Calibration Technology and Application of Mud Logging Sensors Based on Artificial Intelligence

Chang-liang Wu[1,2,3]($\boxtimes$), Zhi-xiong Zhou[1,2,3], Tie-heng Ding[1,2,3], Jian-guo Xiong[1,2,3], Yong-liang Gao[1,2,3], Yang Li[1,2,3], and Xue-li Luo[1,2,3]

[1] KEMBL Petroleum Technology Co., Ltd., Beijing, China
`wuchangliangdr@cnpc.com.cn`
[2] CNPC Engineering Technology R&D Company Ltd., Beijing, China
[3] National Engineering Research Center of Oil and Gas Drilling and Completion Technology, Beijing, China

Abstract. Mud logging serves as the "eyes" of exploration and development, acting as a counselor for drilling safety, the center of information transmission, and holding the first-hand data on oil and gas exploration and development. With the rapid development of informatization, digitization, intelligence, and remote support systems, the demand for high-quality mud logging data has continuously risen, where sensor calibration and calibration technology serve as the foundation for ensuring accuracy and reliability. This paper proposes an artificial intelligence-based comprehensive mud logging instrument sensor calibration and calibration technology, targeting the issues of prolonged service life, low precision, and low inspection rate of traditional mud logging instruments. The technology primarily involves collecting and pre-processing sensor output data such as filtering, sampling to eliminate noise, and improve the dataset's quality. Mathematical models of sensors were constructed using machine learning or deep learning algorithms to analyze the relationship between sensor outputs and actual values, which could also compute sensor errors and uncertainties. Algorithm optimization methods such as wavelet transform and adaptive filtering were used to process and analyze sensor data for different types of sensors and environmental conditions. The adaptive control algorithm was then utilized based on the predicted model results and actual measurement results to calibrate the sensor, ultimately helping to avoid

errors and uncertainty in the traditional manual calibration process. Experimental results show that this technology has higher accuracy and reliability than traditional calibration techniques while maintaining simple operation, fast speed, and cost-effectiveness. This technology improves the level of detection and evaluation technology of comprehensive mud logging instruments, Standardizes mud logging equipment management, and plays an essential role in timely discovering, evaluating oil and gas layers, and optimizing drilling construction safety.

1   Introduction

Informationization, digitization, intelligentization, and remote support systems require high-quality mud logging data. The comprehensive mud logging instrument is the main technical equipment on-site for mud logging, responsible for data collection, processing, analysis, and transmission. It monitors engineering parameters and drilling fluid parameters in real-time during the drilling process, analyzes various gas contents in the drilling fluid. The accuracy and reliability of mud logging data directly affect the quality and safety of drilling projects. It is the basis for timely discovering and evaluating oil and gas layers and optimizing drilling construction safety. Currently, there are several factors that affect the quality of mud logging data.

(1) The harsh installation and usage conditions of the mud logging sensors may cause reduced accuracy, malfunctions, and damages (as shown in Fig. 1).
(2) Mechanical vibration and impact caused by frequent lifting and long-distance transportation can damage equipment in the instrument room. Chipsets and electronic components will have degraded performance as their service time increases, which can lead to abnormal data channels or reduced conversion accuracy.
(3) The performance of the gas analysis system will decrease with production and operation time.
(4) Similar to drilling operations, mud logging operations are located in remote locations with difficult-to-control environmental conditions. Existing indoor testing and assessment devices have low integration and large size and weight, which cannot meet on-site testing needs, resulting in delayed and incomplete testing and assessment.
(5) Comprehensive mud logging instruments of different brands and periods have significant differences in quality, configuration, performance, etc. Many instruments have been in service for more than 10 years.
(6) Most mud logging companies mainly focus on individual testing, calibration, and verification, lacking a unified and systematic comprehensive mud logging instrument testing and evaluation device and technical specifications.

Therefore, major petroleum companies at home and abroad attach great importance to mud logging work, regarding improving mud logging equipment performance and ensuring mud logging data quality as the basis for improving mud logging quality and

engineering technical data quality. In order to eliminate the impact of the above unfavorable factors, research on comprehensive mud logging instrument testing and evaluation technology and equipment has been conducted, forming a set of complete technical specifications for comprehensive mud logging instrument testing and evaluation, and developing a comprehensive mud logging instrument testing and evaluation device that can adapt to fieldwork.

