

Article

Power Transformer Breathing System Condition Monitoring Based on Pressure–Temperature Optical Sensing and Deep Learning Method

Jiabi Liang ¹, Jian Shao ¹, Peng Wu ¹, Qun Li ¹, Yuncai Lu ¹, Yalin Wang ² and Zhaokai Lei ^{2,*}

¹ State Grid Jiangsu Electric Power Co., Ltd., Electric Power Research Institute, Nanjing 210024, China

² School of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

* Correspondence: yikeanee@163.com

Abstract

During long-term operation of power transformers, oil temperature and pressure exhibit strong non-stationarity and multi-scale coupling, which makes early-stage breathing system faults difficult to detect accurately. To address this issue, this paper proposes an integrated diagnosis and early-warning method for transformer breathing systems. It combines a multi-parameter optical sensor with a deep-learning algorithm. The pressure–temperature optical sensing system based on Fabry–Pérot (F–P) interferometry and fiber Bragg grating (FBG) technology is developed to achieve high-precision synchronous measurement of pressure and temperature. To handle the non-stationary and multi-scale characteristics of the measured signals, a swarm-intelligence-optimized variational mode decomposition (VMD) method is employed to adaptively decompose time series temperature and pressure data. On this basis, a joint forecasting model integrating a temporal convolutional network (TCN) and an inverted Transformer (iTransformer) is constructed to capture both local temporal dynamics and long-term dependencies. Furthermore, based on the pressure equilibrium mechanism of transformer breathing systems, oil temperature and equivalent oil level are inferred, and abnormality criteria suitable for both multi-point and single-point monitoring are established. Experimental and field tests on a 220 kV transformer demonstrate that the proposed method outperforms conventional models in prediction accuracy.

Keywords: optical sensing; prediction; variational mode decomposition; deep learning; breathing system

Academic Editor: Christophe Volat

Received: 13 January 2026

Revised: 5 February 2026

Accepted: 11 February 2026

Published: 24 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the continuous expansion of power systems and increasingly complex operating environments, transformers play a critical role in voltage transformation and energy transmission. The safety of transformers is directly related to the stability of the power grid and the reliability of the power supply. Internal oil temperature and oil pressure, together with other state variables, reflect key operating characteristics, including thermal stability, insulation condition, and load level [1–3]. Online monitoring and trend prediction of temperature and pressure help reveal potential degradation processes in a timely manner. This supports maintenance decision-making and reduces economic losses and outage risks caused by sudden faults.

During long-term service, the transformer breathing system acts as a critical auxiliary subsystem for pressure regulation and moisture isolation. It maintains the oil–paper insulation environment by balancing internal and external pressure variations. It also prevents moisture and contaminants from entering the transformer through the oil conservator. Breathing-system failures (e.g., breather blockage, diaphragm rupture, conservator air leakage, or malfunction of pressure-relief components) disrupt normal pressure exchange. They may cause abnormal conservator pressure fluctuations, accelerate oil aging, and increase moisture in the insulation system. In practice, these faults often develop gradually and manifest intermittently, making early-stage identification challenging [4,5].

Early studies relied on gas-state equations, oil thermal-expansion models, and pressure-equilibrium relationships. These models describe conservator volume variation and gas exchange driven by oil-temperature changes. Combined with the characteristics of pressure-relief valves, breathers, and drying devices, they provide nominal operating ranges for breathing systems. Severe breathing system faults were commonly identified by setting upper and lower pressure thresholds for the oil conservator or by monitoring abnormal pressure rises and sustained negative pressure conditions. Although these approaches are physically intuitive and straightforward to implement, their diagnostic performance is highly sensitive to model parameters and operating assumptions, limiting their ability to accurately characterize progressive degradation and detect early-stage or concealed faults.

Advances in online monitoring technologies have enabled researchers to assess breathing system conditions using operational features such as oil pressure, oil level variation rate, and breathing frequency [1,5]. For instance, analysis of phase relationships, amplitude ratios, and hysteresis characteristics between oil pressure and oil temperature curves can indicate the presence of blockage or leakage within the breathing path. Similarly, statistical evaluation of the correspondence between breathing frequency and load variation can provide insights into gas exchange capability. While these methods improve the identification of abnormal operating states to some extent, most remain dependent on empirical rules or single-variable statistical indicators, making it difficult to comprehensively describe coupled evolutionary behaviors under complex operating conditions.

With the rapid development of AI techniques for power equipment assessment, machine learning and deep learning have been increasingly used to model breathing-system-related variables. Existing studies have applied recurrent models (e.g., LSTM/GRU), convolution-based models (e.g., CNN/TCN), and attention-based Transformers to predict oil temperature and pressure. These data-driven approaches can learn nonlinear temperature–pressure relationships and capture dynamic variations under different operating conditions. Compared with rule-based methods or simplified physical models, they are more effective for nonstationary time series, multi-scale fluctuations, and complex coupling among variables.

Nevertheless, each model family has limitations. RNN-based models may suffer from degraded gradients and delayed responses for long sequences, convolution-based models are limited in capturing long-range dependencies, and standard Transformers can be sensitive to data quality and architecture choices in high-frequency industrial monitoring. In addition, interpretability and engineering deployability remain critical concerns. Therefore, it is necessary to develop a practical assessment method that balances accuracy, robustness, and applicability. Motivated by this need, we propose a framework that integrates signal decomposition with deep learning to better characterize the state evolution of transformer breathing systems [6–8].

This paper presents an integrated monitoring and predictive analytics framework for transformer breathing systems. It focuses on developing and implementing a multi-parameter optical sensing solution combined with a deep-learning-based forecasting model

to enable accurate state assessment and early fault identification. Specifically, the research involves: (1) the development and experimental validation of a Fabry–Pérot interferometer and Fiber Bragg Grating-based optical sensor for simultaneous in situ measurement of pressure and temperature within transformer oil. (2) the establishment of a Swarm Intelligence-optimized Variational Mode Decomposition, Temporal Convolutional Network, and inverted-structure Transformer (Swarm Intelligence-VMD-TCN-iTransformer) model for high-precision prediction of pressure and temperature trends. (3) the proposal of novel fault diagnosis criteria for breathing system anomalies, such as blockages, based on the predicted pressure information and derived parameters like calculated oil temperature and equivalent oil level. The ultimate goal is to enhance the reliability of power transformers through a multi-parameter optical sensing and a data-driven approach that facilitates proactive maintenance.

2. Optical Sensing Technologies for Transformer Breathing System Monitoring

2.1. Pressure and Temperature Sensing

(1) Pressure Sensing Technology Based on Fabry–Pérot Effect

The Fabry–Pérot (F–P) interferometer is an important optical instrument based on the principle of multiple-light interference and is widely used for the analysis of fine spectral line structures. The interferometer consists of two mutually parallel plates, as shown in Figure 1. To enhance the interface reflectivity, the inner surfaces of the M_1 and M_2 interfaces are coated with silver films in certain applications [9–12].

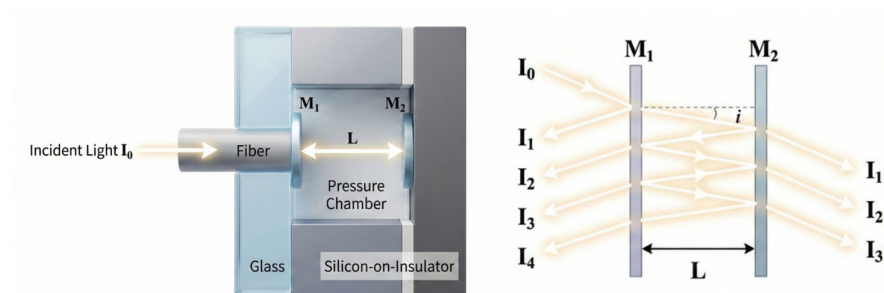


Figure 1. Schematic Diagram of Fabry–Pérot Interferometer.

When a light I_0 is incident at a certain angle onto the parallel surfaces of a medium with refractive index n and thickness L , the incident light is split at the upper surface M_1 into a reflected light and a refracted light. The refracted light is reflected at the lower surface M_2 , while part of it is transmitted out. This process repeats, resulting in an infinite series of reflected lights I and transmitted lights I' . Let i denote the angle of incidence within the cavity. The reflection and transmission coefficients from outside to inside the interface are denoted by r and t , respectively, while r' and t' represent the reflection and transmission coefficients from inside to outside. The intensity reflection coefficient is given by $R = r^2$.

If the amplitude of the incident light I_0 is A , then the amplitudes of the first reflected and transmitted beams are $A \cdot r$ and $A \cdot t$, respectively. By continuing this process, the amplitude of the reflected beams can be expressed as:

$$\left\{ \begin{array}{l} A_1 = A \cdot r \\ A_2 = A \cdot t \cdot r' \cdot t' \\ A_3 = A \cdot t \cdot r'^3 \cdot t' \\ \dots \end{array} \right\} \left\{ \begin{array}{l} A_1 = Ar \\ A_2 = Atr't' \\ A_3 = Atr'r'r't' = Atr'^3t' \\ \dots \end{array} \right. \quad (1)$$

The amplitude of the transmitted beams is:

$$\left\{ \begin{array}{l} A'_1 = A \cdot t \cdot t' \\ A'_2 = A \cdot t \cdot r'^2 \cdot t' \\ A'_3 = A \cdot t \cdot r'^4 \cdot t' \\ \dots \end{array} \right. \quad (2)$$

When the reflection coefficient is much smaller than 1 and $t \approx t' \approx 1$, the amplitudes of the first two reflected lights are approximately equal and significantly larger than those of the subsequent reflected lights. Under this condition, it is sufficient to consider only the first and second reflected lights. However, when the reflection coefficient r is relatively large, the infinitely many reflected lights generated within the cavity must be taken into account, and their amplitudes should be summed to obtain the total optical field amplitude. By applying the summation formula of an infinite geometric series, the transmitted light intensity can be expressed as

$$I_T = \frac{I_0}{1 + \frac{4R \sin^2(\phi/2)}{(1-R)^2}} \quad (3)$$

The reflected light intensity can be expressed as:

$$I_R = I_0 - I_T = \frac{I_0}{1 + \frac{4R \sin^2(\phi/2)}{(1-R)^2}} \quad (4)$$

$$\phi = \frac{2\pi\Delta}{\lambda} = \frac{4\pi nL \cos i}{\lambda} \quad (5)$$

where I_0 is the incident light intensity, R is the end face reflectance, λ is the wavelength of the light wave, ϕ denotes the phase difference between two lights. When the reflectivity is much smaller than 1, and the two parallel reflecting surfaces have identical reflectivities, Equation (4) can be further simplified as:

$$I_R = 2R(1 - \cos \frac{4\pi L}{\lambda})I_0 \quad (6)$$

In practical engineering applications, various demodulation methods, such as intensity demodulation and wavelength demodulation, are commonly employed. Since intensity demodulation is highly sensitive to the selection of the operating point and is easily affected by light source fluctuations and system drift, it is not suitable for the research and application of high-precision sensors. Therefore, the sensor described in this paper adopts

a wavelength demodulation method. Specifically, the interferometric intensity spectrum is transformed from the wavelength domain into the frequency domain, where the frequency-domain signal is acquired and processed using a fast Fourier transform. Subsequently, the characteristic frequency corresponding to the interference spectrum is determined from the peak position in the Fourier spectrum, and the cavity length is inversely calculated accordingly. The basic computational procedure is outlined as follows.

