
[3] Geosteering & Logging Research Institute, Sinopec Matrix Corporation, Qingdao 266071, China; hansuper713@hotmail.com
* Correspondence: s21020020@s.upc.edu.cn (B.H.); xuyuqiang@upc.edu.cn (Y.X.)

Abstract: Drilling-fluid loss has always been one of the challenging issues in the field of drilling engineering. This article addresses the limitations of a single fluid-loss pressure mechanism model and the challenges in predicting positive drilling-fluid-loss pressure. By categorizing fluid losses of various types encountered during drilling, different geological formations associated with distinct mechanisms are considered. The actual drilling-fluid density in the wellbore at the time of fluid-loss occurrence is taken as a reference value for calculating the positive drilling-fluid-loss pressure of the already drilled well. Building upon this foundation, a combined model utilizing the Sparrow Search Algorithm (SSA) and Long Short-Term Memory (LSTM) neural network is constructed. This model effectively explores the intricate nonlinear relationship between well logging, logging engineering data, and fluid-loss pressure. By utilizing both data from the already drilled wells and upper formation data from ongoing drilling, precise prediction of positive drilling formation fluid-loss pressure can be achieved. Case studies demonstrate that the approach established in this paper, incorporating upper formation data, reduces the average absolute percentage error of fluid-loss pressure prediction to 2.4% and decreases the root mean square error to 0.0405. Through the synergy of mechanistic models and data-driven techniques, not only has the accuracy of predicting positive drilling formation fluid-loss pressure has been enhanced, but also valuable insights have been provided for preventing and mitigating fluid losses during drilling operations.

Keywords: mechanism model; leakage pressure; SSA-LSTM; during the drilling process

Citation: Chen, D.; He, B.; Wang, Y.; Han, C.; Wang, Y.; Xu, Y. Prediction of Leakage Pressure during a Drilling Process Based on SSA-LSTM. *Processes* **2023**, *11*, 2608. https://doi.org/10.3390/pr11092608

Academic Editor: Qingbang Meng

Received: 28 July 2023
Revised: 23 August 2023
Accepted: 29 August 2023
Published: 1 September 2023

1. Introduction

Deep and ultradeep oil and gas resources represent the primary frontier for future energy supply. However, due to the intricate and ever-changing geological conditions in these depths, coupled with inadequate predrilling knowledge of subsurface formations, drilling operations frequently encounter challenges such as leaks, surges, collapses, and sticking, among other downhole complexities. Among these issues, wellbore leakage has emerged as one of the most prevalent complexities in recent years. This not only escalates drilling costs and diminishes drilling efficiency but also poses a significant risk of wellbore collapse, surges, and even blowouts, leading to major safety incidents [1–3]. At the drilling site, the occurrence of fluid loss is typically determined by monitoring changes in drilling-fluid volume and the flow rate at the wellhead. Wang [4] discussed the leakage judgment under different working conditions and found that the leakage law of drilling fluid also varies under different working conditions. Sun et al. [1] pointed out that the leakage channel, leakage pressure, and leakage rate are the three main characteristics of oil-well leakage. Accurately determining the nature of the leakage will strongly support the scientific screening of leak prevention and blockage technologies. However, due to the complexity of underground rock formations and the complexity and variability of pressure

systems, accurately predicting the leakage pressure of complex and volatile formations remains a challenge, and there is currently no complete set of leakage pressure prediction methods [5,6].

In previous studies, the rupture pressure as a safe drilling-pre-fluid density window can no longer meet engineering requirements, and it is necessary to use accurate leakage pressure as its upper limit. In the traditional leakage pressure mechanism model, Zhang et al. [7] established a mechanical model for the intersection of cracks and the wellbore based on the stress state of the wellbore-surrounding rock. However, the solution of this model is complex, and its applicability is poor. Subsequently, Fang [8] proposed a method for calculating formation leakage pressure based on logging data, and Zhai et al. [9] conducted research on the prediction and control model of shale-induced fracture leakage pressure. They proposed a dynamic model for leakage pressure based on leakage time, rate, and other factors, and demonstrated the accuracy of the model. However, the occurrence of drilling-fluid leakage is not only related to the formation leakage pressure and wellbore pressure but also to construction parameters and human factors. Therefore, traditional methods based on wellbore pressure balance have certain limitations.

With the continuous indepth integration of machine learning, big data, and other technologies in the field of petroleum engineering, the data-driven method for identifying and early warning of lost circulation risk shows obvious advantages over traditional model methods. Mohammad Sabah et al. [10] conducted research on intelligent prediction models such as decision trees (DT), adaptive neural fuzzy inference systems (ANFIS), artificial neural networks (ANN), and genetic algorithms multilayer perception hybrid artificial neural networks (GA-MLP), and confirmed that machine learning has certain advantages in predicting leakage. Pang et al. [11] selected 16 comprehensive logging parameters with the strongest correlation with drilling loss rate for model training and established a complex relationship between logging parameters and mud loss rate through a subgaussian mixed density network, confirming that the model can evaluate drilling loss risk in real-time. Matinkia et al. [12] validated multiple models using logging data, and the results showed that convolutional neural networks (CNN) models have significant advantages in feature extraction, especially for data with high volatility such as logging data. Song et al. [13] conducted research on the LSTM and back propagation (BP) combined model, and took the formation pore pressure of two wells (the whole well section is considered as the true value) as the training set and one as the verification, proving the feasibility and accuracy of the model. In summary, there are two major difficulties in constructing a leakage pressure profile. First, the model has numerous parameters that are difficult to determine, resulting in poor generalization ability; The second is that the model is simple but lacks accuracy. However, it is obvious that a single model can no longer meet the needs of safe drilling in the project. This paper builds a leakage pressure profile for the whole well section by evaluating the actual leakage of drilled wells and using different mechanism models for different types of leakage. By constructing a drilled leakage pressure profile using mechanism methods, it serves as a machine-learning data sample for predicting leakage pressure while drilling in the drilling formation. The selected LSTM model can perform feature learning well. Considering the impact of model hyperparameters on prediction results, this paper uses the SSA algorithm to optimize hyperparameters, striving to achieve higher accuracy of the model during the drilling process.