Fig. 1. Partial sensors of comprehensive mud logging instrument

Significant progress has been made in recent years with the application of artificial intelligence technology. Many researchers apply AI to sensor-based health and sports biomechanics [1–7], while others utilize it for intelligent industrial manufacturing [8–10]. In the area of using artificial intelligence for sensor calibration, many scholars have also conducted extensive research and achieved significant progress [11–16].

2 Performance Testing, Data Acquisition and Preprocessing of Sensor

2.1 Performance Testing of the Sensors

Sensor performance testing and data acquisition and preprocessing form the foundation of artificial intelligence-based calibration technology for comprehensive logging instrument sensors. A complete set of sensor performance testing equipment was developed using a Siemens 16-bit high-precision PLC, high-precision pressure source, high and

low-temperature constant temperature water tank, standard current generator, and sup-porting coils and high-precision resistors, and an accompanying system was programmed in C#.

Sensor performance testing: Standard testing equipment is used to test the perfor-mance of the logging sensors and record measurement results and error data. Sensor performance testing usually includes the following indicators: sensitivity, resolution, accuracy, and response time. The performance of the sensor is determined by testing it through methods such as adding a known quantity to the sensor or directly placing it in a changing environment and collecting feedback signals. After completing the sensor performance testing, the sensor is evaluated against specific application requirements. In short, sensor performance testing is an important step in ensuring data accuracy and reliability.

2.2   Data Collection and Preprocessing

The data collection system is used to collect the data obtained from the above tests and preprocess it for feature extraction by machine learning algorithms. Different signal acquisition methods, either analog or digital, are employed depending on the type of sensor. Then, the collected data must be preprocessed to remove noise, artifacts, and other unwanted signals.

Filtering is a common data preprocessing technique that can separate useful signals from noisy signals by applying filters. Depending on the filtering method, it can be classified into various types such as low-pass filtering, high-pass filtering, and band pass filtering. Sampling refers to the process of discretizing raw data by converting continuous analog signals into discrete digital signals, making them easier to store and process.

Filtering is a process that removes or retains certain components of a signal via filters. Its mathematical principles are based on signal processing theory. Common filtering methods include moving average filtering, median filtering, IIR low-pass filtering, FIR low-pass filtering, and frequency domain filtering. For logging sensor signal filtering, FIR low-pass filtering is used.

The mathematical formula for FIR low-pass filter can be expressed as:

$$y(n) = \sum_{k=0}^{M} h(k)x(n-k) \tag{1}$$

where $x(n)$ represents the original signal, $y(n)$ represents the filtered signal, and $h(k)$ represents the coefficients of the filter. The order of the filter is denoted by M.

A linear phase FIR filter is adopted, and the specific formula for calculating its coefficients is as follows:

$$h(k) = \begin{cases} 2f_c, k = \frac{M}{2} \\ \frac{\sin(2\pi f_c(k-\frac{M}{2}))}{\pi(k-\frac{M}{2})}, k \neq \frac{M}{2} \end{cases} \tag{2}$$

In the above formula, $h(k)$ represents the coefficients of the filter, M is the order of the filter, and f_c is the cutoff frequency of the filter. After calculating the filter coefficients using the aforementioned formula, they can be applied to convolution operations to filter the original signal and obtain the filtered results.

After applying the standard excitation signal generated by the high-precision standard source to the logging sensor, an analog current signal will be generated by the sensor, which is generally located between 4–20 mA. With the developed signal acquisition instrument, the current signal can be read and converted into corresponding physical quantities to complete the calibration of the sensor. Figure 2 shows the sensor signal acquisition instrument, and Fig. 3 shows the working of the electric torque sensor calibration device.

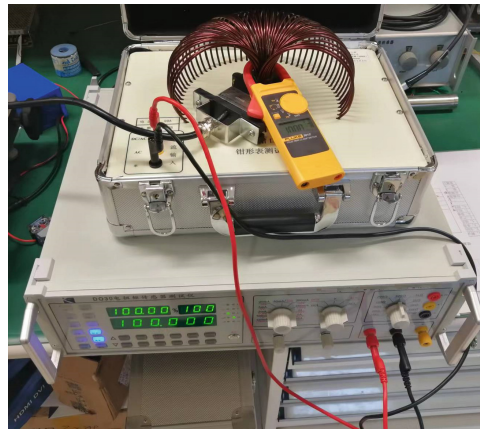**Fig. 2.** SDA-01 Sensor Data Acquisition Instrument

