HYDROGEN ENERGY STORAGE SAFETY EARLY WARNING MODEL BASED ON MULTI-SOURCE SENSING

Ning Feng, Bo Yao, Xuanmin Huo, Wei Zhang, Hua Dang, Yujie Jia, Hao Su*
State Grid Xinxiang Electric Power Supply Company, Xinxiang 453700, China

Abstract - In the specific field of hydrogen storage safety early warning, there are challenges such as difficulties in sensor fusion and the ease with which leakage signals can be masked by noise. Therefore, we propose a graph-augmented time-series model that integrates multiple sensing methods for early warning of hydrogen storage safety. Our model is an MP-GETM model, which integrates four steps: sensing, graph augmentation, inference, and early warning. It utilizes data from multiple sensors, including hydrogen concentration, infrared thermal imaging, and acoustic emission, combined with wavelet transform and variational autoencoders to enhance weak signals, thereby relying on the graph-augmented time-series autoencoder for state reconstruction. Experimental results show that in a 322-hour real-world test, our model improves upon the optimal baseline CAD by 40.3% and 6.6 seconds. Similarly, under threat scenarios, our model achieves a high MTW of 35.2 milliseconds. The experimental results also demonstrate that our model maintains a high F1 score of 86.7% even with a very low signal-to-noise ratio, indicating strong robustness. The interpretability contribution reaches 0.29, indicating that the model's decision-making is interpretable and transparent.

Keywords: Hydrogen energy storage, Safety early warning, Multi-source sensing fusion, Graph neural network, Weak signal enhancement, Dynamic threshold, Explainable AI.

1. Introduction

As low-carbon and environmental protection become important goals of social development, the energy structure tends to be cleaner. Hydrogen energy is a high-density, low-carbon, and healthy energy carrier, which is particularly important under the dual carbon goal. The realization of hydrogen energy storage technology depends on the large-scale and long-term storage of hydrogen energy, which can serve as a backup energy source for intermittent and fluctuating renewable energy [1]. However, hydrogen itself has flammable, explosive, and leak-prone physical and chemical properties, which makes its storage extremely difficult. The explosion limit is extremely low, the ignition energy is extremely low, the diffusion rate is high, and the gas is colorless and odorless. This makes hydrogen storage an event with safety risks throughout its entire life cycle. Once hydrogen storage leaks or hydrogen explosions occur, it will cause significant casualties and property losses [2].

Therefore, in this context, developing a safe and efficient hydrogen storage leak detection instrument has become an important tool for promoting hydrogen storage and energy development. However, existing safety detection equipment related to hydrogen storage mainly relies on threshold-based trigger alarms. Although such alarm devices are simple in structure, low in cost, and have a certain degree of reliability, they still face the problem of not being able to fully depict the dynamic evolution process of hydrogen diffusion and propagation. Moreover, this threshold-based judgment mechanism is prone to causing the problem of insufficient early warning performance after risk identification [3]. In actual environments, various sensors are extremely susceptible to electromagnetic interference, generating additional interference signals and affecting the early warning function of hydrogen storage [4].

In recent years, with the rapid development of emerging technologies such as the Internet of Things and big data, more and more studies in the field of early warning have tried to use multi-modal heterogeneous fusion data for early warning detection. At present, by fusing data from multiple modalities such as gas concentration, infrared thermal imaging, acoustic emission, and video surveillance, and on this basis, realizing

dx.doi.org/10.17683/ijomam/issue24.26

comprehensive monitoring and early warning of hydrogen storage, it has become a new development trend of hydrogen storage. Deep learning models, such as graph neural networks, can model data from different sources according to time nodes and spatial nodes, and can mine hidden patterns from massive real-time data [5, 6]. However, this model has not been applied to hydrogen storage early warning. Therefore, this paper combines multi-modal heterogeneous data to build a hydrogen storage early warning system with sensitivity automation and strong robustness, which can intelligently identify weak signals and improve the reliability of hydrogen storage identification by mixing multiple modal information, becoming a favorable technical tool in the hydrogen storage process [7].

2. Literature Review

2.1 Safety Risk Characteristics of Hydrogen Energy Storage Systems

Hydrogen storage benefits from its physicochemical properties, which leads to safety risks during storage. Therefore, timely early warning and detection of hydrogen storage leaks are very important [8]. High-pressure storage tanks and other storage methods are prone to sealing failure or material fatigue, which can cause hydrogen leaks. Liquid storage systems may cause overpressure risks due to rapid vaporization of hydrogen caused by insulation failure. Therefore, a lot of research is on hydrogen storage leak detection technology. Early studies tried to use electrochemical or catalyst-based methods to detect and alarm based on hydrogen sensors based on threshold concentration [9]. Lv et al. [10] used 70 MPa hydrogen storage screen inlet and outlet and set up sensors to achieve rapid response to leaks. However, the coverage of such methods is very limited and the robustness is not strong under different environmental conditions. In recent years, some non-contact technologies have also received widespread attention. Shi et al. [11] applied infrared thermal imaging to the detection of temperature anomalies on the surface of hydrogen storage to detect the leak point of hydrogen storage. Xiao et al. [12] used acoustic emission technology to intelligently capture the high-frequency signal generated by hydrogen cracks in the pottery jar, thereby realizing the early diagnosis of cracking of hydrogen storage materials.

2.2 Multi-source Sensing Fusion and Anomaly Detection Methods

In terms of anomaly detection, by integrating multi-source heterogeneous data, the robustness of anomaly detection can be improved, and the outliers generated by single-source data can be avoided, which may cause missed detection or false detection. Multi-source heterogeneous data is more reliable and accurate than single data source [13]. Currently, there are three methods for integrating multiple data sources in the industrial field: integrating datasets, feature sets and decision sets. For dataset integration, direct integration requires that various types of data be highly synchronized in time and space [14]. This method has extremely high data requirements. Feature clustering integration can be associated after extracting various key features, which has low data requirements and can maintain high accuracy. For example, in the diagnostic research of Qiu et al. [15], in the detection of leaks in natural gas pipelines, pressure sound waves and vibration signals were integrated, and the features were reduced in dimensionality by combining principal component analysis, and finally input into the vector set for binary classification detection. In addition, some traditional methods still maintain good performance in anomaly detection, such as statistical process control, but in the specific field of light storage, due to the diversity of hydrogen storage scenarios and the diversity of interference signal sources [16], their use is very limited. At present, most research still focuses on leakage detection in general industrial fields, while for the specific field of hydrogen, multi-source fusion mechanisms [17], such as how to coordinate the concentration of gases around hydrogen, the thermal field and sound field generated by hydrogen, and multi-heterogeneous spatiotemporal data, there is still a lack of specific framework models [18].