2. Methodology

2.1. Applicability Analysis

This study is based on ultradeep wells on land and aims to calculate the leakage pressure of various fractures, karst caves, and high permeability formations using various mechanism models. The actual data of leakage is included in the mechanism model validation, providing more accurate leakage pressure data for the drilled formations. The SSA algorithm optimizes the parameter search by simulating sparrows' foraging behavior and is suitable for various optimization problems. The LSTM model performs well in time-

series data analysis; therefore, it also has potential value in other tasks that require time correlation modeling, such as weather forecasting, financial data analysis, and other fields. In addition, this method provides insights for research in other geological engineering fields. This cross-domain migration potential will help expand the applicability of our method. The SSA-LSTM model used provides a good solution for predicting the leakage pressure in forward drilling.

2.2. Background Introduction to Block L

The high and steep structures in this block belong to the mountain front and mountain area [14], with adverse surface terrain conditions and underground structures coexisting. The strong folding and orogeny not only make the terrain height difference change greatly but also make the high and steep structure appear from the surface, making the stratum dip angle very large. Due to long-term erosion of surface water and weathering, fractures and karst caves are developed, and lost circulation is very serious. According to the statistical data of the leakage layer rock core in high and steep structures, the opening size, distribution shape, and filling material of the cracks vary greatly, with significant differences [15], but generally speaking, there are the following rules: from Shaximiao to Xujiahe, the leakage channels are mainly porous leakage channels. When drilling into poorly cemented sandstone and mudstone formations, permeability leakage often occurs; on the main part of the structure, due to long-term surface-water erosion and weathering, the leakage channels are mainly karst caves and large fractures, the width of the fractures is generally more than 10 mm, and most of them are nonfilling inclined fractures with an inclination of more than 30° and extend far to the depth of the stratum. Some of them are connected to the surface. When encountering such lost circulation channels, very serious well losses occur, and handling them is also quite difficult [16]. In addition, the geological environment of the block also has old strata exposed to the surface, weathered mountain gravel, and mostly broken surface of the old strata. Due to differences in the sedimentary age, sedimentation time, and fragmentation of the surface rocks, there are significant differences in the horizontal distribution of strata, resulting in the frequent occurrence of leakage in mid-to-shallow formations. When encountering faults during drilling, there is a phenomenon of venting, which can be confirmed by logging curves to indicate the development of fractures in this section, which can easily lead to malignant leakage.

There are a total of 12 wells in the selected block. Based on the analysis of the actual drilling situation, almost every layer of the wells in Block L, from the Penglai town formation to the Changxing Formation, has experienced well leakage, with a depth ranging from the surface to below 5000 m. However, the severity of well leakage varies greatly among different formations. According to the reasons for the leakage, it can be divided into fracturing leakage, expansion leakage, and differential pressure leakage. After detailed statistics, a total of 259 leaks occurred (multiple leaks at the same depth in continuous operations are recorded as one leak), and, according to different operating conditions, 66% of them were lost during the drilling process, as shown in Figure 1. The main layer is the Qianfo Cliff Formation, with a depth concentration of 1800–3300 m. The leakage type is mainly differential pressure leakage, up to 69%, as shown in Figure 2.

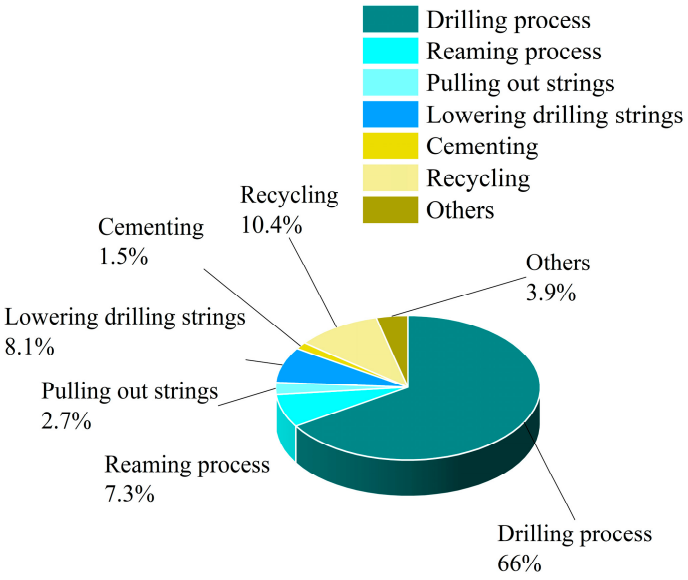

Figure 1. Analysis of Leakage Conditions in Block L.

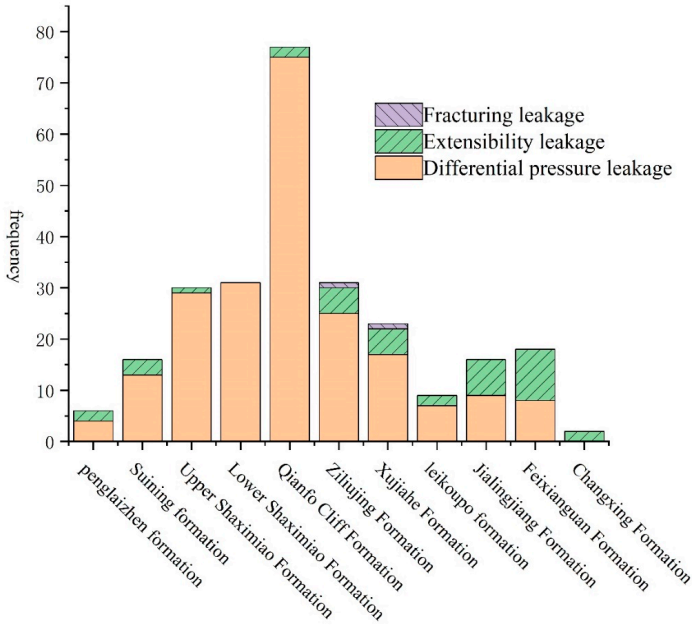

Figure 2. Layered Group Statistics for Different Types of Leakage.