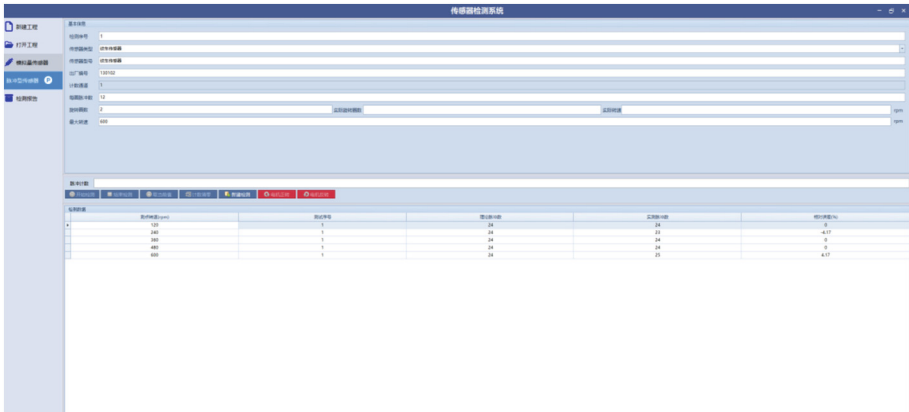**Fig. 3.** Calibration of electric torque sensors

Fig. 4. Interface of the calibration system for comprehensive logging instrument sensors

3   Model Establishment Based on Bayesian Optimization

After obtaining the raw data of mud logging sensors, the optimal model structure and hyperparameter combination are searched through Bayesian optimization to obtain a better theoretical curve.

Bayesian optimization is a black-box function optimization method commonly used in scenarios where a target function needs to be maximized or minimized. When constructing a Bayesian optimization model, we need to define a Gaussian process to describe the overall trend and uncertainty information of the target function. We also need to define a surrogate function to approximate the target function and optimize the surrogate function to find the optimal solution of the target function.

Specifically, the following steps are taken to construct the Bayesian optimization model:

(1) Define the prior distribution of the Gaussian process. In this step, we need to define a mean function and a covariance function for the Gaussian process. The mean function is used to describe the average value of the target function at different input values, while the covariance function is used to describe the correlation between different input values. The typically chosen Gaussian process prior distribution is the zero-mean Gaussian process.
(2) Update the posterior distribution of the Gaussian process based on the existing data. In this step, we need to update the mean function and covariance function of the Gaussian process based on the existing sample data to obtain a more accurate function approximation.
(3) Calculate the next sampling point based on the surrogate function. In this step, we need to use the current Gaussian process to fit the target function, construct a surrogate function, and select the next sampling point by optimizing the surrogate function. Common optimization methods include greedy algorithm and coordinate axis optimization.
(4) Update the posterior distribution of the Gaussian process based on the new sampling point. After obtaining the new sampling point, it can be added to the existing samples,

and these data can be used to update the mean function and covariance function of the Gaussian process.

(5) Repeat steps 3 and 4 until the preset stopping conditions are met.

The entire process of Bayesian optimization can be mathematically expressed as follows:

$$x_{i+1} = arg\max_x EI(x|D_t) = arg\max_x(\mu_t(x) - \xi\sigma_t(x)) \tag{3}$$

In the equation, x_{t+1} represents the next sampling point chosen in the iteration, $EI(x|D_t)$ is the expected improvement metric, representing the expected increase in target function value over the current best known value, given x as input under Gaussian process fitting. $\mu_t(x)$ and $\sigma_t(x)$ represent the mean and standard deviation of the current Gaussian process at x, respectively. ξ is a hyperparameter that controls the balance between exploration and exploitation and is commonly set to 2 or 3".

4   Model Training and Experimental Verification

4.1   Model Training

Neural networks are models composed of neurons that utilize components such as weights, biases, and activation functions to facilitate information transmission and processing. These models possess remarkable fitting and expressive abilities, making them suitable for solving various machine learning and deep learning tasks. Therefore, in this study, the neural networks were used to train the sensor calibration data. Simultaneously, the Cross Entropy loss function and the Stochastic Gradient Descent (SGD) optimizer was selected as key components of the neural network and combined to train a more accurate and efficient model.

Model Training: Using a large-scale dataset to train the model, constantly updating the model parameters to improve prediction accuracy and robust performance.

To train the model with a large-scale dataset, the following steps are required:

(1) Data collection and preparation: First, it is necessary to obtain enough data to train the model, and the data should be representative and able to cover various possible situations. Then, the data needs to be cleaned, transformed, and normalized, so that the model can better understand and process it.