By substituting the relationship between the optical wavelength λ , optical frequency ν , and the speed of light c , i.e., $\lambda = c/\nu$, into Equation (6), the reflected light intensity can be expressed as a function of frequency and cavity length:

$$I_R\left(\frac{c}{\nu}\right) = 2R\left(1 - \cos\frac{4\pi\nu L}{c}\right)I_0 \quad (7)$$

According to Equation (7), the frequency f corresponding to the output signal is given by:

$$f = \frac{2L\nu}{c} \quad (8)$$

By performing a fast Fourier transform on Equation (7), the characteristic frequency corresponding to the cavity length can be obtained. Subsequently, the absolute cavity length of the Fabry–Pérot cavity is calculated according to Equation (8), and the cavity length is then converted into pressure using the optical model or calibration curve of the sensor. This method features high demodulation speed, high measurement accuracy, and high resolution, making it well-suited for engineering applications of high-precision Fabry–Pérot sensors.

(2) Temperature Sensing Technology Based on Fiber Bragg Grating

When the incident light emitted from a broadband light source illuminates a Fiber Bragg Grating (FBG), a narrowband component that satisfies the Bragg condition is reflected, and its central reflected wavelength is determined by the grating period. When the ambient temperature changes, both the grating period and the effective refractive index of the fiber Bragg grating vary accordingly, resulting in a shift in the reflected central wavelength. The demodulator converts the sensing signal encoded in the wavelength domain into a digital signal, which is then processed by a computer to achieve temperature calculation and real-time monitoring [13,14].

For a uniform FBG, the coupling coefficient is constant, and the derivative of the phase is zero. When the light wave is incident axially along the z -axis onto the fiber grating, and the fiber's self-coupling coefficient is zero, the center wavelength corresponding to the maximum reflectivity is given by:

$$\lambda_{\max} = \left(1 + \frac{\delta n_{\text{eff}}}{n_{\text{eff}}}\right)(2n_{\text{eff}}\Lambda) \quad (9)$$

From the wavelength matching condition, it can be derived that:

$$\lambda_B = 2n_{\text{eff}}\Lambda \quad (10)$$

where n_{eff} denotes the effective refractive index of the fiber core, δn_{eff} represents the variation in the effective refractive index, and Λ is the grating period. The central reflected wavelength of the Fiber Bragg Grating varies with changes in n_{eff} and Λ . By differentiating Equation (10), the following expression can be obtained:

$$\Delta\lambda_B = \lambda_B(\xi + \alpha_f)\Delta T \quad (11)$$

where ξ is the thermal optical coefficient of the optical fiber, α_f is the coefficient of thermal expansion of the optical fiber, ΔT represents the temperature change measured by the FBG. It can be seen that changing the temperature surrounding the FBG will affect its structure. These effects are manifested by shifts in the FBG's center wavelength. Figure 2 illustrates the sensing principle of the FBG.

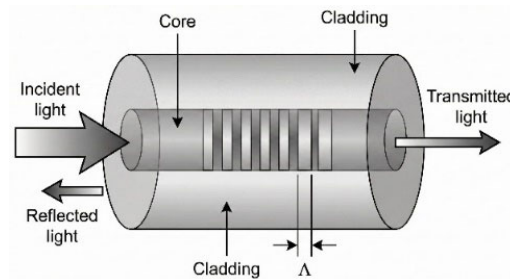


Figure 2. Schematic Diagram of the FBG Sensing.

In analyzing the sensing model of an FBG, the influence of other external factors can be neglected, and only the effect of temperature on the Bragg wavelength λ_B is considered. Under this assumption, the temperature sensitivity of the FBG is mainly determined by the thermo-optic coefficient ξ and the thermal expansion coefficient α_f of the fiber material. Since FBGs are typically fabricated using silica optical fibers with relatively consistent material properties, the temperature sensitivities of different FBGs exhibit good similarity with high linearity.

2.2. Pressure and Temperature Demodulation Method

For a composite FBG–Fabry–Pérot interferometer (FPI) sensor, the measured spectrum is a superposition of the FBG reflection spectrum and the FPI interference spectrum. In principle, the two components can be separated to estimate the Bragg wavelength and the FPI cavity length. However, this separation is computationally complex and may reduce accuracy. Moreover, because the composite signal exhibits strong cross-sensitivity to temperature and pressure, direct decoupling is not preferred for engineering applications [15–17].

The FBG sensor measures temperature by monitoring the shift in the center wavelength of its reflection spectrum. Temperature changes induce both thermal expansion and thermo-optic effects in the fiber material, which in turn alter n_{eff} and Λ , resulting in a shift in the Bragg wavelength. The relationship between the wavelength shift $\Delta\lambda_B$ and temperature change ΔT can be expressed as Equation (11). In practical applications, high-resolution wavelength demodulation techniques—such as peak detection, Gaussian fitting, or central wavelength matching—are used to accurately determine the ambient temperature from the FBG reflection spectrum.

The cavity length of the FPI is not only influenced by pressure but is also susceptible to temperature-induced variations, including thermal expansion of the diaphragm material and the temperature dependence of the refractive index within the cavity. As a result, its output wavelength often contains coupled components of both pressure and temperature and can be regarded as a function of temperature, varying gradually with temperature. Therefore, the cavity length L can be expressed as:

$$L(T) = L_0(T) + L'(T) \quad (12)$$

where $L_0(T)$ represents the static cavity length varying with temperature, and $L'(T)$ represents the dynamic change in cavity length caused by pressure, which also varies with temperature. This sensor adopts a decoupling strategy that combines temperature

measurement based on FBG and pressure measurement based on FPI and performs temperature compensation. Specifically, the ambient temperature T is precisely obtained first through the demodulation of the FBG signal, and this temperature serves as the external input parameter of the F-P pressure demodulation module. Since the temperature sensitivity of the F-P sensor can be pre-characterized through experimental calibration, the initial cavity length variation obtained by demodulation of the F-P interference signal based on the temperature provided by the fiber grating can be corrected to compensate for the cavity length displacement caused by temperature.

To obtain the correlation of temperature and pressure on the FPI characteristics, a test was performed to measure the wavelength of the FPI sensor under various temperatures and pressures. The schematic diagram for the relationship between the F-P reflection wavelength and pressure–temperature, and the test results are shown in Figure 3. After determining the temperature, the curve of the corresponding temperature can be queried to match the pressure and wavelength one by one. In this way, the output of the F-P sensor is extracted from the original pressure–temperature coupling signal, generating a pure pressure measurement. The flowchart of the sensor for measuring temperature and pressure is shown in Figure 4.

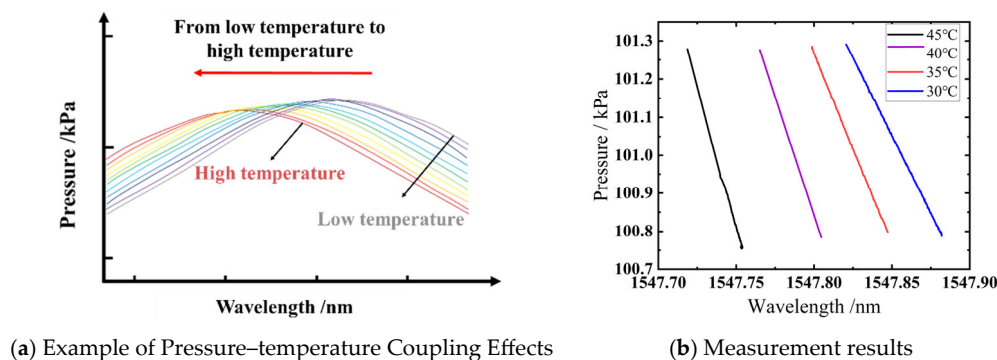


Figure 3. Effects of Pressure and Temperature on the F-P Wavelength.

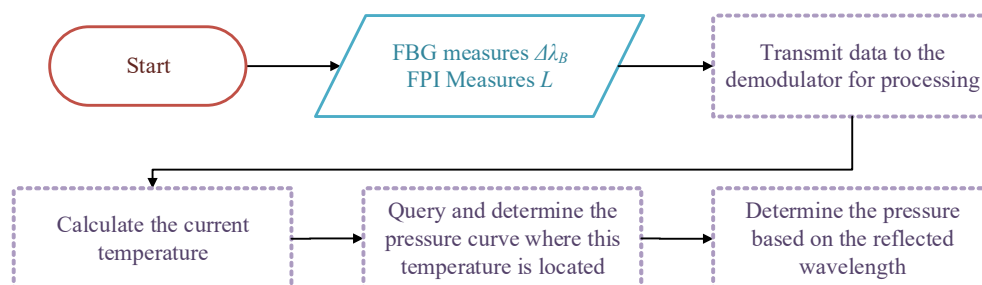


Figure 4. The Flowchart of Demodulation of Temperature and Pressure.