2.3. Drilled Leakage Pressure Model

In response to the above statistical results, it is necessary to conduct indepth research and construct a set of leakage pressure prediction methods during the drilling process. This article takes the cause of leakage as the classification standard and uses different mechanism models for different leakage formations. Since regardless of the type of leakage, the leakage pressure can be regarded as the drilling fluid column pressure in the wellbore when the formation experiences leakage, which means that the drilled leakage pressure profile calculated based on traditional mechanism models needs to be corrected. Research has shown that the values of some model coefficients are also a challenge, and the examples provided in this article have already provided some model parameters in Section 3.1. The specific process is shown in Figure 3.

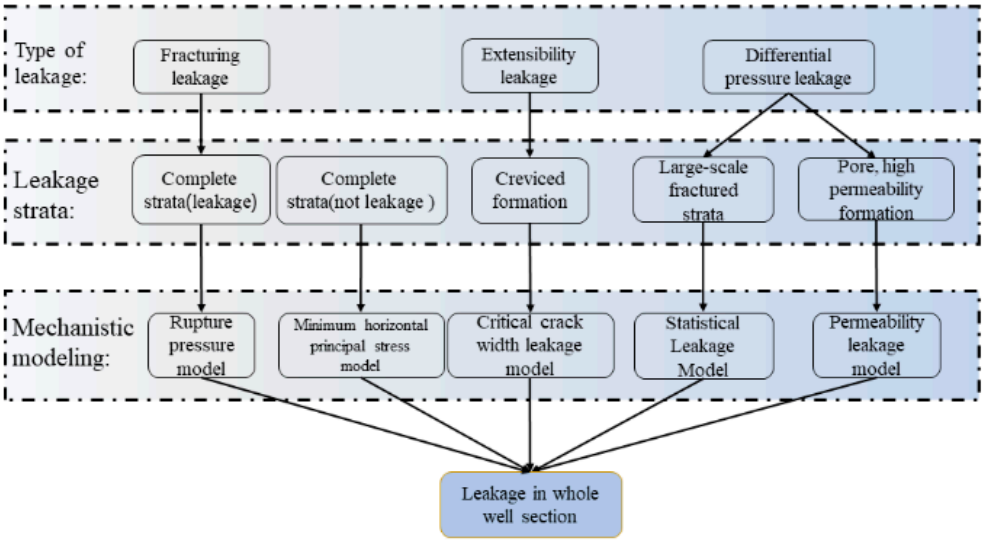

Figure 3. Construction process of the leakage pressure mechanism model for the entire well section.

2.3.1. Critical Crack Width Leakage Model

For extended leakage [17], a critical crack width leakage model is adopted, which assumes the existence of a critical crack width. When the crack width is less than the critical crack width, the drilling fluid will form a sealing layer with a certain pressure-bearing capacity inside the crack, and the drilling fluid will filter normally; when the fracture width exceeds the critical fracture width, it will be converted into fracture leakage. The deformation of cracks follows a power function form, and the flow of drilling fluid in cracks follows a cubic law. The relationship between crack width and effective stress is as follows [17]:

$$\omega = \omega_0 \left\{ A \left[\left(\frac{\sigma}{\sigma_0} \right)^a + 1 \right] \right\}^{-1} \tag{1}$$

where, ω is the dynamic width of the crack, mm; ω_0 is the crack width when the wellbore pressure is equal to the formation pressure, mm; σ is the effective stress on the vertical crack surface, MPa; σ_0 is the effective stress on the vertical fracture surface when the wellbore pressure is equal to the formation pressure, MPa; A and a is an undetermined coefficient; therefore, there is no reason.

Taking a single vertical joint as an example, ignoring the stress concentration around the wellbore, the relationship between the wellbore fluid column pressure and the effective normal stress on the fracture surface is obtained as follows [15]:

$$\sigma = \sigma_h - p'_f \tag{2}$$

where, p'_f is the effective liquid column pressure in the wellbore, MPa; σ_h is the minimum horizontal principal stress, MPa.

Combine Equations (1) and (2) above to obtain the relationship between the dynamic width of fractures and the pressure of the drilling-fluid column:

$$\omega = \omega_0 \left\{ A \left[\left(\frac{\sigma_h - p'_f}{\sigma_h - P_p} \right)^a + 1 \right] \right\}^{-1} \tag{3}$$

From the equation, it can be seen that there is a positive correlation between the dynamic width of fractures and the wellbore fluid column pressure. When the crack

width reaches the critical crack width, fractured leakage occurs underground. The leakage pressure calculation model [17] based on the critical crack width ω_c is:

$$P_{L1} = \sigma_h - \left(\frac{\omega_0}{A\omega_c} - 1\right)^{\frac{1}{a}}(\sigma_h - P_p) \quad (4)$$

where, P_{L1} is the crack propagation pressure, MPa; ω_c is the critical crack width, mm; P_p is the formation pore pressure, MPa.

2.3.2. Permeability Leakage Model

For permeability leakage [18], drilling-fluid leakage can be reflected by the Poromechanics formula. Assuming that the borehole is regular, it can be regarded as cylindrical. Since the drilling fluid is generally non-Newtonian fluid, when the Bingham model is used to describe the Rheology Constitutive equation of the drilling fluid, the relationship between the pressure and flow rate of drilling-fluid seepage can be expressed as the following formula [18]:

$$P_{L2} = P_p + \frac{Q_L \times 10^3 \eta_p}{2\pi kh} \ln\frac{r_e}{r_h} + \frac{7}{6000}\tau_0\sqrt{\frac{\phi}{5k}}(r_e - r_h) \quad (5)$$

where, Q_L is the average flow rate of drilling-fluid leakage, L/s; η_p is the plastic viscosity of Bingham plastic, Pa·s; τ_0 is the yield stress of Bingham plastic, Pa·s; ϕ is the porosity of the formation; therefore, there is no reason; P_{L2} is the permeability leakage pressure, MPa; k is the formation permeability, D; r_h is the wellbore radius, m; r_e is the leakage radius of the formation, m; h is the thickness of the leakage layer, m.

