

A Variational Auto-Encoder Method in Early Fault Alarm of Smart Meters

Zhou Yang^{1*}, Yuxiang Peng¹, Wenqian Jiang¹, Juntao Pan¹, Yanfu Li²

¹ China Southern Power Grid Guangxi Power Grid Co Ltd., China

² Tsinghua University, China

yangzhouwanshui@163.com, 18587748446@163.com, jiang_wq@gx.csg.cn,

panjuntao1985@163.com, liyanfu@tsinghua.edu.cn

Abstract

This study presents an effective solution for detecting faults early on in smart meters. The proposed method involves using a variational auto-encoder to evaluate the operational conditions of the meters by creating a variational feature state space matrix from the power consumption data. The use of a variational auto-encoder automates the generation of a state model, which allows for accurate detection of early faults. The method works by calculating a predefined state model and the state model, and comparing it to a self-generated threshold. This enables prompt alerts when issues are detected. The proposed approach has been experimentally validated and has shown better performance compared to a comparison method that relies on high-dimensional feature space. The accuracy rate of up to 91.67% has been achieved for alarming smart meter failures.

Keywords: Smart meters, Early fault alarm, Variational auto-encoder, State model

1 Introduction

It is proposed that the inevitable option for achieving the emission peak and carbon neutralization aim is to develop a new power system with new energy as its main component. This is made possible by AMI (Advanced Metering Infrastructure), a critical component of the smart grid [1], which performs the essential task of terminal sensing for the modern power system and integrates with the Internet of Things, cloud computing, and new communication technology to create a vast smart meter network, which serves as the “last mile” of sensing and controlling carbon emissions [2]. The two primary purposes of AMI are automatic meter reading (AMR) and advanced data mining (ADM), which together provide an overview of the technology (Figure 1).

The volume and complexity of AMI are growing as a result of the emergence of advanced IT technologies including the Internet of Things (IoT), Cloud computing, and 5G. In China, there will be roughly 600 million smart meters by 2020, and by 2025, there may be more than 1 billion of them. Such a sizable IoT system would not only contribute to the development of new national infrastructure but also

carry significant hazards related to the security of data and electricity.

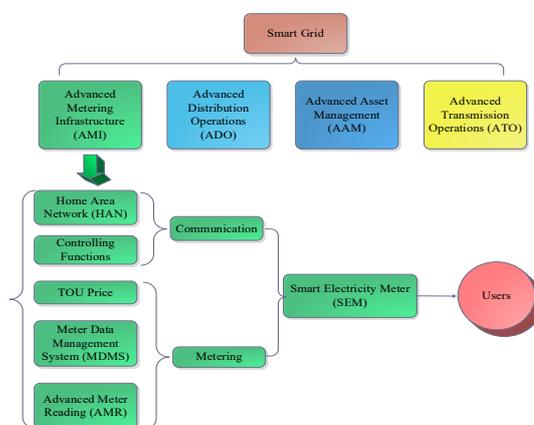


Figure 1. Smart grid's functions including smart meters [3]

In recent years, smart meters have become an essential component of power grid, facilitating the great achievement of the Smart Grid information collecting system [4-5]. As the primary measurement equipment in the system, smart meters play as a foundation in communicating original power measurement data, achieving smart meters information integration [6]. However, the efficient maintenance of a large number of smart meters remains a significant challenge for power grid safety. To reach balance of the maintenance cost and the financial benefits of smart meters, detecting early defect alert for devices is imperative. Traditional methods of evaluating the operating status of devices solely by human are extremely challenging, as the accuracy and timing of the findings cannot be trusted due to the sheer volume of operation data and the diverse range of potential defects [7]. Artificial intelligence with their exceptional data processing capabilities, offer a potential solution to this problem and are increasingly being applied in defect diagnosis. Zhang et al. [8] proposed a ideally ensemble empirical mode decomposition-convolutional deep belief network and achieved good results in fault diagnosis of reciprocating compressors. Zou et al. [9] achieved accurate bearing fault diagnosis by using an anti-noise one-dimension convolutional neural network. Xue et al.