In the proposed sensing system, distinct demodulation strategies are employed for temperature and pressure signals based on their respective sensing mechanisms. Specifically, the FBG temperature sensor utilizes a wavelength-based demodulation method by tracking the Bragg wavelength shift, which provides a direct and stable mapping between wavelength variation and temperature. In contrast, the Fabry–Pérot pressure sensor employs a spectral interferometry-based demodulation method, where the interference spectrum is analyzed in the frequency domain to extract cavity length variations and convert them into pressure values. Regarding the optical source, the system operates in the near-infrared C-band with a central wavelength around 1550 nm. This wavelength range is

selected to minimize transmission loss in standard single-mode optical fibers and to leverage the maturity and stability of commercial optoelectronic components in this band. Although the demodulation algorithms rely on wavelength shifts and interference phases rather than absolute wavelength values, the choice of the C-band ensures optimal long-term monitoring performance.

It is worth noting that some FBG interrogation techniques, such as peak detection, Gaussian fitting, or center wavelength matching, were discarded due to their fundamental incompatibility with the complex, multi-peaked spectra of Fabry–Perot sensors, where they introduce unacceptable nonlinearity, ambiguity, and instability. Spectroscopy was selected as the superior alternative because it directly aligns with the physics of F-P interference, enabling linear, absolute, and robust cavity length measurement. This approach eliminates mode-hopping errors, provides an extensive dynamic range, and delivers the precision required for sensitive physical sensing applications.

2.3. Fabrication of Pressure and Temperature Optical Sensors

In this study, the F–P pressure sensor was fabricated based on MEMS processing technology. Its core components include a circular sensitive diaphragm with good elastic response characteristics and a stable interferometric cavity structure. The specific fabrication procedure is as follows. First, a double-sided polished silicon wafer was selected as the substrate material. The geometric shapes of the sensitive diaphragm and cavity region were defined on the silicon wafer surface using standard photolithography. Subsequently, deep reactive ion etching (DRIE) was applied to perform anisotropic etching of the silicon wafer, forming a circular sensitive diaphragm with a designed thickness and the corresponding cavity structure. By precisely controlling the etching depth and diaphragm thickness, high-pressure sensitivity and good linear response characteristics can be achieved while maintaining sufficient mechanical strength.

After completion of the MEMS cavity structure fabrication, optical coating treatments were applied to the cavity end face and the surface of the sensitive diaphragm. High-reflectivity metallic or dielectric thin films were deposited on the cavity reflecting surfaces using physical vapor deposition (PVD) technology to enhance the reflectivity of the F–P interferometric cavity and the contrast of the interference fringes, thereby improving the stability and signal-to-noise ratio of the interference signal. After coating, the MEMS chip was cleaned and dried to remove potential contaminants introduced during processing, ensuring the reliability of subsequent optical assembly. After fiber cleaving and polishing, the fiber end face was coaxially aligned with the cavity in the MEMS chip along the axial direction to ensure the formation of a stable F–P interferometric cavity between the fiber end face and the sensitive diaphragm. After alignment, ultraviolet-curable adhesive was used to fix and seal the optical fiber to the MEMS chip, and rapid curing was completed under ultraviolet irradiation, thereby improving the mechanical stability and vibration resistance of the structure. Ultimately, an F–P pressure sensing unit based on a MEMS sensitive diaphragm was formed, whose cavity length can undergo repeatable and measurable modulation in response to external pressure variations, as shown in Figure 5a.

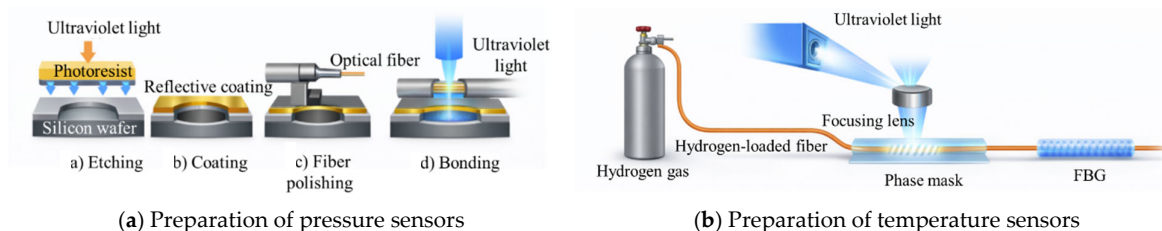


Figure 5. Fabrication of Optical Sensors.

The FBG temperature sensor was fabricated using the ultraviolet laser phase mask method. During fabrication, the single-mode optical fiber was first hydrogen-loaded to enhance its photosensitivity to ultraviolet laser irradiation. Subsequently, under stable ultraviolet laser exposure, periodic refractive index modulation structures were inscribed in the fiber core through a phase mask, forming a uniform fiber Bragg grating. By precisely controlling the ultraviolet exposure energy, exposure time, and phase mask parameters, the central reflected wavelength and spectral profile of the FBG can be accurately tailored. After fabrication, the grating was annealed to improve its long-term stability under high-temperature conditions, thereby obtaining an FBG temperature sensing unit suitable for transformer operating environments, as shown in Figure 5b.

To enable the optical fiber sensors to operate under complex conditions, such as high temperature and high pressure, appropriate armoring and protection of the optical fiber are required to prevent direct exposure to harsh environments, while also imposing higher requirements on the mechanical strength of the sensor. The sensor probe is externally protected by a stainless steel sleeve, which effectively ensures the safety of the optical fiber sensor. A threaded opening is machined at the front end of the steel tube, allowing the sensor to directly interface with the external environment while ensuring that the pressure response time of the F–P cavity is not constrained. The rear end of the sensor is connected to the demodulator via an optical fiber and adopts a double-layer armored structure: the inner layer consists of polyimide-coated optical fiber, while the outer layer is a composite shielding layer composed of stainless steel braided mesh and aluminum foil. The packaging structure of the sensor is shown in Figure 6.

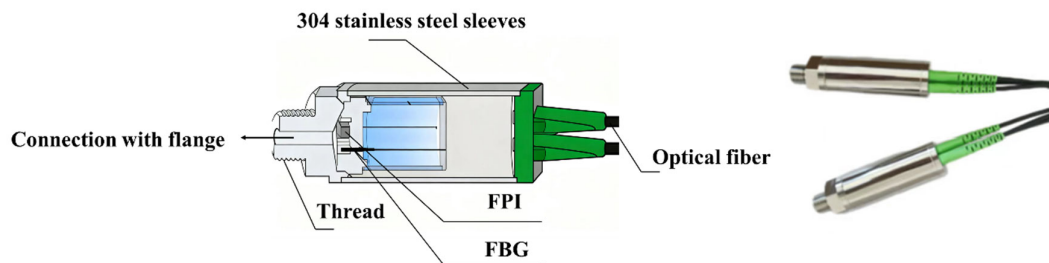


Figure 6. Schematic of the Sensor Packaging Structure.

2.4. Performance Testing of Optical Sensors

To verify the effectiveness of MEMS-based multi-parameter optical fiber sensors, a pressure–temperature multi-parameter optical sensor test platform was established, as shown in Figure 7. The platform is composed of an acrylic shell with a length of 50 cm, a width of 50 cm, and a height of 60 cm to form the oil tank. The tank is filled with transformer oil inside, and sensors are installed at the top, middle, and bottom positions through flanges at the front of the oil tank. The back of the oil tank is connected to the circulating oil device with valves and pumps, which can separately control the inflow and outflow of oil. In addition, a breathing valve for the transformer is also installed on the back. When the oil temperature inside the oil tank changes, leading to a pressure change in the air above the oil, it balances the internal and external air pressure and prevents moisture and impurities in the air from entering the oil. The top of the oil tank is a round acrylic plate, which is fixed to the oil tank with an O-ring and screws to seal it. A hole is made on the top to install a heating rod to heat the transformer oil, thereby simulating the internal overheating fault of the transformer. Through comparative experiments conducted under different operating conditions using a calibrated thermometer and pressure gauge, the measured parameters are summarized in Table 1. The results indicate that the

measurement error of the proposed multi-parameter optical sensor does not exceed 0.02% of the full-scale range (approximately 60 Pa).

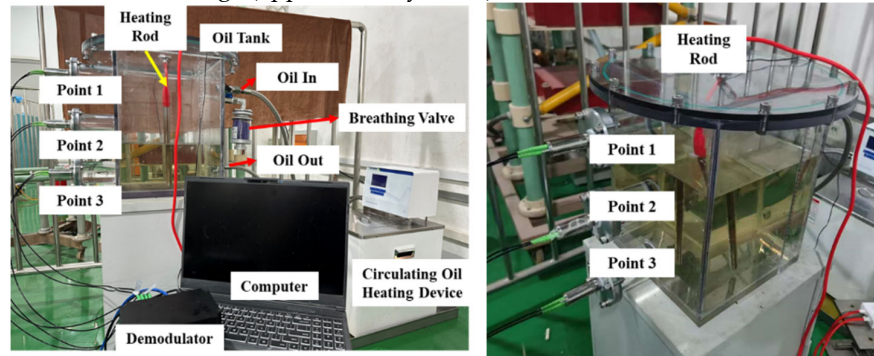


Figure 7. Multi-Parameter Optical Sensing Test Platform.

Table 1. Test results of sensors under different working conditions.

	Temperature/°C						Pressure/kPa					
Actual value	−55	20	50	100	150	200	0.1	10	50	100	150	250
Measured value	−54.3	20.45	50.35	100.3	150.38	200.43	0.101	9.991	50.013	99.995	149.988	249.989
Error value	0.5	0.5	0.45	0.4	0.45	0.5	0.002	−0.009	0.013	−0.005	−0.012	−0.011

In addition, acoustic signals at different frequencies were generated by driving a sound source with a signal generator, and the signal intensity was adjusted by varying the driving voltage using a high-voltage amplifier. The diaphragm vibration amplitude was measured using a laser vibrometer and compared with the diaphragm amplitude obtained from the optical sensor. The laser vibrometer operated with a measurement range of 50 nm, and the average of three randomly selected measurements was taken as the diaphragm amplitude under a given sound pressure. For the optical sensor demodulation unit, the average of 50 demodulation results was used as the diaphragm amplitude demodulated under the same sound pressure. Tests were conducted using sound sources at 30 kHz and 30 Hz. The experimental results are shown in Figure 8.