2.3.3. Fracture Pressure Model

Usually, due to excessive pressure in the drilling-fluid column or rapid increase in drilling-fluid density in the well, a large amount of pressure is formed, exceeding the maximum pressure-bearing capacity of the weak layer underground, leading to rock fracture and the formation of cracks, or the expansion of closed cracks in the rock, resulting in leakage [19]. The intact formation did not experience fracturing during the drilling process, therefore, its value is approximately equal to the fracturing pressure value, as shown in Equation (7). However, based on the analysis of well history data in this block, when the practical drilling-fluid density is much less than, fracturing leakage still occurs, which is mostly related to insufficient pressure-bearing capacity when drilling to thin and weak layers or lithological interfaces. Therefore, this article modifies Equation (6) by assigning a correction coefficient.

$$P_{L3} = P_f = 3\sigma_h - \sigma_H - \alpha P_p + S_t \quad (6)$$

$$P_{L3} = K_i P_f = K_i(3\sigma_h - \sigma_H - \alpha P_p + S_t) \quad (7)$$

where, P_f is the formation fracture pressure, MPa; S_t is the tensile strength of the rock, MPa; P_{L3} is the fracturing leakage pressure, MPa; α is the effective stress coefficient, dimensionless; K_i Correction coefficient for different depths, with a value range of 0.55~0.92.

2.3.4. Statistical Leakage Model

Differential pressure leakage refers to the presence of large-scale fractures, karst caves, and fracture karst-cave networks connected to the wellbore. Multiple pressure systems coexist in the L block where fractures intersect. In order to obtain the differential pressure of drilling-fluid leakage, the first step is to refer to well history data and classify the leakage based on the cause. Then, the leakage pressure difference and leakage rate are calculated and fitted for this type of differential pressure leakage; the fitting model is as follows [8]:

$$P_{L4} = P_p + \Delta p = P_p + KQ^n \quad (8)$$

where, P_{L4} is the differential leakage pressure, MPa; Δp is the leakage pressure difference, MPa; Q is the leakage rate, m^3/h; n is the fitting coefficient, dimensionless.

2.3.5. Minimum Horizontal Principal Stress Model

In traditional understanding, a complete formation without leakage can only occur when hydraulic fracturing occurs during drilling. However, for the development of fractures in this block, the stability of the rock layer is poor, and the leakage pressure is much lower than the fracturing pressure. Therefore, this article adopts the minimum horizontal principal stress model for the formation without leakage accidents. The minimum horizontal principal stress model believes that, for the formation with bedding, joints, and closed fractures, the fluid pressure that causes the crack to open only needs to overcome the ground stress on the vertical crack surface; that is, the leakage pressure is approximately equal to the minimum horizontal principal stress [20].

$$P_{L5} = \sigma_h = (\frac{\mu}{1-\mu} + \omega)(\sigma_z - \alpha P_p) + \alpha P_p \tag{9}$$

where, P_{L5} the leakage pressure determined for the small horizontal principal stress model, MPa; ω is the stress coefficient of horizontal construction, dimensionless; μ is Poisson's ratio, dimensionless; σ_z is the pressure of the overlying rock layer, MPa.

2.4. Data-Driven Approach

The process of oil drilling is a continuous process of generating data and deepening our understanding of reservoirs. The acquisition of underground data is very valuable, and how to fully explore the connections between data has become an urgent problem to be solved. Incorporating logging and engineering data into leak-pressure prediction provides a new approach. This article uses data from drilled wells as training samples, and it is a feasible method to calculate the leakage pressure of drilling wells. However, in order to achieve ideal prediction accuracy, adding data from the upper strata of the drilling well (previous drilling) to the previous samples will effectively improve model accuracy. At the same time, using an SSA algorithm to optimize LSTM hyperparameters can eliminate the blindness of manually setting parameters and improve timeliness.

2.4.1. SSA-LSTM Model

SSA [21] is an algorithm proposed in 2020, inspired by sparrows' predatory and antipredatory behavior. This algorithm has advantages such as avoiding falling into local optima, fast convergence speed, high convergence accuracy, and strong search ability. The algorithm and its variants have good performance in continuous optimization problems. The modeling process of this algorithm can be summarized as follows: assuming that the position of each sparrow is $x = \{x_1, x_2, \ldots, x_D\}$, its fitness is $f = f[x_1, x_2, \ldots, x_D]$, the population of sparrows is set to m, and n sparrows with the best population position are selected as producers in each generation, while the remaining m-n are selected as scavengers. Compared to other sparrows, individuals with higher fitness will prioritize the discovery of food. In addition, producers will always find abundant food and provide directions for all scavengers to search for food. Therefore, producers will obtain a larger search range. Therefore, the producer's location update is as follows (10). If R_2 is less than ST, then there are no predators and producers begin to conduct more extensive searches; If R_2 is greater than or equal to ST, it indicates that a certain number of sparrows have discovered predators, and an alarm is issued. All sparrows in the population must fly to safer areas to forage.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} \cdot \exp(\frac{-i}{\alpha \cdot iter_{\max}}), R_2 < ST; \\ X_{i,j}^{t} + Q \cdot L, R_2 \geq ST; \end{cases} \tag{10}$$

where, $X_{i,j}$ is the position information of the i th sparrow in the j th dimension; t is the current number of iterations; α is a random number between $(0, 1]$; $iter_{\max}$ is a constant; R_2 is the alarm value, which belongs to the range [0, 1]; ST is the safety threshold, which belongs to the range [0.5, 1]; Q is the value of a simple random distribution; L is a $1 \times d$ matrix with each element of 1.

The position of the scavengers is shown in Equation (11) below. When $i > n/2$, it means that the i th scavenger with lower adaptation is not getting food, so at this time the scavengers need to fly to other places to find food. When the sparrows realize the danger, they will abandon the immediate food and enter the warning state; the specific expression is shown in Equation (12) below.