2.3 Intelligent Security Early Warning Model

At present, with the rapid development of information technology and artificial intelligence, hydrogen-based safety early warning is transitioning from the traditional threshold-based method to the intelligent model-based prediction framework. The integration of data-driven, online learning and adaptive update mechanisms to build an intelligent model combining hydrogen multi-heterogeneous data has become a new research direction. For example, in the study, Diao et al. [19] applied computational fluid dynamics to the detection of hydrogen diffusion data and fused the real-time data of the sensors to train a convolutional neural network to predict the range of hydrogen diffusion. Finally, the features extracted by the convolutional neural network were used for classification and regression prediction. In the study, Xu et al. [20] applied federated learning to the distributed early warning process. Under the premise of protecting the privacy of multiple different hydrogen refueling stations, the parameters of the model were jointly optimized to improve the early warning capability of the model. However, there are still huge difficulties in this field, specifically the scarcity of data. Labeling

data requires a lot of manpower, material resources and financial resources. On the other hand, there is a huge semantic gap in multi-heterogeneous data, which usually cannot be changed by deep learning models [21].

3. Innovative Hydrogen Energy Storage Safety Early Warning Model and Method

3.1 Overall Model Architecture

Figure 1, we propose a safety early warning model for hydrogen storage that integrates multi-modal heterogeneous data and combines graph time-series modeling. This model employs heterogeneous data from multiple modalities at the perception layer, including time-series data from hydrogen concentration sensors, image data from infrared thermal imaging, audio data from acoustic emission propagation sensors, and time-series data from pressure transmitters. Temperature probes and cameras transmit time-series and image data, respectively. Therefore, we need to deploy these sensors at various points along the hydrogen tank and pipeline at the perception layer.

At the enhancement layer, we add a module for enhancing weak anomalous signals. The core of this module is the fusion of wavelet decomposition and variational autoencoder to improve the signal-to-noise ratio of signals generated by early leaks. At the inference layer, we fuse the physical topology of the device and data associations based on a hybrid dependency graph. The resulting time-enhanced autoencoder is used for data reconstruction and sensitive anomaly detection. At the prediction layer, we combine uncertainty quantification techniques and interpretability analysis to determine the safety level of hydrogen storage [22, 23].

The design philosophy of our model is that a single set of data cannot comprehensively characterize the safety of hydrogen storage. Secondly, traditional threshold-based methods often miss a large number of leaks due to insufficient sensitivity to early warning signals. Effective safety warnings not only require high accuracy but also interpretability and robust engineering practicality. Therefore, our model employs an unsupervised or semi-supervised approach to adapt to the lack of labeled data in the real world.

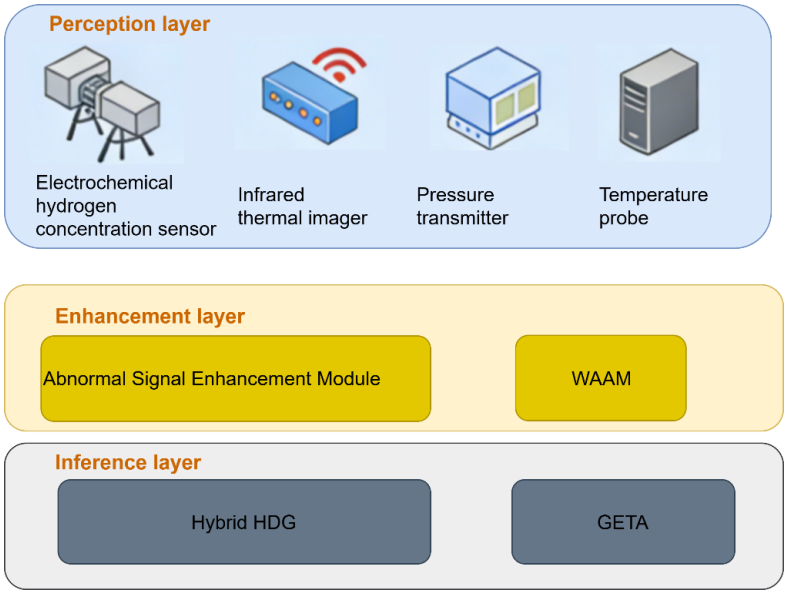

Figure 1: Model Framework

3.2Data and Enhancement of Weak Signals

After obtaining diverse and heterogeneous data, we need to preprocess this data. The core issue is to address the inconsistency in sampling frequency, different dimensions, and noise interference during the data processing. Therefore, we designed a detailed preprocessing procedure. In the early stages of a hydrogen leak, the data variation is very low. This weak signal is extremely easy to be masked by noise emitted by various sensors in the environment. Therefore, we first need to perform a standard preprocessing procedure on data from different sources. Then, we pass these signals through a weak anomaly signal enhancement module to maximize the feasibility of early warning for hydrogen leaks. Assuming there are a total of n different types of sensors in our system, we assume that the raw observation of each type of sensor i at a certain time t can be represented as $\mathbf{x}_i^{(t)} \in \square^{d_i}$, where the dimension of this feature can be represented as d_i . Since the forms of different features are different, including images, text data, etc., and the sampling frequencies of different data are also different, we first need to unify the sampling frequency. For high-

frequency signals, we use a sliding sampling method, as shown in Formula (1) [24, 25].

$$\bar{\mathbf{x}}_i^{(t)} = \frac{1}{k}\sum_{\tau=t-k+1}^{t}\mathbf{x}_i^{(\tau)}, \quad k = \left\lfloor \frac{f_i}{f_0} \right\rfloor \tag{1}$$

After standardizing the data sampling frequency, it is necessary to process the steel mill differences in the data. We adopt the z-score standardization process, as shown in Formula (2).

$$\mathbf{x}_i^{(t)} = \frac{\bar{\mathbf{x}}_i^{(t)} - \mu_i}{\sigma_i} \tag{2}$$

In μ_i Formula (2). The two parameters, σ_i mean and standard deviation, represent the statistical values obtained from the training set data, respectively.

For each dimension of the time series data, we perform a small-step decomposition process with four levels. This process yields several different sub-band coefficients, which can be filtered using energy entropy, and an enhancement factor is obtained, as shown in Formula (3).

$$\tilde{s}_j(t) = \sum_{k\in \mathrm{K}} \alpha_k \cdot \mathrm{IDWT}(c_j^{(k)}) \tag{3}$$

In Equation (3), IDWT is used as a value for the inverse wavelet transform, $c_j^{(k)}$ representing the coefficient of the corresponding k-th subband in the formula. To further amplify the hidden signal of hydrogen leakage, we added a lightweight variational word encoder. This encoder is used to further process the signal to enhance the hydrogen leakage signal, as shown in Equation (4).

$$\begin{aligned} &\text{Encoder:}\boldsymbol{\mu} = \mathbf{W}_\mu \tilde{\mathbf{s}} + \mathbf{b}_\mu, \quad \boldsymbol{\sigma} = \exp\left(\frac{1}{2}(\mathbf{W}_\sigma \tilde{\mathbf{s}} + \mathbf{b}_\sigma)\right) \\ &\text{Latent:} \quad \mathbf{h} = \boldsymbol{\mu} + \boldsymbol{\sigma} \square \grave{\mathbf{o}}, \quad \grave{\mathbf{o}} \sim \mathrm{N}(0, \mathbf{I}) \\ &\text{Decoder:} \quad \hat{\mathbf{s}} = \mathbf{W}_d \mathbf{h} + \mathbf{b}_d n \end{aligned} \tag{4}$$

In Equation (4), the variational autoencoder can be trained using KL loss, as shown in Equation (5).

$$\mathrm{L}_{\mathrm{VAE}} = \tilde{\mathbf{s}} - \hat{\mathbf{s}}\ {}_2^2 + \lambda D_{\mathrm{KL}}\left(\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)\ \mathrm{N}(0, \mathbf{I})\right) \tag{5}$$

In $\lambda = 0.1$ Formula (5), the output $\hat{\mathbf{s}}$ serves as the input to the next module. This processing improves the ability to identify weak abnormal signals and increases the detection rate of hydrogen leaks.