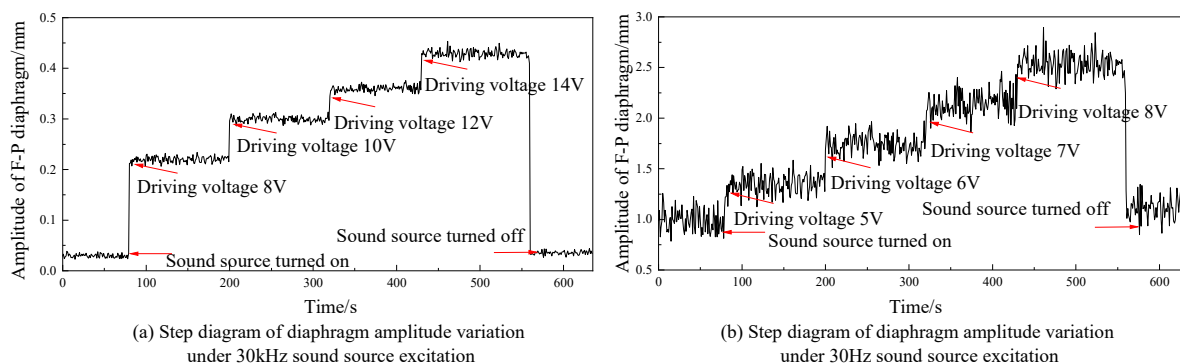


Figure 8. Experimental Results.

The results show that when the sound source is not activated, the output of the sensor remains at a relatively stable state. After the calculation, the error is less than 0.02%, so these noises can be ignored. As the sound source is gradually driven, the output of the sensor shows a clear stepwise increase. Within each steady-state platform, the signal

fluctuation amplitude is small, and no obvious sudden changes or abnormal jumps are observed, which indicates that the sensor has good output consistency and anti-noise ability under steady-state conditions. The step boundaries between different driving stages are well defined, and the mean values of adjacent plateaus differ significantly, demonstrating a high resolution in response to variations in external excitation. After the acoustic source is turned off, the output signal rapidly returns to a level close to the initial value without noticeable lag or drift.

3. Fault Identification of Transformer Breathing Systems

3.1. Principle of Pressure Equilibrium in the Breathing System

The transformer breathing system consists of the external atmosphere, the breather, the rubber bag (or diaphragm), the oil conservator, the transformer tank, and the associated piping. Its primary function is to maintain a dynamic balance between the internal pressure of the transformer and the external atmospheric pressure. The breather generates a pressure difference through the oil level variation in the oil seal cup, thereby enabling the exhalation and inhalation processes. The specific operating principle is as follows:

(1) Exhalation Process

When the transformer load increases or the temperature rises, the oil volume expands, leading to an increase in internal pressure. From the pressure equilibrium relationship, it can be obtained that:

$$P = P_0 + \rho_0 g h_{ex} + \rho_{oil} g \Delta h \quad (13)$$

where P_0 denotes the external atmospheric pressure, $\rho_0 g h_{ex}$ represents the pressure difference generated by the oil level difference in the oil seal cup during the exhalation process, $\rho_{oil} g \Delta h$ denotes the pressure difference caused by the height difference between the oil level in the conservator and the specified position in the transformer tank.

The increase in internal pressure compresses the rubber bag, increasing the internal pressure of the bag. When the condition $P_{bag} > P_0 + \rho_0 g h_{ex}$ is satisfied, air is discharged from the breather to the external atmosphere, thereby restoring pressure equilibrium.

(2) Inhalation Process

When the transformer load decreases or the temperature drops, the oil volume contracts, leading to a reduction in internal pressure. Similarly, it can be obtained that:

$$P = P_0 - \rho_0 g h_{in} + \rho_{oil} g \Delta h \quad (14)$$

where $\rho_0 g h_{in}$ represents the pressure difference generated by the oil level difference in the oil seal cup during the inhalation process.

The decrease in internal pressure causes the rubber bag to expand, resulting in a reduction in the internal pressure of the bag. When the condition $P_{bag} < P_0 - \rho_0 g h_{in}$ is satisfied, the breather draws air from the external atmosphere, thereby restoring pressure equilibrium.

Based on the above analysis, the liquid level height in the oil seal cup is generally on the order of a few centimeters. Therefore, the gas pressure inside the rubber bag is essentially maintained close to atmospheric pressure. The transformer can be treated as a closed cavity, with the breathing system being equivalent to a valve connecting the transformer to the external environment. The principle of pressure balance in the breathing system is shown in Figure 9.

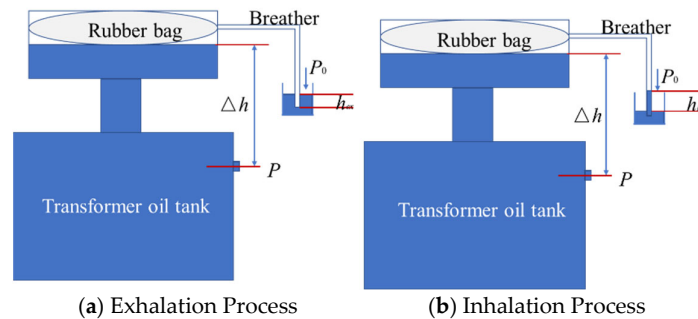


Figure 9. Principle of Pressure Equilibrium in the Breathing System.

3.2. Fault Analysis of Transformer Breathing Systems Based on Optical Temperature and Pressure Sensing

(1) Calculate the average oil temperature of the transformer based on the bottom pressure value

Given the sensor installation heights H_1 , H_2 , and H_3 (from top to bottom) and the measured pressures P_1 , P_2 , and P_3 , the average oil temperatures T_1 and T_3 between the sensors can be calculated based on hydrostatics. Here, T_1 and T_3 represent the transformer oil temperatures below and above the middle sensor, respectively:

$$\begin{cases} P_3 - P_2 = \rho_3 g(H_2 - H_3) = \rho_{20} [1 - \beta(\frac{T_2 + T_3}{2} - 20)] g(H_2 - H_3) \\ P_2 - P_1 = \rho_2 g(H_1 - H_2) = \rho_{20} [1 - \beta(\frac{T_2 + T_1}{2} - 20)] g(H_1 - H_2) \end{cases} \quad (15)$$

(2) Determine whether the calculated temperature is normal

Plot the calculated transformer oil temperature together with the sensor-measured temperature and examine whether their trends and magnitudes are consistent. If the calculated oil temperature is significantly higher than the sensor-measured oil temperature, it can be used as a fault criterion.

(3) Calculate the equivalent oil level of the transformer

Given the known positions of the upper, middle, and lower sensors, the actual oil level can be computed using the pressure relationship, and the sensor readings can be reused for cross-validation. The transformer's equivalent oil level can then be used to assess potential fault conditions.

$$P_1 = \rho_{20} [1 - \beta(T_1 - 20)] g(H_{oil} - H_1) \quad (16)$$

Based on the above formulation, the equivalent oil level of the transformer can be calculated. When multiple sensors are deployed, multi-point transformer data are available, enabling fault identification directly via the equivalent oil level. Corresponding thresholds can be defined, and real-time early warning can be implemented based on the absolute pressure values.

When only a single sensor is used, relative-pressure information cannot be leveraged to infer the internal condition of the transformer. In this case, differential processing of the measured pressure can be applied. If pronounced, non-random spike-like features appear in the differential signal, they can be interpreted as abnormal oil-level fluctuations. By setting appropriate thresholds, potential blockage conditions can be identified. The setup of the multi-optical sensor system for breathing system faults is shown in Figure 10.

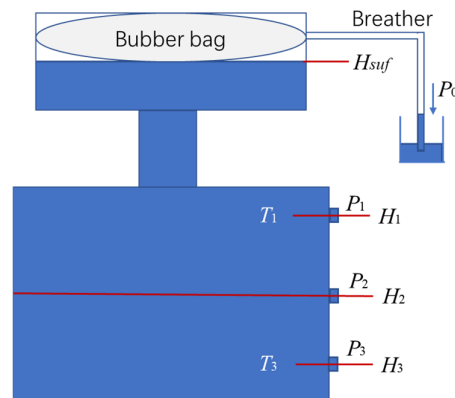


Figure 10. Multi-Optical Sensor Setup for Breathing System Fault Diagnosis.

4. Deep Learning Data-Driven Pressure–Temperature Forecasting of Power Transformer

Based on synchronous pressure–temperature measurement and the breathing-system pressure-equilibrium mechanism, the operating state of the transformer breathing system can be characterized from a physical perspective. Section 3 establishes quasi-static criteria using multi-point pressure to infer oil temperature and equivalent oil level. However, under practical operating conditions, temperature and pressure are influenced by load fluctuations, ambient changes, and thermal inertia. Their evolution is highly nonstationary and multi-scale. Therefore, relying only on real-time measurements and static threshold rules is insufficient for early warning.

4.1. Algorithm Architecture

During long-term operation of transformers, oil temperature and pressure are jointly influenced by load variations, heat dissipation conditions, environmental factors, and equipment aging. As a result, their time-series data typically exhibit pronounced nonstationarity, multi-scale fluctuations, and abrupt changes. Such complex evolutionary characteristics make it difficult for traditional prediction methods based on physical mechanisms or statistical assumptions to simultaneously ensure high accuracy and adequate dynamic responsiveness. In this section, a deep learning-based data-driven model for transformer oil temperature and pressure prediction is developed. By integrating signal decomposition with deep neural network modeling, the proposed approach enables high-precision characterization of multi-scale temporal evolution patterns.