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp(\frac{X_{worst}^t - X_{i,j}^t}{i^2}), i > n/2; \\ X_P^{t+1} + \left| X_{i,j}^t - X_P^{t+1} \right| \cdot A^+, otherwise; \end{cases} \tag{11}$$

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \cdot \left| X_{i,j}^t - X_{best}^t \right|, f_i > f_g \\ X_{i,j}^t + K \cdot (\frac{\left| X_{i,j}^t - X_{worst}^t \right|}{(f_i - f_w) + \varepsilon}), f_i = f_g; \end{cases} \tag{12}$$

where, X_{worst} is the current global worst position; X_P^{t+1} is the optimal position of the producer at the t + 1th iteration; A is a matrix of $1 \times d$ with elements in the range $[-1, 1]$ and which has to satisfy $A^+ = A^T \left(AA^T\right)^{-1}$; K is a randomized measure and belongs to the range $[-1, 1]$; β is a step control parameter with a random normal distribution; ε is a minimal constant which avoids the error of having a zero denominator in the division; X_{best} is the current globally optimal position; f_i is the fitness value of an individual sparrow at this point in time; f_w and f_g are the current global worst and best fitness, respectively.

If $f_i > f_g$, it means that at this time the sparrow is at the edge of the population, and there is a great possibility of being attacked by predators. If $f_i = f_g$, it means that the sparrows located in the center of the population have found the danger; in order to avoid being preyed upon, they need to approach the other sparrows.

LSTM [21] is a kind of recurrent neural network with complex and powerful asymptotic processing ability. Due to the existence of before and after correlation of logging data, the model can extract the logging sequence feature data along the depth, respectively, forward and backward, and the change rule of formation pressure before and after depth extraction. x_t and x_{t-1} represent the input states at the current time and before time respectively; H_t and H_{t-1} denote the current time and the previous time; C_t and C_{t-1} represent the cell states that send the current time and the previous time, respectively. σ denotes the sigmoid activation function with the range of [0, 1] and Tanh is the hyperbolic tangent function with the range of $[-1, 1]$.

First, the LSTM decides what information to discard from the cell state through the forgetting gate. The formula is as in Equation (13):

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{13}$$

f_t is the output of the oblivious gate; W_f is the weight matrix of the oblivious gate; b_f is the bias term of the oblivious gate. Next, the LSTM determines the information that needs to be updated through the input gate according to Equation (13):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{14}$$

i_t is the output of the input gate, W_i is the weight matrix of the input gate, and b_i is the bias term of the input gate. Then we update the cytosolic state C_t. The candidate cytosolic state $\hat{C}_t$ and the current cell state C_t are expressed as Equations (15) and (16).

$$\hat{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{15}$$

$$C_t = f_t C_{t-1} + i_i \hat{C}_t \tag{16}$$

where, W_c is the weight matrix of the candidate vector, and its deviation is represented by b_c. Finally, the LSTM determines the output information state of the unit through the output gate. The output of the output gate is shown in Equation (17):

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{17}$$

O_t is the output of the output gate; W_o is the weight matrix output gate; b_o is the bias term of the output gate. The hidden state output of the LSTM unit is shown in Equation (18):

$$h_t = O_t tanh(C_t) \tag{18}$$

However, the hyperparameters of this model have a significant impact on the model, and during the drilling process, different parameters should be used at different intervals or lithological layering. Considering the effectiveness of the drilling process and the addition of new knowledge in the drilled section, the practical sparrow algorithm is a good way to optimize the hyperparameters of the model. The optimization process is shown in Figure 4:

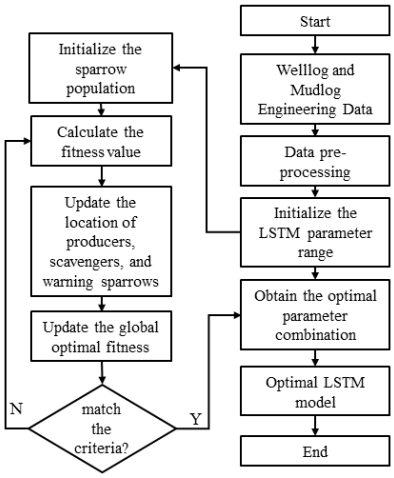

Figure 4. SSA-LSTM algorithm flowchart.

2.4.2. Data Preparation and Evaluation Indicators

The burrs or abrupt changes in logging and logging curves may be abnormal jumps, or they may indeed be sudden changes in the rock environment. In order to eliminate the influence of abnormal data on correlation and model, data smoothing is also required before this, and a fast Fourier transform smoothing method is used. Normalization processing takes into account the dimensional impact between data at different scales and normalizes the logging data to a range of 0~1. After normalization, the logging data y_i is shown in Equation (19).

$$y_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}, 1 < i < n \tag{19}$$

where, y_i is the normalized logging data at the corresponding depth H_i, and x_i represents the original logging data at the measured depth H_i.

The maximum information coefficient (MIC) is used to measure the degree of correlation between two variables x and y, as well as the strength of linearity or nonlinearity. Compared to conventional Pearson correlation analysis, MIC is more suitable for complex nonlinear relationships and has the advantages of low computational complexity and higher robustness. The range of values for the MIC correlation coefficient is [0, 1], and the closer it is to 1, the stronger the correlation.