3.3 Graph Augmented Temporal Autoencoder (GETA) Modeling

The graph-augmented temporal autoencoder (GETA) model is shown in Figure 2. During hydrogen storage, there are certain spatial and temporal dependencies between different monitoring points. Therefore, we combine a graph-augmented temporal encoder to perform dependency graph mixing. We assume that there are M different hydrogen monitoring nodes in the system, which are used to monitor different locations of the hydrogen storage tank, including the top, bottom, inlet, and outlet. In this way, each node will generate a corresponding set of features. We assume that the graph can be represented by Equation (6).

$$\mathrm{G} = (\mathrm{V}, \mathrm{E}) \tag{6}$$

In Equation (6), the node set can be represented as $\mathrm{V} = \{v_1, \ldots, v_M\}$. The set of edges can be represented as E consisting of two types of edges. The first type is the physical connection edges, meaning that the two nodes are interconnected on a device or pipe. This can be represented as $A_{ij}^{\mathrm{phys}} = 1$. Another type of edge is the data-related edge, which refers to edges where node signals are interconnected at different stages of training. Transfer information is used to measure specific input formula (7).

$$I(x_i; x_j) = \sum_{x_i, x_j} p(x_i, x_j) \log \frac{p(x_i, x_j)}{p(x_i)p(x_j)} \tag{7}$$

In Formula (7), if the mutual information is greater than the threshold, which is usually 0.3, then this side is set to 1.

Therefore, we can obtain the final adjacency matrix, as shown in Equation (8).

$$\mathbf{A} = \alpha \mathbf{A}^{\mathrm{phys}} + (1-\alpha)\mathbf{A}^{\mathrm{data}}, \quad \alpha = 0.6 \tag{8}$$

We also designed a GETA encoder, which incorporates temporal modeling on top of graph convolution. We assume the input node feature matrix is represented as $\mathbf{H}^{(0)} \in \square^{M \times d_h}$. Therefore, the update of the graph convolution in the l-th layer can be expressed as Equation (9).

$$\mathbf{H}^{(l+1)} = \sigma\left(\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right) \tag{9}$$

In $\mathbf{A} = \mathbf{A} + \mathbf{I}$ Equation (9), $\mathbf{W}^{(l)} \in \square^{d_h \times d_h}$ the weight matrix composed of the degrees of each node represents the parameter vector that can be learned in the model, and $\sigma(\cdot)$ represents the activation function. After obtaining such graph-reinforced sequence data, we further input it into a bidirectional long short-term neural network, as shown in Equation (10).

$$\begin{aligned} &\vec{\mathbf{o}}_t = \mathrm{LSTM}_{\mathrm{fwd}}(\mathbf{H}_t), \quad \overleftarrow{\mathbf{o}}_t = \\ &\mathrm{LSTM}_{\mathrm{bwd}}(\mathbf{H}_t), \quad \mathbf{o}_t = [\vec{\mathbf{o}}_t; \overleftarrow{\mathbf{o}}_t] \end{aligned} \tag{10}$$

In the decoder, we will further reconstruct the model's input, as shown in Equation (11).

$$\mathbf{H}_t = \mathrm{MLP}(\mathbf{o}_t) \tag{11}$$

Therefore, in order to ensure the accuracy of state reconstruction and the sensitivity of anomaly detection, we use a two-branch training function, as shown in Equation (12).

$$\begin{aligned} \mathrm{L}_{\text{total}} &= \underbrace{\beta \frac{1}{MT} \sum_{t,i} \mathbf{H}_{i,t} - \mathbf{H}_{i,t}\ {}_2^2}_{\mathrm{L}_{\text{rec}}} + \\ &(1-\beta)\underbrace{\mathrm{E}[\log D(\mathbf{H}_{\text{anom}})] + \mathrm{E}[\log(1 - D(\mathbf{H}_{\text{anom}}))]}_{\mathrm{L}_{\text{adv}}} \end{aligned} \tag{12}$$

In $\beta = 0.7$ Formula (12) D, as the discriminator in the model weights, is used after training to express the definite fraction of the existentialist field as Equation (13).

$$S_t = \frac{1}{M} \sum_{i=1}^{M} \mathbf{H}_{i,t} - \mathbf{H}_{i,t}\ {}_2^2 \tag{13}$$

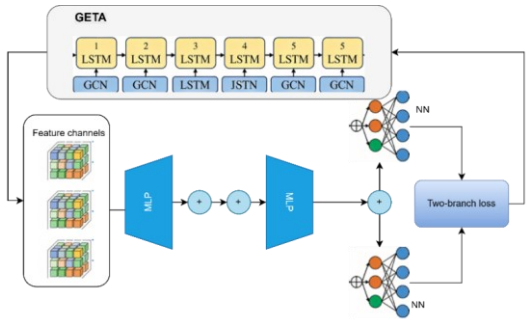

Figure 2: Graph-Enhanced Temporal Autoencoder (GETA) Modeling

3.4 Dynamic Risk Assessment

As shown in Figure 3, the ultimate goal of the early warning model is not simply to identify whether storage clearing has occurred, but to provide a more reliable basis for risk decision-making. Therefore, we have incorporated an early warning mechanism that integrates uncertainty quantification.

Specifically, we first use Monte Carlo methods to estimate uncertainty. Therefore, we retain the dropout layer during the inference phase, performing 20 forward propagations to obtain a set of anomaly scores, specifically representing... $\{S_t^{(k)}\}_{k=1}^{K}$ For outlier scores, we calculate the variance and mean, as shown in Formula (14).

$$\begin{aligned} \bar{S}_t &= \frac{1}{K} \sum_{k=1}^{K} S_t^{(k)}, \\ \mathrm{Var}(S_t) &= \frac{1}{K} \sum_{k=1}^{K} (S_t^{(k)} - \bar{S}_t)^2 \end{aligned} \tag{14}$$

In Equation (14), a higher variance indicates that the current state is uncertain and may be subject to some unseen minor leakage patterns. Therefore, it is necessary to improve the detection capabilities.