(1) Variational Mode Decomposition

Variational Mode Decomposition (VMD) is an adaptive signal decomposition method grounded in variational theory, whose core idea is to decompose a complex nonstationary signal into a set of mode components with finite bandwidths and well-defined center frequencies. Compared with traditional empirical mode decomposition methods, VMD imposes global constraints on modal bandwidths in the frequency domain, thereby effectively avoiding endpoint effects and mode mixing. As a result, it exhibits superior decomposition stability and stronger mathematical interpretability.

Let the original time-series signal be denoted as $f(t)$. The objective of VMD is to decompose $f(t)$ into K mode components $\{u_k(t)\}_{k=1}^K$, where each mode is compactly distributed around its corresponding center frequency ω_k . To this end, VMD applies the Hilbert transform to each mode to construct its analytic signal and performs frequency shifting in the frequency domain to translate the spectrum of each mode to the baseband, leading to the following variational optimization model:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^K u_k(t) = f(t) \end{cases} \quad (17)$$

where $\delta(t)$ denotes the Dirac delta function, $*$ represents the convolution operator, and $\partial_t(\cdot)$ is the temporal differentiation operator. The constraint in the above formulation ensures that the sum of all mode components can exactly reconstruct the original signal. To solve the above constrained variational problem, a quadratic penalty term and a Lagrange multiplier $\lambda(t)$ are typically introduced to transform it into an unconstrained optimization problem. The corresponding augmented Lagrangian function can be expressed as:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \\ & \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle \end{aligned} \quad (18)$$

where α is the penalty factor, which is used to balance the trade-off between the reconstruction error and the modal bandwidth constraint. Through iterative optimization, the modal center frequencies can be adaptively updated in the frequency domain using an energy-weighted averaging scheme:

$$\omega_k^{(n+1)} = \frac{\int_0^\infty \omega |\hat{u}_k^{(n+1)}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{(n+1)}(\omega)|^2 d\omega} \quad (19)$$

where $\hat{u}_k(\omega)$ denotes the Fourier transform of the K mode component $u_k(t)$.

(2) Adaptive Parameter Selection Method for VMD-Based Swarm Intelligence

The decomposition performance of VMD is highly sensitive to the selection of the mode number K and the penalty factor α . If K is chosen too small, different frequency components are prone to mode mixing. Conversely, an excessively large K introduces redundant modes and amplifies noise. Meanwhile, the penalty factor α directly determines the strength of the bandwidth constraint imposed on each mode. An inappropriate value of α may lead to excessive spectral spreading or over-concentration of modal components, thereby degrading subsequent feature extraction and predictive modeling performance. Therefore, to address the limitations of conventional VMD that relies on empirical parameter selection and exhibits poor generalization, this study introduces a swarm intelligence-based optimization strategy to adaptively determine the key VMD parameters K and α .

(1) The VMD parameter selection problem is formulated as a continuous-discrete mixed optimization problem, in which the optimization variables are defined as:

$$\mathbf{x} = [K, \alpha] \quad (20)$$

where $K \in \mathbb{Z}^+$ denotes the number of decomposition modes and $\alpha \in \mathbb{R}^+$ is the penalty factor. Considering engineering experience and computational complexity constraints, the search space is defined as:

$$K \in [K_{\min}, K_{\max}], \alpha \in [\alpha_{\min}, \alpha_{\max}] \quad (21)$$

(2) Considering that VMD serves as the front-end decomposition module of the prediction model in this study, the quality of its parameters is ultimately reflected in the

predictive performance. Therefore, the prediction error on the validation set is adopted as the optimization objective, and the following fitness function is constructed:

$$J(K, \alpha) = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i(K, \alpha)}{y_i} \right| \quad (22)$$

where y_i and $\hat{y}_i(K, \alpha)$ denote the ground-truth value and the predicted value obtained under a given parameter combination (K, α) through VMD, sub-sequence prediction, and reconstruction, respectively, and N is the number of validation samples. This objective function can directly reflect the influence of different parameter configurations on the overall prediction accuracy.

(3) Swarm intelligence optimization algorithms achieve a dynamic balance between global exploration and local exploitation by mimicking cooperative behaviors observed in natural swarms. The algorithm first randomly initializes the positions of the population individuals as follows:

$$\mathbf{x}_i^{(0)} = [K_i^{(0)}, \alpha_i^{(0)}], i = 1, 2, \dots, M \quad (23)$$

where M denotes the population size. Subsequently, during each iteration, the parameter vectors are updated according to the current state of each individual, the global best solution of the population, and a random perturbation term, as follows:

$$\mathbf{x}_i^{(t+1)} = \mathbf{x}_i^{(t)} + \Delta \mathbf{x}_i^{(t)} \quad (24)$$

where the $\Delta \mathbf{x}_i^{(t)}$ is adaptively generated by the internal behavioral rules of the algorithm and is used to guide the search process toward potential optimal regions.

For the discrete variable K , rounding or boundary-mapping strategies are applied after each update to ensure the integer constraint is satisfied; for the continuous variable α , it directly participates in the continuous search process. After each parameter update, the VMD and the training of the prediction model are re-executed, and the corresponding fitness value is computed.

(4) When the maximum number of iterations is reached, or the convergence criterion of the fitness function is satisfied, the global optimal parameter combination is obtained.

$$(K^*, \alpha^*) = \arg \min_{K, \alpha} J(K, \alpha) \quad (25)$$

Based on this optimal parameter combination, the original signal is finally decomposed using VMD, and the resulting multi-scale mode components are obtained as input features for the subsequent prediction model.

(3) Modeling Method Based On Temporal Convolutional Network

After applying swarm intelligence-optimized parameter selection, VMD can decompose the original nonstationary signal into multiple mode components with well-defined frequency-band characteristics, effectively alleviating multi-scale coupling and noise interference in the signal. However, each mode component still exhibits pronounced non-linear temporal evolution and long-term dependency characteristics along the time dimension. Relying solely on linear modeling or conventional convolutional structures is insufficient to fully capture such temporal correlations. Therefore, on the basis of the optimally decomposed VMD modes, this study introduces a Temporal Convolutional Network (TCN) to further model each mode component, thereby enhancing the model's capability to represent complex temporal dynamics.

Let the optimal VMD obtain K^* modes $u_k(t)$, and combine them into multi-channel inputs at the same time:

$$\mathbf{U}(t) = [u_1(t), u_2(t), \dots, u_K(t)]^T \in R^{K*} \quad (26)$$

Then, the TCN-based modeling of the VMD-optimized data and the corresponding output prediction can be directly expressed as:

$$\hat{y}(t + \Delta t) = g(\text{TCN}(\mathbf{U}_{t-L+1:t})) \quad (27)$$

where $\mathbf{U}_{t-L+1:t}$ denotes a historical window of length L , $\text{TCN}(\cdot)$ represents the feature extractor composed of stacked causal dilated convolutions, and $g(\cdot)$ denotes the linear mapping layer.

(4) High-Order Time-Series Prediction Modeling Based on Inverted-Transformer

The feature sequences processed by the TCN have effectively captured the nonlinear dynamic characteristics of each VMD mode at local temporal scales. However, this modeling paradigm primarily relies on convolutional operations within a fixed receptive field, which limits its ability to represent global temporal dependencies across multiple scales and periods. In particular, in power equipment operating data, state variables such as pressure and temperature are often influenced by daily load cycles, environmental variations, and operating condition transitions, resulting in pronounced long-term dependencies and global correlations. To further capture long-range temporal dependencies that are difficult for TCNs to model explicitly, this study introduces an inverted-structure Transformer (iTransformer) for predictive modeling on top of the TCN-based feature extraction.

Let the high-order temporal feature representation output by the TCN be denoted as:

$$\mathbf{H} = \{\mathbf{h}(t-L+1), \dots, \mathbf{h}(t)\} \in R^{L \times d} \quad (28)$$

where d denotes the feature dimension. Unlike conventional Transformers that model attention relationships along the temporal dimension, iTransformer performs self-attention computation in the feature dimension, thereby efficiently modeling global dependencies among different modal features and their cross-scale representations.

In iTransformer, after transposing the dimensions of the TCN output feature matrix $\mathbf{H}$, the Query, Key, and Value representations are constructed as follows:

$$\mathbf{Q} = \mathbf{H}^T \mathbf{W}_Q, \mathbf{K} = \mathbf{H}^T \mathbf{W}_K, \mathbf{V} = \mathbf{H}^T \mathbf{W}_V \quad (29)$$

where $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V$ are learnable parameter matrices. Through the above mechanism, the iTransformer can explicitly model the global dependencies among the multimodal features extracted by the TCN while keeping the computational complexity manageable. After passing through multiple iTransformer encoder layers, a feature representation that integrates global correlations is obtained, and the final prediction is produced via a linear mapping. The flow chart of the proposed prediction algorithm is shown in Figure 11.