The mean absolute percentage error ($MAPE$) [13] and root mean square error ($RMSE$) of the indicator are shown in Equations (20) and (21), respectively.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{20}$$

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{21}$$

3. Example Analysis and Comparison

3.1. Examples of Lost Pressure in Drilled Wells

A total of twelve wells have been drilled in this block. This article uses two wells as validation to simulate the drilling process, while the rest are adjacent wells. Through the calculation of formation pore pressure and leakage pressure in the previous section, three adjacent well-pressure profiles were constructed and compared based on the actual leakage situation in the well history. The following is an example of the L4 well calculation, as shown in Figure 5. The well has experienced a total of 27 losses, with the types of losses being expansionary and differential pressure losses:

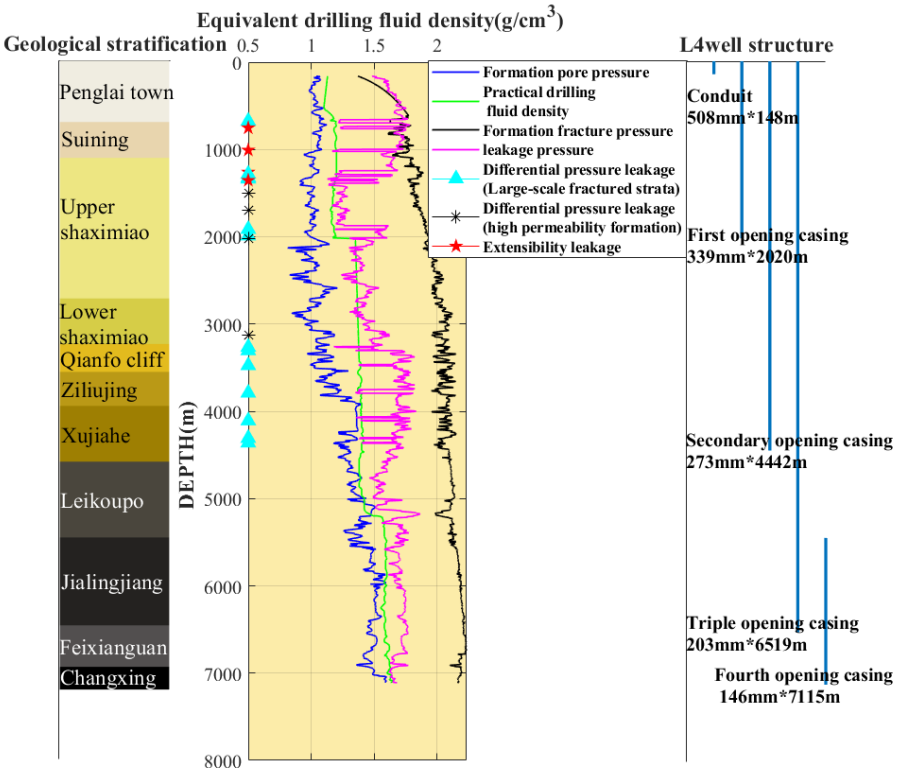

Figure 5. Pressure profile of L4 well.

(1) At 674 m–1397 m (from Penglaizhen Formation to Suining Formation), microfractures coexist with large-scale fractures. For example, at 1008 m, extended leakage occurs. When the wellbore pressure is equal to the formation pressure, the crack width is 0.5 mm, the critical crack width is 2.5 mm, the formation pore pressure is 10.56 MPa, the minimum horizontal principal stress is 16.85 MPa, A is 0.15, a is 6, and the leakage pressure at this depth is calculated to be 11.61 MPa;

(2) Intermittent permeability leakage occurs in a large section from 1398 m to 1890 m (Shang Shaximiao Formation). The leakage well section is long, the leakage volume is large, and the leakage rate is low, but the success rate of plugging is high. At 1398 m, the average leakage flow is 0.56 L/s, the plastic viscosity is 0.035 Pa·s, the permeability is

0.25 D, the hole radius is 0.22 m, the average leakage half diameter is 5 m, the porosity is 3%, the drilling-fluid yield stress is 7 Pa·s, the formation pore pressure is 12.56 MPa, and the leakage pressure is 16.55 MPa;

(3) There are many large-scale fractures in the 3475–4361 m (artesian well group) formation, resulting in lateral differential pressure leakage. The leakage rate is relatively high, and the leakage amount is relatively large. Based on the actual differential pressure leakage situation of adjacent wells (C5, C6, C7), the relationship between pressure difference and leakage rate is obtained as follows: $\Delta P = 1.47Q^{0.562}$.

3.2. Data Sample Instance

In order to demonstrate the improvement effect of the previous drilling data on the drilling formation model, this article constructed adjacent well data sample 1 and sample 2 integrated into the drilled formation, respectively, to predict the leakage pressure while drilling. MIC analysis included a total of 19 feature parameters, excluding parameters with a correlation less than 0.4 and retaining 12 parameters with a strong correlation with leakage pressure such as DEPTH, MwIN, MwOUT, and Inlet Resilience for model training, as shown in Figure 6. Similarly, the L4 well was used as validation, and training sets were constructed using L5, L6, and L9. The constructed sample dataset 1 (8294 × 11) and the data from the first two wells of L4 were integrated into the drilled dataset (12,197 × 11).