In addition, we have added an adjustment to the dynamic East China window azimuth in the model. Assuming the window length w is set to 3,600, which is one hour, we will calculate the two azimuths of 95% and 99% of the azimuth, respectively, as shown in Formula (15).

$$\begin{aligned} \theta_t^{\text{med}} &= Q_{0.95}(\{S_\tau\}_{\tau=t-W}^{t-1}), \\ \theta_t^{\text{high}} &= Q_{0.99}(\{S_\tau\}_{\tau=t-W}^{t-1}) \end{aligned} \tag{15}$$

From Formula (15), we can obtain the following early warning mechanism: low risk can be defined as $\bar{S}_t < \theta_t^{\text{med}}$ if simultaneously satisfying... $(\theta_t^{\text{med}} \le \bar{S}_t < \theta_t^{\text{high}})(\mathrm{Var}(S_t) < \delta$ $\delta = 0.05$ It is then defined as medium risk. If it meets the following conditions... $\bar{S}_t \ge \theta_t^{\text{high}}$ And $\mathrm{Var}(S_t) < \delta$ it can then be further defined as high risk. If $\bar{S}_t \ge \theta_t^{\text{high}}$ Furthermore, $\mathrm{Var}(S_t) \ge \delta$ it is indicated as an emergency risk, often representing the encounter with some unseen anomalies.

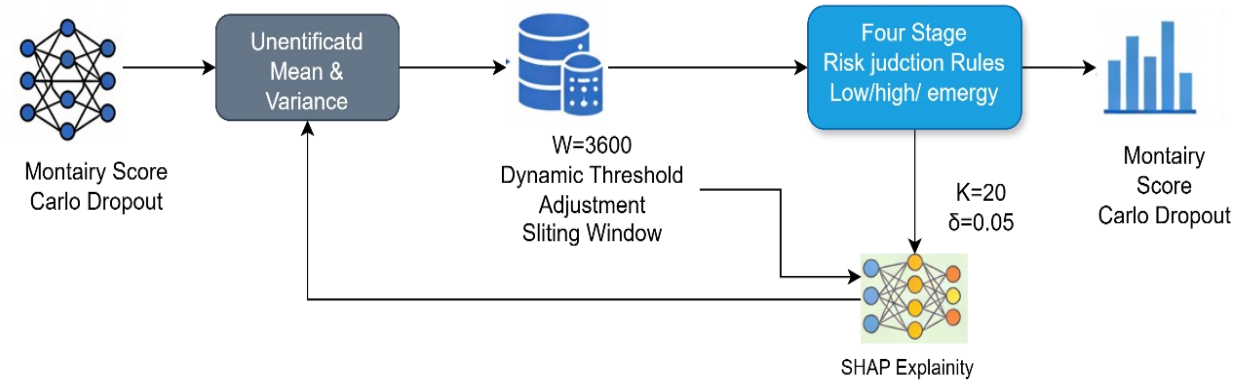

Figure 3: Dynamic Risk Assessment

4. Experimental Evaluation

4.1 Experimental Design

To comprehensively analyze the effectiveness of the proposed MP-GETM in hydrogen energy storage safety early warning, we conducted experimental investigations in a simulated environment. We used a national-level hydrogen energy storage safety experimental platform and simulated a 70 MPa high-pressure pump gaseous storage system. We employed six types of sensors: hydrogen concentration sensor, infrared sensor, acoustic emission sensor, pressure sensor, temperature sensor, and visible light sensor. The sampling frequency for all sensors was uniformly defined to 10 Hz. Therefore, we obtained a total of 280 hours of real-time data and manually labeled data, ultimately obtaining a total of 322 hours of time-series data.

In the model, the evaluation metrics mainly include common classification evaluation metrics, such as accuracy, precision, and recall. In addition, we defined some key metrics for early warning, such as the average early warning time. This metric is defined as the time difference between the model's first encounter with a high-risk warning and the gold standard of human annotation. We selected five mainstream models for comparison. First, we adopted an unsupervised anomaly detection method combining ensemble learning, namely Isolation Forest (IF). LSTM-AE is a temporal encoder that combines LSTM and attention mechanisms.

MTAD-GAT represents a multi-variable fusion-based anomaly detection graph attention network; USAD represents an unsupervised anomaly detection method based on adversarial training; and TranAD is an emerging anomaly detection model, with its core based on Transformer. We perform the same data processing operations on all mechanical models and test them in the same experimental environment, aiming to analyze the robustness and reliability of the models through comprehensive testing.

4.2 Experimental Results

Figure 4 shows the distribution of the model constructed in this paper in two-dimensional space after signal enhancement and graph enhancement time encoder processing.

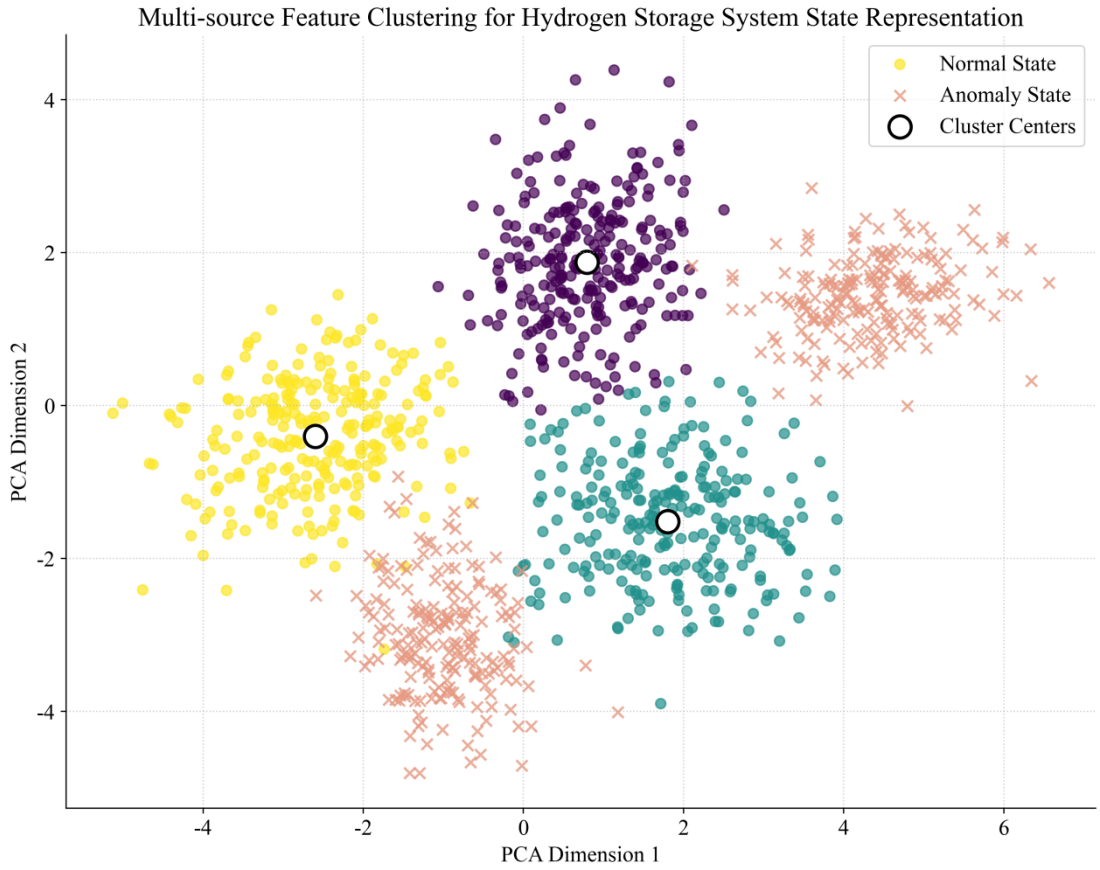

Figure 4: Visualization of state clustering of hydrogen energy storage system based on multi-source feature fusion