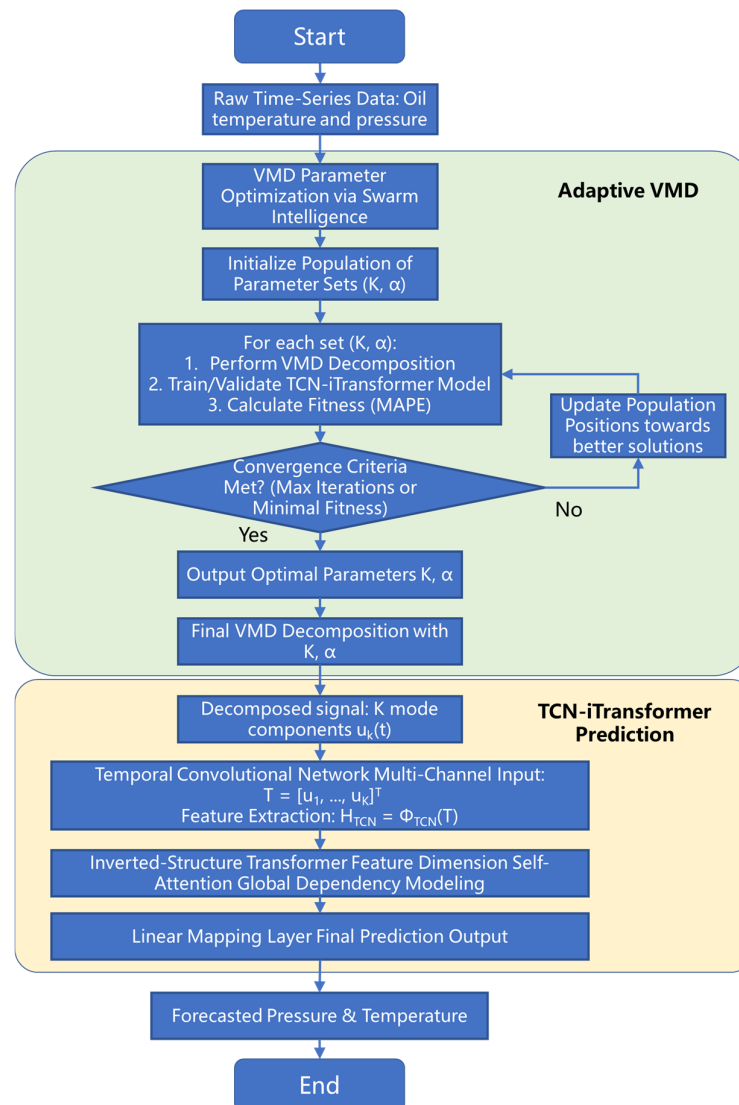


Figure 11. Flow chart of proposed prediction algorithm.

4.2. Algorithm Validation

Particle Swarm Optimization (PSO) is a widely used swarm intelligence algorithm for continuous optimization problems, which simulates the behavior of bird flocks foraging to search for the optimal solution. This article uses PSO as an optimization algorithm. In this study, online monitoring data from a transformer were selected to validate the proposed algorithm. The dataset consists of physical measurements collected over three consecutive days, comprising six groups in total, with each group containing 295,200 data points. The prediction model was implemented using the PyTorch 2.1.0 framework.

During the model training stage, to ensure predictive performance and model stability, the input data were standardized by removing the mean and scaling to unit variance, thereby eliminating the influence of dimensional differences among different physical quantities on the training process. The oil temperature and oil pressure data collected over the three-day period were divided into training and testing sets in chronological order. Data from the first 2.5 days were used for model training, while data from the remaining 0.5 day were reserved for model evaluation, ensuring a reliable assessment of prediction performance on unseen data. Figure 12 presents the validation dataset.

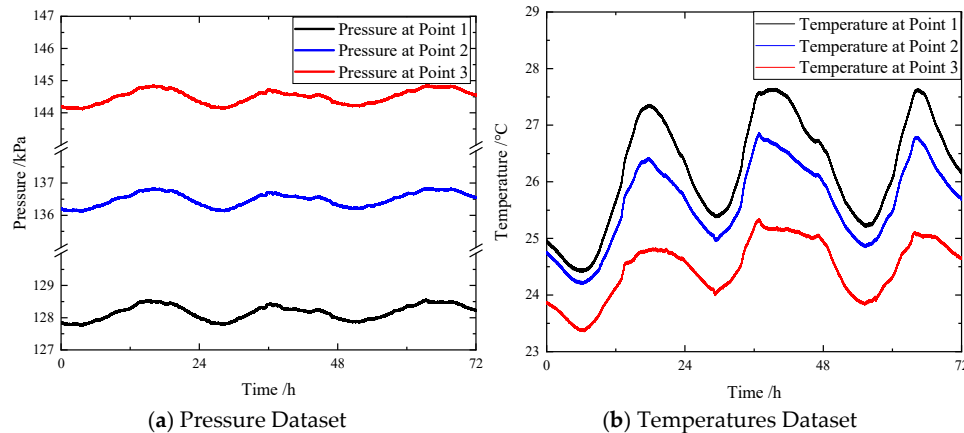


Figure 12. Validation Dataset composed of temperatures and pressures measured through optical sensing of an online power transformer.

In the network architecture, the TCN component employs dilated convolutional structures to expand the receptive field, with the convolution kernel size set to 4 and the number of hidden layers set to 4. For the Transformer component, the feature embedding dimension D is set to 128, the number of heads in the multi-head self-attention mechanism is set to 4, and the dropout rate is set to 0.1.

During training, the Adam optimizer is used to update the model parameters, with an initial learning rate of 0.001. The batch size is set to 64, and the number of training epochs is set to 200. To prevent model overfitting, an early stopping strategy is introduced during training. Specifically, training is terminated when the validation error shows no improvement for 20 consecutive epochs.

Set the initial population size of the Swarm Intelligence to 8 and the number of iterations to 50, thereby obtaining the optimal K value of the temperature signal as 4 and the α value as 4179, the optimal K value of the pressure signal as 6 and the α value as 3150. The iterative curves of the decomposition parameters, their center frequencies, and the optimal K values after Swarm Intelligence-VMD are shown in Figure 13. The original content sequence is decomposed into smooth low-frequency high-amplitude components and regular high-frequency low-amplitude components, and the spectra of each IMF component do not overlap.

The normalized data after modal decomposition is input into the TCN-iTransformer and trained and predicted based on the optimal model. To further verify the effectiveness of the prediction model proposed in this paper, it was compared with the TCN-iTransformer, VMD-TCN-iTransformer, and Swarm Intelligence-VMD-TCN-iTransformer models. The temperature and pressure prediction results at each point are shown in Figure 14.

From the overall prediction results, the three models show significant differences in the fitting accuracy and dynamic response capability of the transformer oil temperature and oil conservator pressure signals.

As can be seen from Figure 15, the TCN-iTransformer model predicts the original signal directly. The prediction value cannot follow the fluctuation of the real value, and the prediction error MAPE and RMSE are 19.61 and 0.23, respectively. If the VMD of the original signal is used and then the prediction is made, the prediction results show significant improvement, and the prediction error decreases by 12.13 and 0.1, indicating that VMD can effectively reduce the non-stationarity of the original data to further improve the prediction accuracy. Compared with the VMD-TCN-iTransformer model, the model in this paper adopts Swarm Intelligence to optimize the modal decomposition results of

the VMD, which further improves the order of the modal decomposition parameters and helps to further improve the prediction accuracy. The MAPE and RMSE of this paper's model are 3.48 and 0.04, respectively, which are the lowest among these prediction results.

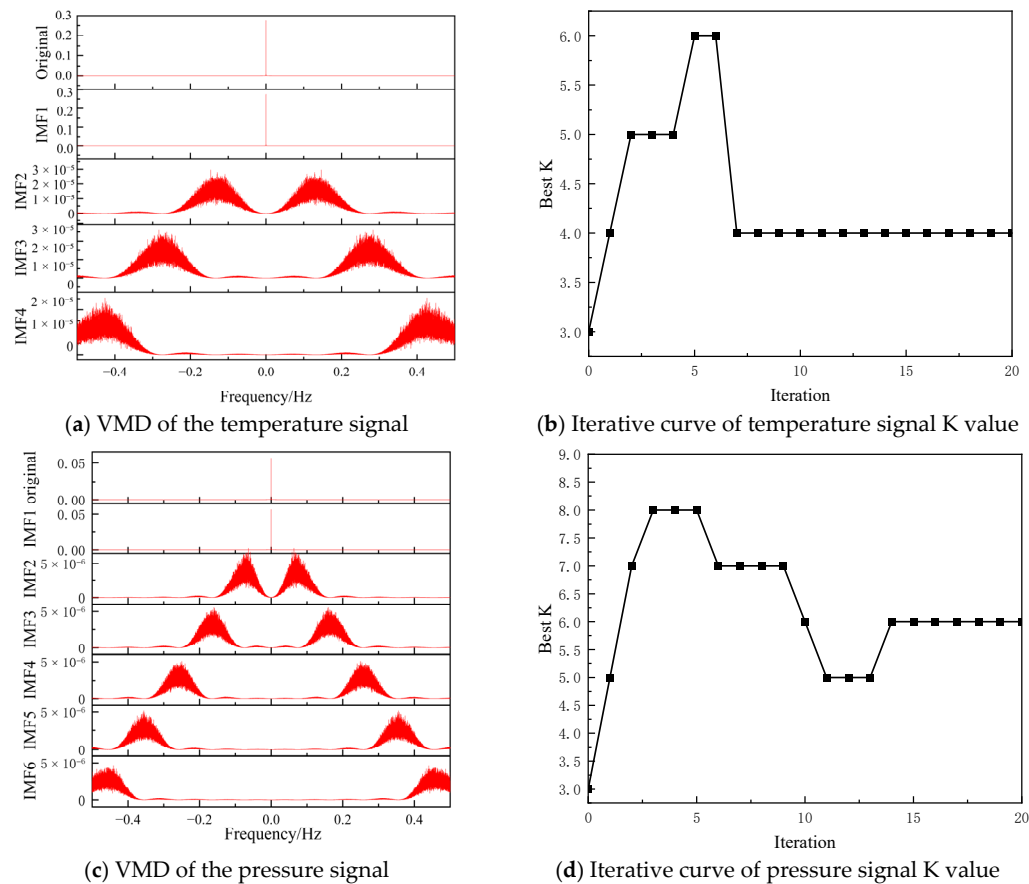


Figure 13. VMD Result.

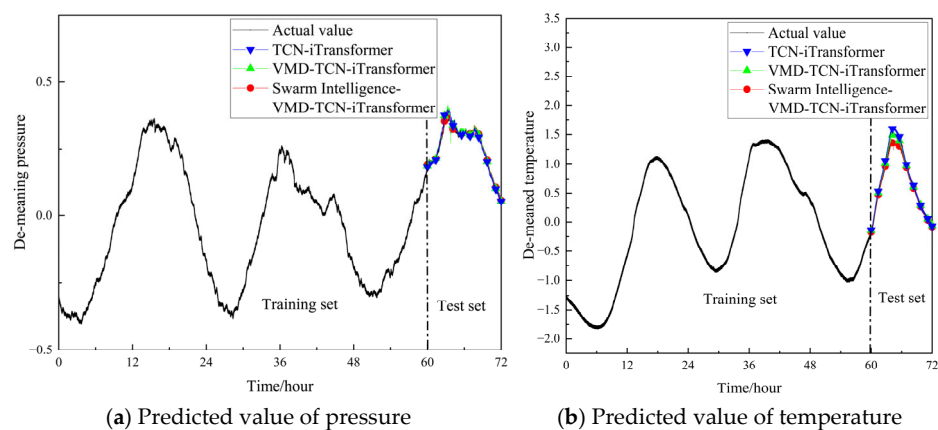


Figure 14. Prediction Results of Temperature and Pressure.