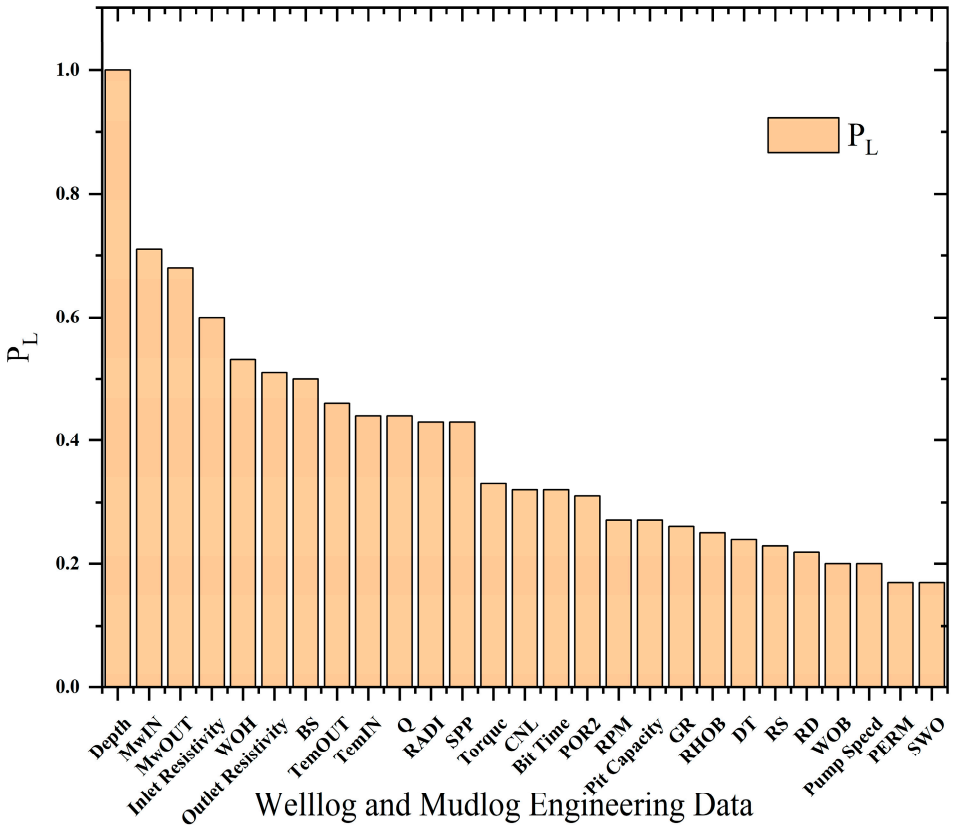

Figure 6. MIC correlation coefficient.

3.3. SAA-LSTM Model Parameter Settings

The first layer neuron range is [5–150], the second layer neuron range is [5–150], and the dropout ratio is [0.05–0.8]. The batch size is [8–32], the sparrow algorithm searches for the optimal combination are (10, 20, 0.6788, 8), the sparrow population size is set to 20, and the maximum number of iterations is 100. The proportion of producers in the population is 20%, and its safety threshold is 0.8.

3.4. Comparison of Positive Drilling Model Predictions

Through the data samples and model construction constructed in the previous section, the comparison results of Figure 7 were obtained, and the two predicted results were compared with the actual leakage pressure. Using only drilled data samples for training, testing was conducted on a positive drilling well (with a depth range of 4450–7200 m). Sample 1 showed an RMSE of 0.053 and a MAPE of 2.8%. However, for a positive drilling well, incorporating the upper drilling data can reduce the RMSE to 0.0405 and the MAPE to 2.4%. And this data sample can better reflect the leakage pressure situation of vulnerable formations (pressure profile fluctuations). Furthermore, it has been proven that accurate prediction of leakage pressure is difficult, and using adjacent well data as samples alone cannot effectively characterize the situation of leakage pressure in normal drilling, especially in areas with uneven lateral distribution of the formation. After adding the upper strata data, this not only proves the applicability of the SSA-LSTM model but also demonstrates the impact of data quality on its results.

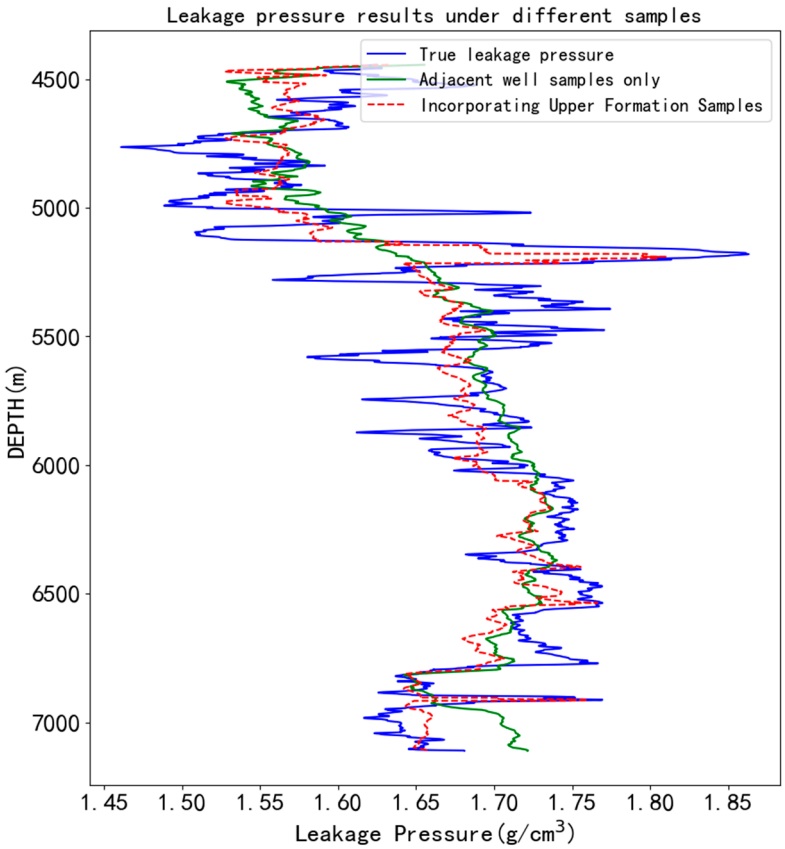

Figure 7. Prediction results of drilling leakage pressure under different samples.

4. Conclusions

This article proposes a method for predicting the leakage pressure during the drilling process, which utilizes various traditional mechanism models to calculate the leakage profile of the drilled well. The constructed leakage samples are used as the output of the machine-learning model to achieve the inversion of the leakage pressure in the drilling formation while drilling. In order to highlight the importance of upper formation information in drilling, we also compared two different datasets and obtained the following summary:

1. The leakage situation of the block was analyzed, and a leakage pressure mechanism model suitable for the entire well section of the block was obtained. Instead of using a single model to calculate the leakage pressure, different models are used for different leakage formations to build the leakage pressure of the whole well section. And the actual

leakage situation will be checked, and the accurate calculation of the leakage pressure of the drilled well will provide a basis for the training of the forward drilling model;

2. The constructed well leakage profile data can be used as the output of the model to effectively predict the leakage pressure of the drilling formation. Different data samples have been constructed to prove that incorporating the upper formation data of the drilling into the samples can effectively improve the accuracy of the model, with an RMSE of only 0.0405 and a MAPE of only 2.4%. And it can better reflect the leakage pressure of vulnerable formations (pressure-profile fluctuations);

3. Future work will start with software integration of input features and combined models, integrating more logging while drilling data into the model, and achieving leakage pressure prediction by combining geological and engineering data. We will always keep up with onsite requirements and integrate this method into the software to achieve risk prediction of drilling leakage and assist in adjusting the drilling construction plan.