The figure shows that the two directions represent the principal components, corresponding to dimension one and dimension two of the principal components. The figure clearly shows five clusters of data points. Normal samples are mainly concentrated in the directional region of the figure, indicating that the model can effectively capture the common features of the system in a stable state.

Different types of anomalous samples are distributed in the other clusters. There are obvious coarse boundaries between micro-leakage and local overheating, indicating that the features extracted by the model, although clustered around the central point, can still have enhanced discriminative ability to distinguish different types of anomalous samples.

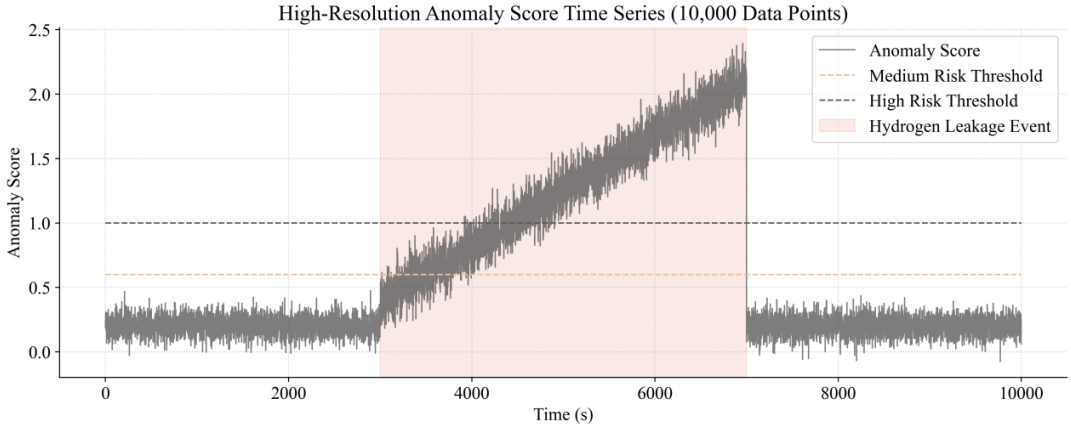

Figure 5: Correspondence between high-resolution anomaly scoring time-series curves and leakage events

Figure 5 clearly illustrates the dynamic detection process of the proposed MP-GETM model in a hydrogen leak event. The figure covers a total of 1000 seconds and 10,000 data points. The horizontal axis represents time, and the vertical axis represents the model's anomaly score, ranging from 0 to 2.5. The solid line in the figure represents the score curve where anomalies are evenly distributed, while the gray area represents the score fluctuation range under normal, leak-free conditions.

The figure also marks the threshold ranges for medium and high risk, with the red vertical line representing the initial moment of the leak. As can be seen from the figure, before the leak occurs, i.e., before t is less than 350 seconds, the score is at a low level. After the leak occurs, the curve continuously increases from 350 to 600 seconds, eventually exceeding the high-risk threshold of 1.8, triggering the early warning mechanism.

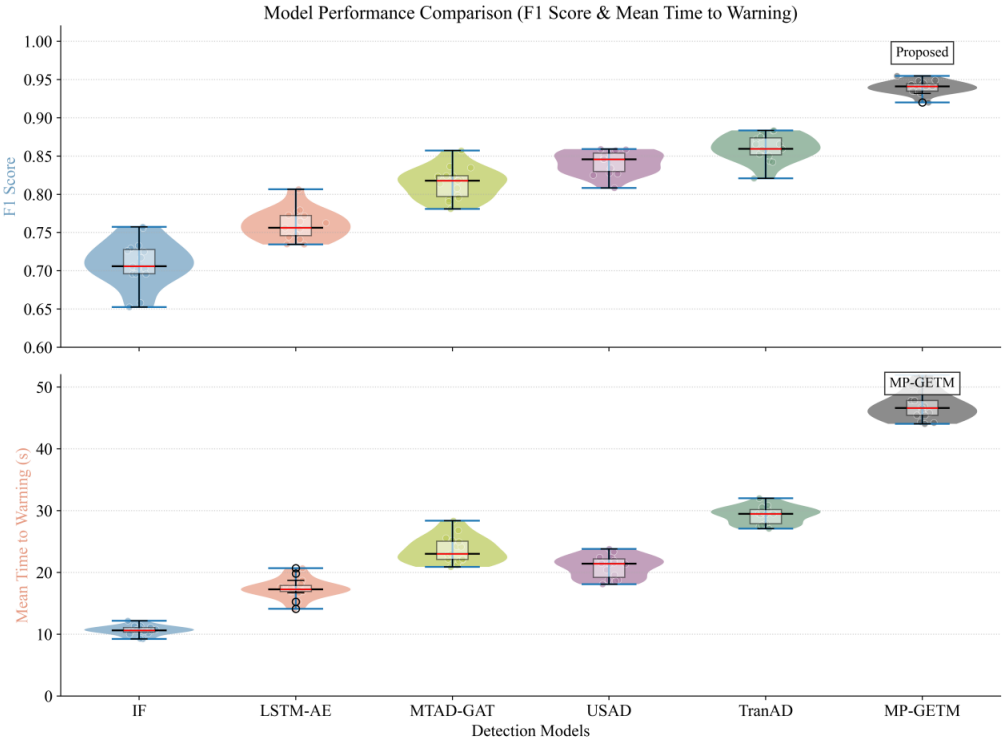

Figure 6: Performance comparison of multiple models

Figure 6 shows the performance comparison of our proposed mpjetm model with several other baseline methods in a hydrogen leak early warning task. We use a box plot to illustrate the distribution

of the results. The upper part of the box plot represents the distribution of F1 scores for different models, and the lower part represents the distribution of average warning time.

The results show that our proposed model achieves an F1 score of 0.92, compared to the highest of 0.877 achieved by other models. The F1 score distribution of our proposed model is more concentrated, indicating that the model has stronger prediction accuracy and more stable predictions.