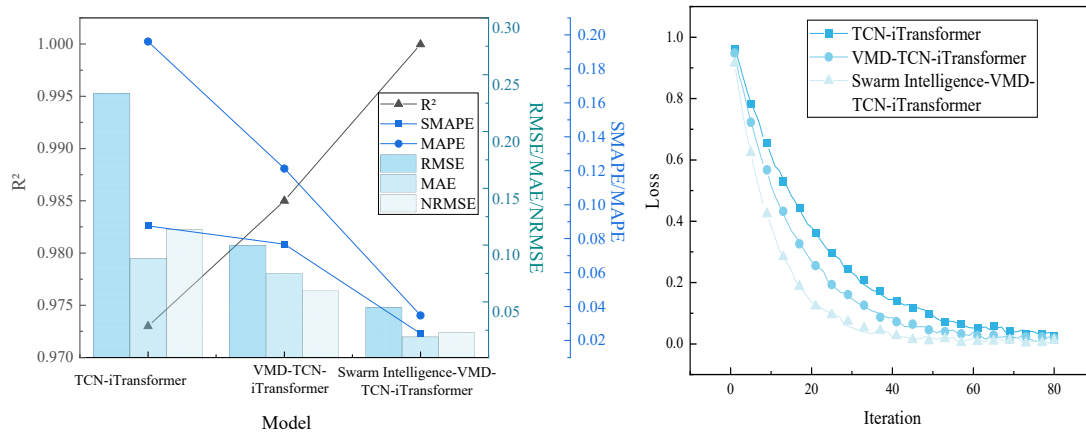


Figure 15. The Evaluation Indicators of Three Different Models.

In addition to the aggregate performance metrics, the statistical dispersion of the prediction residuals was analyzed to evaluate the model's reliability under varying conditions. Based on the calculation of 100 consecutive test groups, the proposed model achieved a Coefficient of Determination (R^2) of approximately 0.995, with residuals centered closely around zero.

As shown in Figure 16, the Proposed Swarm Intelligence-VMD-TCN-iTransformer model exhibits a significantly narrower error distribution, with a Residual Standard Deviation (STD) of 0.0028, compared to the baseline VMD-TCN-iTransformer (STD = 0.0145) and TCN-iTransformer (STD = 0.0291). This reduced variance indicates superior robustness against signal fluctuations. Furthermore, a comparative assessment with the field data provided in Section 5.2 confirms that this tight error distribution is consistent across both the validation dataset and the real-world 220 kV transformer deployment, proving that the model's stability does not degrade during the transition from validation to practical engineering scenarios.

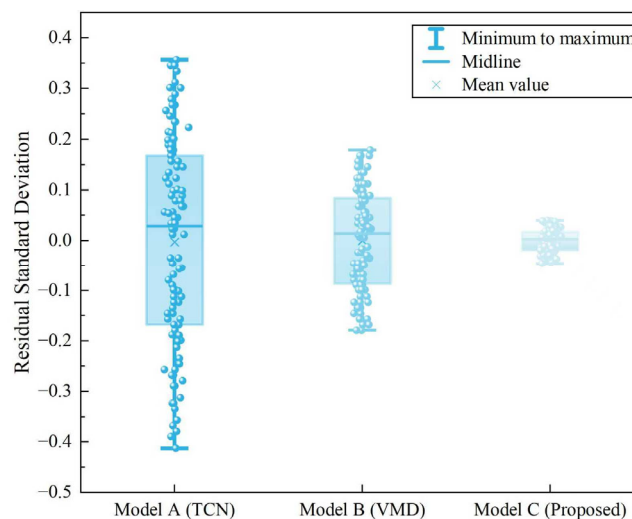


Figure 16. The Distribution of Residual Standard Deviations for the Three Models.

5. Online Monitoring and Predictive Applications Based on Optical Sensors

To further verify the engineering applicability of the developed optical sensor and its monitoring and prediction capabilities, the sensor was deployed on a 220 kV power

transformer operated by Jiangsu Electric Power Company, Taizhou, China. This transformer operates under complex conditions, including long-term high voltage, high temperature, and strong electromagnetic interference, thereby providing a comprehensive evaluation of the operational performance and reliability of the proposed sensing system.

5.1. Settings of Sensor Online Monitoring

In the field application, the multi-parameter sensor probes were introduced into the transformer oil tank through dedicated pressure-resistant sealing flanges and deployed at three representative locations: (1) the top of the oil tank, used to monitor variations in the oil pressure at the tank head; (2) the middle oil region, primarily reflecting oil temperature and pressure changes near the winding area; and (3) the bottom of the oil tank, close to the core and windings, used to acquire liquid column pressure and bottom oil temperature signals. This deployment scheme not only covers different physical state regions within the transformer but also facilitates multi-point comparison and differential analysis. The on-site arrangement of the sensors is shown in Figure 17.

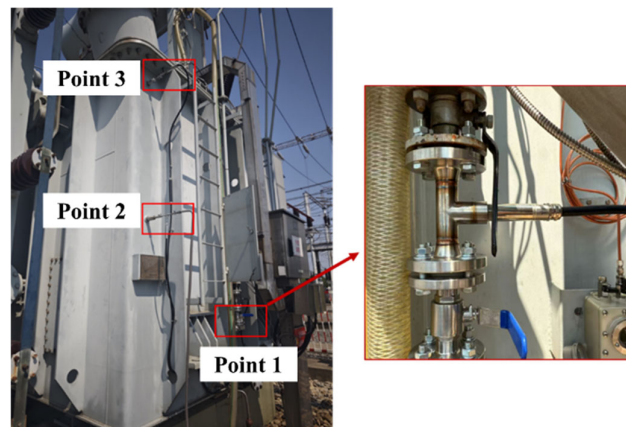


Figure 17. Sensor arrangement for a 220 kV power transformer.

5.2. Results and Analysis of Online Monitoring and Prediction

By continuously acquiring internal transformer temperature and pressure signals using the optical sensors, experiments were conducted in which the main connecting valve of the oil conservator and the breather were actively closed in turn, and the corresponding transformer data were collected. After the main connecting valve of the oil conservator was closed, the pressure equilibrium between the oil conservator and the external oil and gas paths was disrupted. As a result, at measurement points 2 and 3, the pressure increased due to temperature rise. Once the main connecting valve was reopened, the pressure data returned to normal levels. When the breather was blocked, the oil conservator system completely lost its buffering path to the atmosphere, and the transformer effectively became a closed container. Consequently, the pressures at all three measurement points increased with temperature. After the breather was subsequently reopened, the pressures recovered to their original levels. The above results verify the effectiveness of multi-point pressure monitoring in state identification and abnormal condition diagnosis of transformer breathing systems.

On this basis, the prediction algorithm proposed in the preceding sections is further employed to perform rolling prediction analysis on the acquired pressure time-series signals. Specifically, a sliding time-window strategy is adopted, in which the collected historical pressure data are used as known inputs, the model is continuously updated by incorporating newly acquired monitoring data, and the pressure variation trend over a future time horizon is predicted, ultimately generating a continuous predicted pressure

sequence. This approach not only reflects the short-term evolution trend of the oil conservator pressure under current operating conditions, but also captures the developing tendency of pressure anomalies in advance, before alarm thresholds are reached, thereby providing more reliable data support for early warning and operational risk assessment of breathing system abnormalities. Pressure and temperature data were continuously collected for 3.4 days, and predictions were extended to 4.1 days. The predicted pressure data and the corresponding measured data are shown in Figure 17.

To ensure the model adapts to gradual changes in the transformer's operating state (such as oil aging, seasonal ambient variations, or sensor drift) without suffering from computational saturation or performance degradation caused by indefinite data accumulation, a fixed-size sliding time-window strategy is strictly defined. In this framework, the training dataset does not grow infinitely. Instead, an optimal data window (e.g., maintaining the most recent 3 to 5 days of high-resolution historical data) is selected for rolling model updates. As new data enters the window, the oldest data is discarded. This approach effectively mitigates the impact of “concept drift”—where historical data distributions no longer match current conditions—and ensures that the Deep Learning model remains sensitive to the immediate state evolution of the breathing system while keeping the computational load constant.

Based on the above experimental results, it can be observed that abnormalities in transformer breathing systems are often first manifested in changes in the evolution trend of pressure signals. These characteristics are not limited to isolated instantaneous mutations, but rather appear as gradually accumulated pressure deviations and dynamic response mismatches in conjunction with temperature variations. Consequently, relying solely on real-time monitoring data or static threshold-based criteria makes it difficult to timely identify the early stages of abnormalities. Therefore, it is necessary to incorporate pressure prediction information into the fault decision-making process. The specific scheme is as follows:

Real-time transformer pressure signals are acquired and predicted using the above Swarm Intelligence–VMD–TCN–iTransformer model to obtain the predicted oil temperature and oil pressure values. The transformer pressure profile is shown in Figure 18.