Author Contributions: Conceptualization, D.C. and B.H.; methodology, Y.X.; software, C.H.; validation, Y.W. (Yanshu Wang), Y.W. (Yucong Wang) and Y.X.; formal analysis, Y.W. (Yucong Wang); investigation, D.C.; resources, D.C.; data curation, B.H.; writing—original draft preparation, B.H.; writing—review and editing, B.H.; visualization, C.H.; supervision, Y.W. (Yanshu Wang); project administration, D.C.; funding acquisition, D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sun, J.; Liu, F.; Cheng, R.; Feng, J.; Hao, H.; Wang, R.; Bai, Y.; Liu, Q. Research progress and prospects of machine learning in lost circulation control. *Acta Pet. Sin.* **2022**, *43*, 10.
2. Wang, C.; Liu, H.; Liu, Y.; Xiao, F.; Zheng, N. Managed pressure drilling technology: A research on the formation adaptability. *Fluid Dyn. Mater. Process.* **2022**, *18*, 1865–1875. [CrossRef]
3. Xie, R.; Zhang, X.; He, B.; Zheng, N.; Yu, Y. Analysis of the applicability of a risk quantitative evaluation method to high temperature-pressure drilling engineering. *Fluid Dyn. Mater. Process.* **2023**, *19*, 1385–1395. [CrossRef]
4. Wang, S. Judgment of drilling fluid leakage and overflow during the drilling process. *West-China Explor. Eng.* **2020**, *32*, 49–52.
5. Yang, J.; Sun, J.; Bai, Y.; Lv, K.; Zhang, G.; Li, Y. Status and prospect of drilling fluid loss and lost circulation control technology in fractured formation. *Gels* **2022**, *8*, 260. [CrossRef] [PubMed]
6. Zhang, X.; Xie, R.; Liu, K.; Li, Y.; Xu, Y. Analysis of the lost circulation problem. *Fluid Dyn. Mater. Process.* **2023**, *19*, 1721–1733. [CrossRef]
7. Zhang, L.; Wang, X.; Xie, T.; Lin, H.; Yuan, W. A New leakage pressure prediction model for natural fractured formation. *Drill. Prod. Technol.* **2018**, *41*, 5.
8. Fang, J.; Yang, F.; Jia, X.; Deng, M. Method of calculating formation leakage pressure based on mud logging data. *Mud Logging Eng.* **2020**, *31*, 5.
9. Zhai, X.; Chen, H.; Lou, Y.; Wu, H. Prediction and control model of shale induced fracture leakage pressure. *J. Pet. Sci. Eng.* **2020**, *198*, 108186. [CrossRef]
10. Sabah, M.; Talebkeikhah, M.; Agin, F.; Talebkeikhah, F.; Hasheminasab, E. Application of decision tree, artificial neural networks, and adaptive neuro-fuzzy inference system on predicting lost circulation: A case study from Marun oil field. *J. Pet. Sci. Eng.* **2019**, *177*, 236–249.
11. Pang, H.; Meng, H.; Wang, H.; Fan, Y.; Nie, Z.; Jin, Y. Lost circulation prediction based on machine learning. *J. Pet. Sci. Eng.* **2022**, *208*, 109364. [CrossRef]
12. Matinkia, M.; Amraeiniya, A.; Behboud, M.M.; Mehrad, M.; Bajolvand, M.; Gandomgoun, M.H.; Gandomgoun, M. A novel approach to pore pressure modeling based on conventional well logs using convolutional neural network. *J. Pet. Sci. Eng.* **2022**, *211*, 110156. [CrossRef]
13. Song, X.; Yao, X.; Li, G.; Xiao, L.; Zhu, Z. A novel method to calculate formation pressure based on the LSTM-BP neural network. *Pet. Sci. Bull.* **2022**, *7*, 12–23.

14. Pu, J. Research on Comprehensive Control of Lost Circulation Technology in Northeast Sichuan Region. Ph.D. Thesis, Southwest Petroleum University, Chengdu, China, 2015.
15. Lietard, O.; Unwin, T.; Guillot, D.; Hodder, M. Fracture width LWD and drilling mud/LCM selection guidelines in naturally fractured reservoirs. In Proceedings of the European Petroleum Conference, Milan, Italy, 22–24 October 1996.
16. Fuh, G.F.; Morita, N.; Boyd, P.A.; McGoffin, S.J. A new approach to preventing lost circulation while drilling. In Proceedings of the SPE Annual Technical Conference and Exhibition, Washington, DC, USA, 4–7 October 1992.
17. Li, D.; Kang, Y.; Liu, X.; Zeng, Y.; Du, C. The lost circulation pressure of carbonate formations on the basis of leakage mechanisms. *Acta Pet. Sin.* **2011**, *32*, 900–904.
18. Zhang, J.; Yue, H.; Zhang, D.; Hou, R.; Liu, K. New analysis and calculation method for permeable leakage layer depth and pressure. *Oil Drill. Prod. Technol.* **2013**, *35*, 12–15.
19. Le, M.; Wang, D.; Xu, H.; Wang, J. Analysis of Leakage Pressure in Different Formation. *Liaoning Chem. Ind.* **2016**, *45*, 195–197.
20. Zhang, L.; Xu, J.; Xie, T.; Lin, H.; Liu, H. Comparison of Several Calculation Models for Loss Pressure of Fractured Formation. *China Pet. Mach.* **2018**, *46*, 13–17.
21. Chen, L. *Ultra-Short-Term Wind Speed Forecast Based on the LSTM Optimized by the Sparrow Search Algorithm*; Jiangxi University of Finance and Economics: Nanchang, China, 2022.