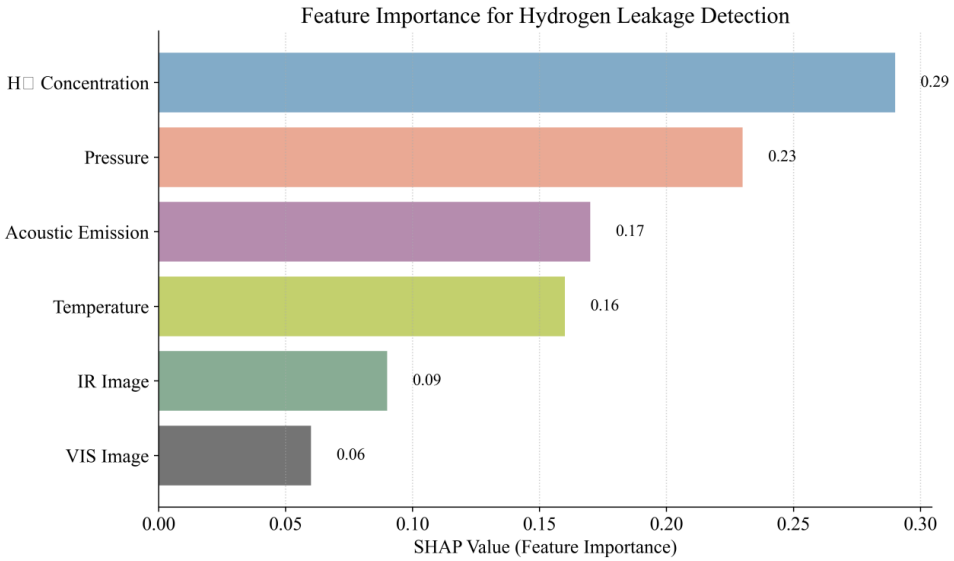

Figure 7: Ranking of importance of various sensing features in hydrogen leak detection

Figure 7 analyzes the contribution mechanism of different modal data to the overall early warning effect of the proposed MEDETM model in the detection of hydrogen leaks. We use the sciap method for analysis. Quantitative analysis shows that the contribution of different modal data to the overall effect varies. The hydrogen sensor contributes the most, reaching 0.29, which is the most direct basis for judging hydrogen leaks. The pressure sensor is also a core component, with a SHAP value of 0.23, indicating that the system will experience pressure changes internally. In the early stages of micro-leaks, detecting depressurization can provide an early warning, ranking third in acoustic emission signals. Temperature, as a secondary indicator, contributes only 0.16, while infrared imaging and visible light images, although rich in information, contribute little in this experimental scenario.

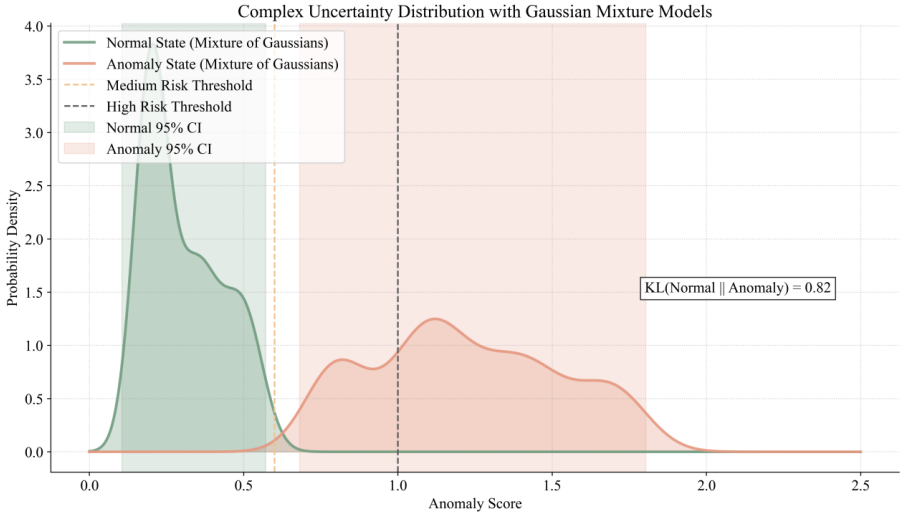

Figure 8: Ranking of importance of various sensing features in hydrogen leak detection

Figure 8 illustrates the impact mechanism of different sensing features on the overall early warning process of the proposed mpgetm model in a hydrogen-based leakage detection task. As can be seen from the figure, the number of components representing normal and abnormal states differs significantly. Normal states are mainly concentrated around 0 to 0.5, while abnormal states are mainly

distributed in the range of 0.5 to 2. This demonstrates that the model constructed in this paper has a significant advantage in identifying normal and routine states.

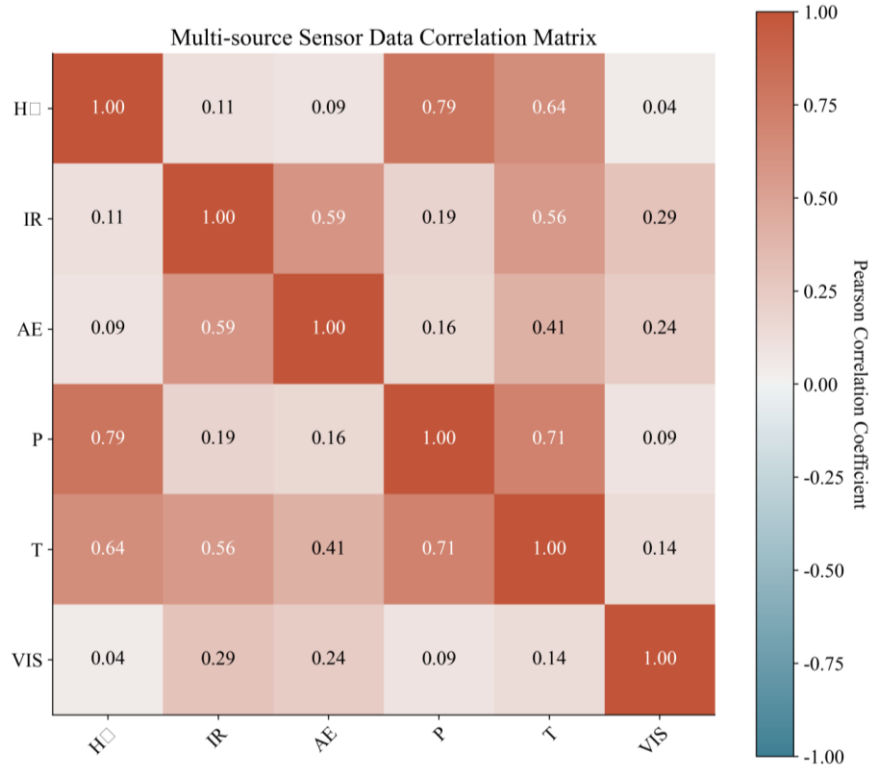

Figure 9: Heatmap of correlation between multi-source sensor data

Figure 9 uses a heatmap to illustrate the correlation analysis of different sensor modalities. The figure shows six sensing modalities. A strong correlation of 7-8% exists between hydrogen concentration and pressure, indicating that a decrease in system pressure often accompanies leaks. Similarly, a strong correlation of 0.64% exists between hydrogen concentration and temperature. Hydrogen diffusion may cause a rise in the temperature of hydrogen storage equipment or lead to heat exchange with the environment. The correlation between concentration and temperature is only 2-1%, indicating a weak correlation. This suggests that thermal imaging has a lag in its response to micro-leakage events.