(2) Model selection and design: Based on the application scenario and data characteristics, select an appropriate model structure and determine the parameters that need to be optimized.

(3) Loss function and optimizer selection: Depending on the task of the model, select an appropriate loss function and optimizer to evaluate the model and adjust the model parameters.

(4) Batch training: Since the dataset is too large to be loaded into memory for training at once, the data needs to be divided into equally-sized batches, and the stochastic gradient descent algorithm (SGD) is used to update the model parameters batch by batch.

(5) Batch normalization and regularization: Performing batch normalization before or after each batch can reduce the bias and variance of input features, thereby improving the model's prediction accuracy and robustness. In addition, methods such as L1 or L2 regularization can constrain the size and number of model parameters, avoiding overfitting and underfitting.

(6) Model evaluation and fine-tuning: Evaluate the model through the training and testing sets to determine the model's prediction accuracy and robust performance. If problems are found in the model, fine-tuning is needed, such as changing the model structure, adjusting the loss function or optimizer, etc.

Through these steps, the model can be trained using a large-scale dataset, constantly updating the model parameters to improve prediction accuracy and robust performance.

When choosing the appropriate network architecture, number of layers, and number of nodes to establish the ANN model and initialize weights, several steps usually need to be performed:

The problem type is determined by firstly clarifying whether a classification problem or a regression problem is faced. This will help determine the network structure and activation function.

Input and output are determined by specifying the number and type of input feature vectors and output predicted values.

An appropriate activation function is chosen based on the problem type, such as sigmoid or ReLU.

The network structure is designed by selecting a network structure that includes determining the range of the number of nodes in each hidden layer, whether to use dropout techniques, and so on.

Weights are initialized by selecting appropriate weight initial values, such as Xavier initialization, etc.

The model is trained by using the training dataset, and parameters are adjusted based on the validation set results.

The model is evaluated by examining its performance using the testing set.

When selecting the network structure and number of nodes, the principle of Occam's Razor should be followed. That is to say, the model structure should be made as simple as possible with reduced node numbers to prevent overfitting. At the same time, when designing the model, common deep learning frameworks such as TensorFlow and PyTorch can be considered. They provide a series of optimized structures, numbers of layers, and nodes, as well as pre-trained weights, which can reduce some manual parameter tuning work.

Taking the casing pressure sensor as an example, with a measuring range of $0 \sim 70MPa$, it should be calibrated using a standard pressure pump source calibrated by the Beijing Institute of Metrology and Measurement. The training samples are shown in the table below.

Table 1. Training sample for calibration of casing pressure sensor.

Input	Output	Result
0MPa	4.00 mA	Pass the calibration
0MPa	4.01 mA	Pass the calibration
0MPa	4.02 mA	Pass the calibration
0MPa	4.03 mA	Pass the calibration
…	…	…
0MPa	4.08 mA	Pass the calibration
0MPa	4.09 mA	Failure to pass the calibration
8.75 MPa	6.00 mA	Pass the calibration
8.75 MPa	6.01 mA	Pass the calibration
8.75 MPa	6.02 mA	Pass the calibration
8.75 MPa	6.03 mA	Pass the calibration
8.75 MPa	6.04 mA	Pass the calibration
8.75 MPa	6.04 mA	Pass the calibration
…	…	…
8.75 MPa	6.12 mA	Pass the calibration
8.75 MPa	6.13 mA	Failure to pass the calibration
…	…	…

4.2 Experimental Verification and Data Visualization

Test the model, compare the experimental data with the predicted data, evaluate the reliability and accuracy of the model, and adjust and improve it accordingly.

Based on the experimental test results, present the data in the form of charts and analyze the sources and trends of errors to provide visual support for sensor calibration and testing.

As can be seen from the figure below, the data predicted by AI technology is in very good agreement with experimental data, with a maximum error of only 0.15%, thereby proving the effectiveness of this method.