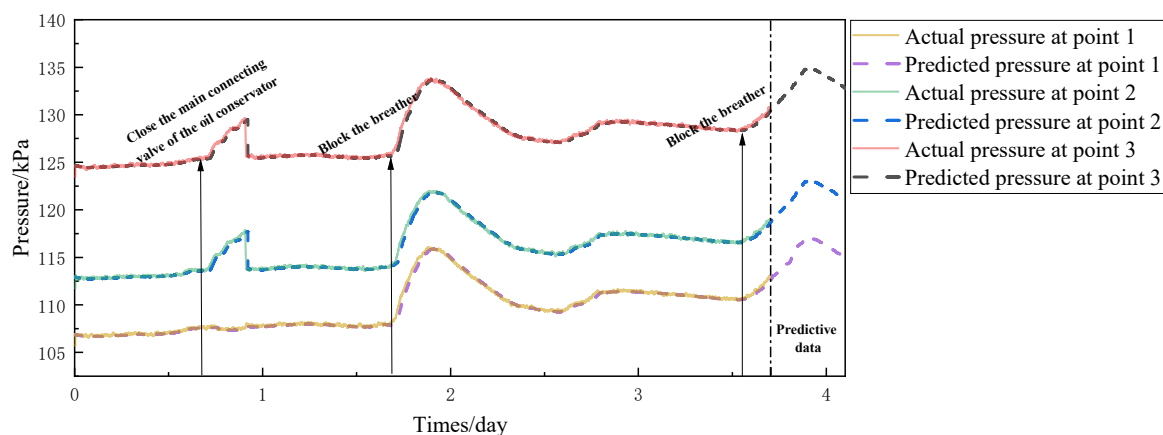


Figure 18. Actual and Predicted Pressure of Transformer Breathing System blockage.

According to the transformer oil temperature calculation formula in Equation (13), the calculated oil temperature can be obtained, as shown in Figure 19. If the calculated oil temperature is significantly higher than the sensor-measured oil temperature, an early warning can be issued. When the breather is blocked, the oil temperature calculated based on pressure inversion is markedly higher than the oil temperature measured by the

sensor. The fundamental reason is that the transformer transitions from an open system to a quasi-closed system, causing the internal pressure to include an additional gas compression component. This component cannot be distinguished by the conventional oil temperature–pressure hydrostatic model, resulting in an overestimation of the calculated oil temperature. This discrepancy effectively reflects abnormal functioning of the breathing system and provides a reliable criterion for early fault identification.

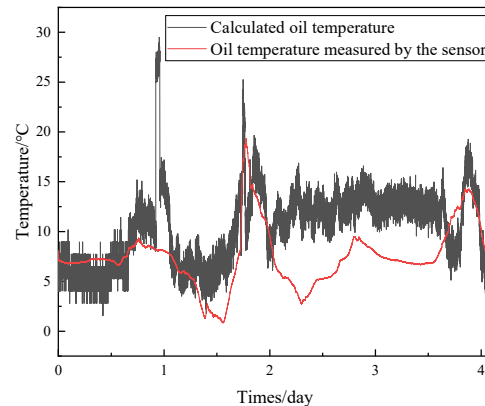


Figure 19. Comparison between calculated and measured temperature.

According to Equation (15), the actual oil level of the transformer is calculated based on the oil temperature. The calculated results are shown in the figure. The oil level is computed using the data from the lower oil tapping point. An increase in the oil level is observed both when the main connecting valve of the oil conservator is closed and when the breather is blocked. The height of the lowest point of the oil conservator can be defined as the normal oil level, while 90% of the maximum oil conservator height can be set as the warning oil level. If the oil level rises significantly beyond this threshold, it can be identified as a fault oil level.

As shown in Figure 20, applying differential processing to the pressure signal from a single sensor can effectively highlight abrupt variations in the pressure evolution process while suppressing the influence of slowly varying temperature on the pressure trend. When the breathing system becomes blocked or gas exchange is restricted, the pressure change rate increases significantly, leading to pronounced non-random spike features in the differential sequence. Therefore, the pressure differential results can serve as an auxiliary criterion for identifying breathing abnormalities under single-point monitoring.

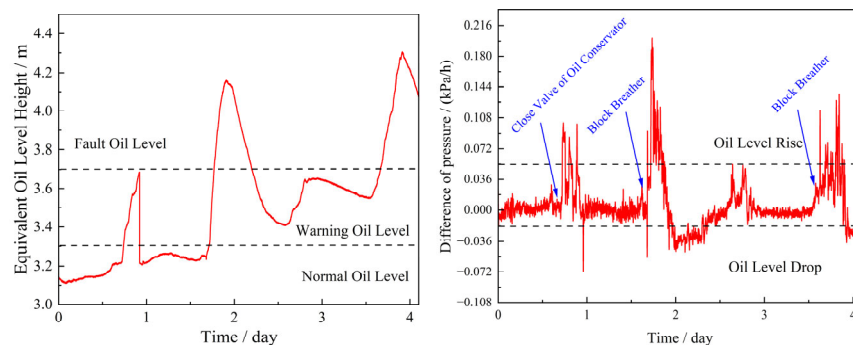


Figure 20. Example of criteria for transformer breathing system blockage.

6. Conclusions

This study focuses on transformer breathing system condition monitoring and early fault identification and conducts a systematic investigation by integrating multi-parameter optical sensing technology with deep-learning-based forecasting methods. The main conclusions are summarized as follows:

(1) A multi-parameter optical sensor suitable for power transformers is proposed and experimentally validated. By integrating FBG temperature sensing with MEMS Fabry–Pérot pressure sensing, a pressure–temperature multi-parameter optical sensing system is developed, enabling stable and accurate measurement.

(2) A swarm intelligence-VMD–TCN–iTransformer pressure–temperature forecasting model is established. To address the strong non-stationarity and complex spectral characteristics of transformer oil temperature and oil pressure signals, a swarm intelligence algorithm is introduced to adaptively optimize the key parameters of VMD, thereby significantly improving the quality of modal decomposition. On this basis, the local temporal modeling capability of the TCN is combined with the global feature-correlation modeling advantage of the iTransformer, enabling high-precision prediction of multi-scale temporal features. Comparative experimental results indicate that the proposed model outperforms non-decomposed or non-optimized benchmark models in terms of both prediction accuracy and dynamic response capability.

(3) Prediction-based abnormality criteria for transformer breathing systems are proposed. By incorporating the pressure–equilibrium mechanism of transformer breathing systems, predicted pressure data are utilized to inversely estimate oil temperature and equivalent oil level variations. This approach reveals the pressure–temperature response mismatch that occurs when the breathing system transitions from an open system to a quasi-closed system, providing an effective basis for identifying abnormal operating conditions.

Author Contributions: J.L. and J.S. contributed to the conceptualization, methodology, and drafting of the original manuscript. P.W. took responsibility for formal analysis, supervision, and overall project coordination. Q.L. was involved in data curation and supported funding acquisition. Y.L. contributed to visualization and assisted in project administration. Z.L. and Y.W. participated in the review and editing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of State Grid Jiangsu Electric Power Co., Ltd., grant number J2024217.

Data Availability Statement: Data is contained within the article. Further inquiries can be directed to the corresponding author since the large dataset is derived from State Grid Company and should not be publicly open.

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

References

1. Wang, M.V.A.J.; Vandermaar, A.J.; Srivastava, K.D. Review of condition assessment of power transformers in service. *IEEE Electr. Insul. Mag.* **2003**, *18*, 12–25.
2. Li, L.; Niu, S.; Ho, S.L.; Fu, W.N.; Li, Y. A novel approach to investigate the hot-spot temperature rise in power transformers. *IEEE Trans. Magn.* **2015**, *51*, 1–4.
3. Susa, D.; Lehtonen, M.; Nordman, H. Dynamic thermal modelling of power transformers. *IEEE Trans. Power Deliv.* **2005**, *20*, 197–204.
4. Meira, M.; Ruschetti, C.R.; Álvarez, R.E.; Verucchi, C.J. Power transformers monitoring based on electrical measurements: State of the art. *IET Gener. Transm. Distrib.* **2018**, *12*, 2805–2815.

5. Saha, T.K. Review of modern diagnostic techniques for assessing insulation condition in aged transformers. *IEEE Trans. Dielectr. Electr. Insul.* **2003**, *10*, 903–917.
6. Torres, J.F.; Hadjout, D.; Sebaa, A.; Martínez-Álvarez, F.; Troncoso, A. Deep learning for time series forecasting: A survey. *Big Data* **2021**, *9*, 3–21.
7. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies* **2018**, *11*, 1636.
8. Yang, F.; Wu, T.; Jiang, H.; Jiang, J.; Hao, H.; Zhang, L. A new method for transformer hot-spot temperature prediction based on dynamic mode decomposition. *Case Stud. Therm. Eng.* **2022**, *37*, 102268.
9. Behniafar, A.; Farshad, M. A machine-learning-based strategy for online prediction of hotspot temperature in dry-type three-phase transformers. *Comput. Electr. Eng.* **2025**, *126*, 110490.
10. Hughes, R.; Haidinger, T.; Pei, X.; Vagg, C. Real-time temperature prediction of electric machines using machine learning with physically informed features. *Energy AI* **2023**, *14*, 100288.
11. Doolgindachbaporn, A.; Callender, G.; Lewin, P.L.; Simonson, E.; Wilson, G. Data driven transformer thermal model for condition monitoring. *IEEE Trans. Power Deliv.* **2021**, *37*, 3133–3141.
12. da Costa Souza, J.P.; Picher, P.; Zinflou, A.; Fofana, I.; Asl, M.B. A comprehensive review on artificial intelligence-based applications for transformer thermal modeling: Background and perspectives. *IEEE Access* **2024**, *12*, 152310–152329.
13. Zhou 2024, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 11106–11115.
14. Wu, H.; Hu, T.; Liu, Y.; Zhou, H.; Wang, J.; Long, M. TimesNet: Temporal 2D-variation modeling for general time series analysis. *arXiv* **2022**, arXiv:2210.02186.
15. Su 2022, H.; Zhao, D.; Heidari, A. RIME: A physics-based optimization. *Neurocomputing* **2023**, *532*, 183–214.
16. Liu, Y.; Hu, T.; Zhang, H.; Wu, H.; Wang, S.; Ma, L.; Long, M. iTransformer: Inverted transformers are effective for time series forecasting. *arXiv* **2023**, arXiv:2310.06625.
17. He 2023, J.; Yang, Q.; Wang, Z. On-line fault diagnosis and fault-tolerant operation of modular multilevel converters—A comprehensive review. *CES Trans. Electr. Mach. Syst.* **2020**, *4*, 360–372.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.