As shown in Table 1, we comprehensively evaluated the detection performance and anomaly detection results under different types of fault conditions. The model demonstrates that in the micro-leakage scenario, the model achieved an F1 score of 92% and a median time over time (MTW) of 35.2 seconds. Compared to the other two fault types, the prediction accuracy and MTW values are higher. This indicates that the model performs best in predicting anomalies of specific diffusion types. The accuracy and recall rates are both high for different fault types, suggesting that the model adheres to a conservative strategy of minimizing false negatives. These results demonstrate high rationality and reliability in industrial-grade applications.

Table 1. Detection performance under different types of fault conditions

Fault type	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 (%)**	**MTW(s)**
Microleakage	94.6	91.5	88.9	90.2	35.2
Local overheating	93.8	90.3	89.1	89.7	26.8
Material deterioration	92.1	88.6	86.2	87.4	22.1

Table 2 shows the contribution of different modules to overall performance through system project analysis. The figure shows that the complete model achieves an F1 score of 92%, while the MTW (Mean Time Tolerance) is only 0.887 seconds. After removing the WAA (Waa Subsystem) module, the F1 score drops to 86.3%, and the MTM (Mean Time Tolerance) also decreases.

This demonstrates that weak signal enhancement plays a crucial role in early warning systems, helping them to more sensitively identify leakage events.

Table 2. Contribution of different modules to overall performance

Model variants	F1 (%)	MTW(s)	ΔF1 vs Full
Full Model	92.0	28.7	—
–WAAM	86.3	19.4	-5.7
–Data Edge in HDG	90.2	25.1	–1.8
–Phys Edge in HDG	90.9	26.3	–1.1
–Adversarial Loss	87.8	23.5	–4.2
–Time-order Split	89.4	26.9	-2.6

Table 3 analyzes the impact of different sampling frequencies on system performance. The experimental results show that as the reference frequency gradually increases from 1 Hz to 10 Hz, the F1 score and MTW both gradually increase. This indicates that using a higher sampling frequency allows the system to obtain more detailed information about leakage events. However, as the sampling frequency increases, the MTW value also gradually increases, reaching a performance saturation point of 92.3% at 50 Hz. This demonstrates that a more detailed sampling rate can capture more details and improve performance, but performance tends to saturate after reaching a certain threshold.

Table 3. Impact of different sampling frequencies on system performance

Sampling strategy	H_2/P/T/VIS (Hz)	AE (Hz)	F1 (%)	MTW(s)	GPU Mem (GB)
Low Sampling	1	1k	85.6	18.3	4.2
Medium Sampling	10	10k	92.0	28.7	8.7
High Sampling	50	100k	92.3	29.1	15.6
Hybrid Sampling (Proposed)	10	100k	91.8	28.5	5.5

Table 4 analyzes the impact of different graph neural network architectures on overall performance. We compare various architectures, including gCN and GAT. Overall, the proposed architecture of hybrid dependency graph and bidirectional long short-term neural network shows significant advantages over the others, especially in terms of MTW (Mean Time To Value). Our model achieves 6.3 seconds faster performance than GAT, demonstrating that it not only distributes attention across different nodes but also further constrains the rationality of the graph structure through theoretical physical prior knowledge.

Table 4. Impact of different graph neural network architectures on overall performance

Figure/Time Series Model	F1 (%)	MTW(s)	Number of parameters (M)
LSTM-AE (No Graph)	80.4	15.6	1.8
GCN	88.1	20.2	2.3
GAT	89.4	22.4	2.5
GraphSAGE	87.9	21.0	2.1
MP-GETM (Ours)	92.0	28.7	3.1

Table 5 analyzes the warning effect under different threshold conditions. We analyze the impact of fixed threshold and moving quantile methods on model performance. The fixed threshold method has a high recall rate, but the precision is only 78.4%, resulting in very frequent false alarms. Using the mean method, it is very effective under stable conditions, but it is prone to false alarms under instantaneous changes such as system startup and shutdown. The moving quantile dynamic threshold method proposed in this paper can improve the system's precision while maintaining a high recall rate, achieving an F1 score of 92%.

After adding system uncertainty correction, the risk of emergency false alarms decreased by 42%, further improving the reliability of the system's warnings.

Table 5. Early warning effect under different threshold conditions

Threshold strategy	Precision (%)	Recall (%)	F1 (%)	Emergency false alarm rate (%)
Fixed Threshold	78.4	93.1	85.1	18.7
3-Sigma	82.6	89.3	85.8	15.2
Sliding Quantile (Ours)	92.4	91.6	92.0	9.8
+ Uncertainty Filter	93.1	90.9	92.0	5.7

5. Conclusions

This paper proposes a novel model, the MPGETM model, which achieves high performance in hydrogen energy storage safety early warning. The model integrates multiple steps, including data perception, graph augmentation, model inference, and early warning modules, forming a coordinated hydrogen energy storage early warning system. At the data perception level, six modal data types are involved, with hydrogen sensors contributing the most to modal data. The graph augmentation module amplifies weak early warning signals, avoiding missed detections due to electromagnetic interference and other factors. Experimental results show a detection accuracy of 92%. The MTW (Mean Time To Warning) early warning value is 28.7 seconds, demonstrating strong robustness compared to other models. Interpretability tools are used to analyze the model results, providing valid evidence for model decision-making. However, this research has some limitations. First, our data primarily uses simulated data, mainly from high-pressure gas storage platforms at tens of megapascals, while data from other routes has not been fully validated. Second, our model relies on a large number of sensors. In practical applications, considering cost constraints, it is not feasible to have so many sensors simultaneously. In the future, we will further expand the data from different scenarios to build a cross-scenario general lightweight anomaly detection model. We will also continue to combine various neural networks such as target information networks and causal reasoning networks to improve the model's generalization ability in different scenarios.