Using the above technical solution, it is possible to utilize artificial intelligence technology for the calibration of logging tool sensors in order to improve testing accuracy and efficiency, reduce human error and testing costs, and provide strong support for the drilling engineering in the oil and gas industry.
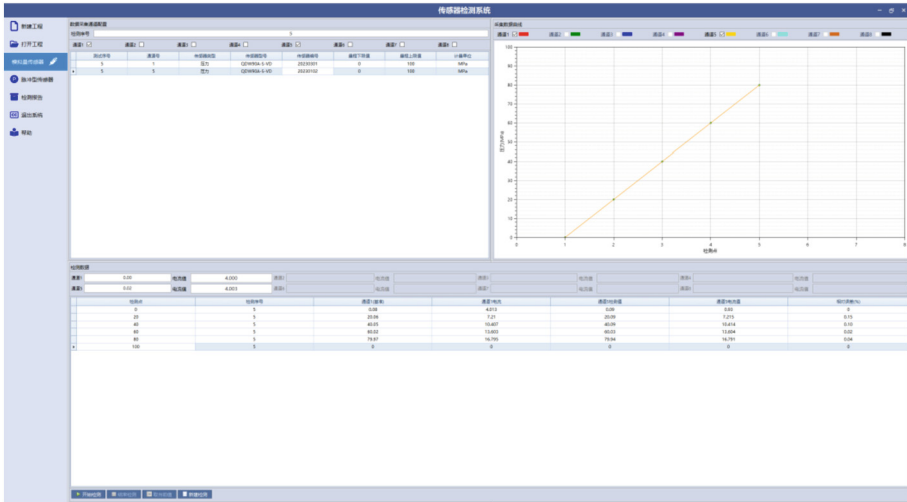
Fig. 5. Experimental Verification and Data Visualization

5　Conclusion

(1) The application of artificial intelligence in mud logging sensors can greatly improve the accuracy and reliability of the measurement results.

(2) The use of machine learning algorithms artificial neural networks (ANNs) can effectively address the problem of nonlinearity and complex interference in mud logging data.

(3) The calibration technology based on these algorithms has been successfully applied to real drilling engineering, achieving excellent results. The study improves the quality of mud logging data and provides a theoretical basis and practical guidance for the further promotion and development of intelligent mud logging technology.

References

1. Balarabe, J.S., Abubakar, I.A., Nuhu, S.A., et al.: Artificial intelligence, sensors and vital health signs: a review. Appli. Sci. **12**(22) (2022)
2. Zhang, C., Cheng, K.: Accurate detection of intelligent running posture based on artificial intelligence sensor. J. Sensors (2022)
3. Chen, Y., Chen, Q.: Gymnastics action recognition and training posture analysis based on artificial intelligence sensor. J. Sensors (2022)
4. Li, K.: Tennis technology recognition and training attitude analysis based on artificial intelligence sensor. J. Sensors (2022)

5. Song, Z., Tian, C.: Influence of the athlete's training physical state test based on the principle of artificial intelligence sensor. Mobile Inform. Syst. (2022)

6. Michael, P., Douglas, B., Wayne, D., et al.: Artificial intelligence, sensors, robots, and transportation systems drive an innovative future for poultry broiler and breeder management. Animal Front. Rev. Mag. Animal Agricul. **12**(2) (2022)

7. Zeng, A., Yu, T., Song S., et al.: Multiview self-supervised deep learning for 6D pose estimation in the amazon picking challenge. In: 2017 IEEE International Conference on Robotics and Automation (CRAIEEE), pp. 386–383 (2016)

8. Zeng, A., Song, S., Yu, K.T., et al.: Robotic pick-and place of novel objects in clutter with multi affordance grasping and cross-domain image matching. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 1–8 (2018)

9. Ewerton M. Neumann G. Lioutikov Ral. Learning multiple collaborative tasks with amixture of interaction primitives C . EEE International Conference on Robotics & AutomationIEEE 2015. 1535–1542

10. Jingsha, Z., Yan, Z.: Research on automatic control of laser sensors based on artificial intelligence. Laser J. **43**(11), 199–203 (2022)

11. Hongwei, S., Na, L.: Automatic correction of ranging error of laser displacement sensors using artificial intelligence technology. Laser J. **42**(10), 167–170 (2021)

12. Xuetong, R.: Research on sensor technology based on artificial intelligence. Mod. Indust. Econ Inform. **10**(05), 60–61 (2020)

13. Zhiwu, W.: Fault diagnosis technology of sensors based on artificial intelligence methods. Rocket Propulsion **05**, 59–62 (2005)

14. Beizhan, P., Lin Dejie, O., Jincheng.: Application of artificial intelligence in the field of sensors. Sensor Technology **03**, 5–7 (2002)

15. Yan, S., Lei, H., Yan, R.: Design of an automatic calibration system for temperature sensors based on robots. Electronic Measure. Technol. **44**(09), 56–65 (2021)

16. Xianghua, H., Feng, J., Shuiwang, Y., et al.: Application of artificial intelligence in field dynamic calibration of vector thrust. Aerospace Measurem, Technol. **39**(03), 51–57 (2019)