References

[1] Yin XF, Lin J, Wang ZG, Cheng YJ, Feng G, Wei M, et al. Size-dependent hydrogen sensing performance of Pd species on g-C3N4: From single atoms to nanoparticles. International Journal of Hydrogen Energy. 2025; 193:6. DOI: 10.1016/j.ijhydene.2025.152391

[2] Xiao JS, He P, Li XF, Bénard P, Yang TQ, Chahine R. Computational fluid dynamics model based artificial neural network prediction of flammable vapor clouds formed by liquid hydrogen releases. International Journal of Energy Research. 2022;46(8):11011-26. DOI: 10.1002/er.7902

[3] Duan QL, Xin J, Zhang HC, Hou ZS, Duan PY, Jin KQ, et al. Hydrogen leakage in underground garages: Impact of leakage locations on gas diffusion and safety considerations. Journal of Energy Storage. 2025; 117:12. DOI: 10.1016/j.est.2025.116123

[4] Du ZY, Dai ZX, Yin SX, Dong SN, Zhang XY, Yin HC, et al. Combining machine learning and multi-objective optimization algorithms to optimize key parameters for underground hydrogen storage. Gas Science and Engineering. 2025; 142:19. DOI: 10.1016/j.jgsce.2025.205713

[5] Wang SL, Bi YB, Shi JH, Wu QL, Zhang CT, Huang SS, et al. Deep learning-based hydrogen leakage localization prediction considering sensor layout optimization in hydrogen refueling stations. Process Safety and Environmental Protection. 2024; 189:549-60. DOI: 10.1016/j.psep.2024.06.122

[6] Deng J, Fan YC, Wang CP, Yang NN. Advances in hydrogen leakage jets for hydrogen storage systems. International Journal of Hydrogen Energy. 2024; 93:585-606. DOI: 10.1016/j.ijhydene.2024.10.439

[7] Zhang LY, Jiang N, Zheng ZC, Chi HW, Huang DH, Ye ZZ, et al. Ultrasensitive detection of H2 based on WO3 nanocubes decorated with PtO nanoparticles. International Journal of Hydrogen Energy. 2024; 94:1464-75. DOI: 10.1016/j.ijhydene.2024.11.015

[8] Dai JX, Chen ZN, Yang RD, Wu ZY, Tang ZA, Hu WB, et al. Early Detection of Hydrogen Leakage Using Fiber Optic Hydrogen Sensor Based on WO3-PdPt-Pt Nanocomposite Films. Nanomaterials. 2025;15(11):11. DOI: 10.3390/nano15110836

[9] Pan XH, Xie QW, Wu DL, Zhang CG, Zang XW, Hua M, et al. Fast response and high contrast of monoclinic MoO3-based hydrogen gasochromic sensor. Applied Surface Science. 2025; 700:10. DOI: 10.1016/j.apsusc.2025.163182

[10] Lv H, Shen YH, Zheng T, Zhou W, Ming PW, Zhang CM. Numerical study of hydrogen leakage, diffusion, and combustion in an outdoor parking space under different parking configurations. Renewable & Sustainable Energy Reviews. 2023; 173:15. DOI: 10.1016/j.rser.2022.113093

[11] Shi S, Lyu N, Jiang X, Song YH, Lu HF, Jin Y. Hydrogen gas diffusion behavior and detector installation optimization of lithium ion battery energy-storage cabin. Journal of Energy Storage. 2023; 67:11. DOI: 10.1016/j.est.2023.107510

[12] Xiao JS, Xu NF, Li YZ, Li GD, Liu M, Tong L, et al. CFD Simulation and ANN Prediction of Hydrogen Leakage and Diffusion Behavior in a Hydrogen Refuelling Station. International Journal of Energy Research. 2024; 2024:17. DOI: 10.1155/2024/8910533

[13] Du ZY, Xu LL, Yin SX, Dong SN, Dai ZX, Ma Y, et al. Enhanced prediction and uncertainty analysis for hydrogen production rate in depleted oil and gas reservoirs using advanced machine learning techniques. Geoenergy Science and Engineering. 2025;249:13. DOI: 10.1016/j.geoen.2025.213795

[14] Lee J, Oh S, Ma B. Performance analysis of optimized machine learning models for hydrogen leakage and dispersion prediction via genetic algorithms. International Journal of Hydrogen Energy. 2025; 97:1287-301. DOI: 10.1016/j.ijhydene.2024.10.183

[15] Qiu F, Shen ZY, Bai YZ, Shan GB, Qu DR, Chen WW. Hydrogen defect acoustic emission recognition by deep learning neural network. International Journal of Hydrogen Energy. 2024; 54:878-93. DOI: 10.1016/j.ijhydene.2023.09.176

[16] Liang Z, Yang YH, Wang YJ, Zhang M, Zhuang YF. Multidimensional quantitative modeling fusion analysis of safety risks in hydrogen refueling stations: A case study of a station in Beijing. Journal of Loss Prevention in the Process Industries. 2026; 100:22. DOI: 10.1016/j.jlp.2025.105865

[17] Wang T, Huang LC, Wu HL, Li WJ, Lu Q, Han R, et al. Ce0.8Gd0.2O1.95 based mixed potential gas sensor: AgRu bimetallic co-regulated WO3 for H2 sensing under high temperature. Sensors and Actuators B-Chemical. 2024; 402:14. DOI: 10.1016/j.snb.2023.135105

[18] Jiang YC, Xing ZX, Xu Q, Wu J, Peng M, Liu YC. Research on fence protection for liquid hydrogen leakage in the storage tank area. Journal of Energy Storage. 2024; 95:14. DOI: 10.1016/j.est.2024.112481

[19] Diao ST, Li HT, Wang JC, Wei CC, Yao YH, Yu MG. Hydrogen leakage location prediction for fuel cell vehicles in parking lots: A combined study of CFD simulation and CNN-BiLSTM modeling. International Journal of Hydrogen Energy. 2025; 109:115-28. DOI: 10.1016/j.ijhydene.2025.02.065

[20] Xu DY, Lu LQ, Wang ZL, Zhang LJ, Pan XH, Jiang JC. Experimental Study on the Influence of Ignition Position on the Overpressure of Hydrogen Jet Flame. Acs Omega. 2024;9(36):37869-81. DOI: 10.1021/acsomega.4c03651

[21] Liu XJ, Chen XY, Xiao Y, Yang ZM, Xu Y, Ault AP, et al. PtNi Nanocrystal-Ionic Liquid Interfaces: An Innovative Platform for High-Performance and Reliable H2 Detection. Acs Sensors. 2025;10(6):3993-4005. DOI: 10.1021/acssensors.4c03564

[22] Yoo BH, Wilailak S, Bae SH, Gye HR, Lee CJ. Comparative risk assessment of liquefied and gaseous hydrogen refueling stations. International Journal of Hydrogen Energy. 2021;46(71):35511-24. DOI: 10.1016/j.ijhydene.2021.08.073

[23] Li SX, Zhou SY, Zhao SY, Jin TF, Zhong MH, Cen ZH, et al. Room Temperature Resistive Hydrogen Sensor for Early Safety Warning of Li-Ion Batteries. Chemosensors. 2023;11(6):22. DOI: 10.3390/chemosensors11060344

[24] Zhang JX, Wang L, Shi MH, Li ZH, Liu BJ. Analysis of risk evolution mechanisms for hydrogen leakage in HECS: A dynamic Bayesian network and scenario deduction approach. International Journal of Hydrogen Energy. 2025; 163:14. DOI: 10.1016/j.ijhydene.2025.150804

[25] Yuan YP, Cui WY, Tong L. Prediction and sensitivity analysis of hydrogen leak diffusion using CFD and data-driven modeling under variable leak and wind conditions. International Journal of Hydrogen Energy. 2026; 202:18. DOI: 10.1016/j.ijhydene.2025.153056