*Corresponding Author: Zhou Yang; E-mail: yangzhouwanshui@163.com

DOI: <https://doi.org/10.70003/160792642024122507009>

[10] proposed a deep convolution neural network and support vector machine-based fault diagnosis method for rotating machinery and the average diagnostic accuracy increased to 98.71%. However, there is limited research for early alarms of smart meters failures.

Several studies in the reliability field have examined smart meter reliability prediction approaches from a variety of perspectives. According to LCM (Life Cycle Management), the cost and accuracy of associated prediction systems should fluctuate noticeably at various stages. As shown in Figure 2, various reliability prediction techniques, including manual prediction, simulation prediction, reliability test, on-site failure data analysis, and PoF (Physics of Failure) analysis, are used in a typical LCM. The ability to anticipate product reliability with products from the demonstration, design, manufacture, operation, and scrapping stages has gradually increased.

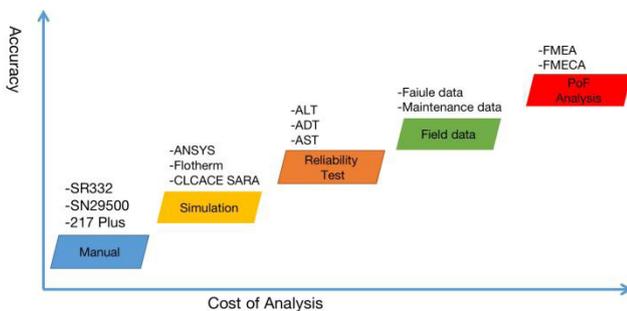


Figure 2. Life cycle reliability analysis

In the product design and demonstration stages, manual prediction tools like the Bellcore SR-332, Siemens SN29500, and ML-217 Plus have been quite helpful [11]. For engineers, it is affordable and simple to use, albeit the accuracy heavily depends on the database. In the iterative stage of product design, simulation prediction methods are frequently used. Structural reliability is typically predicted using ANSYS and Flotherm in the aerospace, shipbuilding, and other industries. Sara from the University of Maryland's Computational Center is also more experienced when it comes to designing electronic products. Mathematical modeling and simulation in Matlab are typically utilized for complicated unusual systems. There are numerous studies and literature in this field. The literature [12] used a BP neural network for switching power supply reliability prediction. The literature [13] proposed a reliability prediction method that is appropriate for highly complex systems. The literature [14] proposed an aqa-based Software reliability prediction method based on the LVQ neural network.

When the technical finalization of the product is carried out to confirm the indicators at the product delivery stage, it is necessary to conduct several product tests through reliability tests, such as accelerated life test (ALT), accelerated degradation test (ADT), accelerated stress test (AST), and so on. The literature [15] is the most productive and concentrated area of dependability work. The issue of no failure data was resolved by the accelerated degradation test, which produced positive findings for the reliability prediction of smart meters.

The on-site operation data can accurately forecast the dependability of operating products and direct the use of upcoming items in real operations [16]. Essentially, a thorough analysis of the effect areas of physical structures, such as materials, structures, assembly processes, and the usage environment, is required to explain and predict the reliability of products. This necessitates performing a failure analysis and looking at the foundational layer of the failure mechanism. From this vantage point, the product prediction can comprehend why it exists. Therefore, numerous well-known companies, like General Motors, Intel, and Boeing, no longer employ the conventional forecast method in favor of a reliability prediction method based on failure mode and failure mechanism [17].

KPCA (Kernel Principal Component Analysis) is a nonlinear data processing technique that may create a high-dimensional feature space from the input data [18]. To create a high-dimensional feature space (HDFS) from the equipment operating data, Zhao et al. [19] employed the KPCA method. By comparing changes to the HDFS, they were able to identify early fault warnings. However, the choice of the kernel function or its parameters has a significant impact on the early defect alert method based on KPCA; poor selection will have a significant impact on the output.

The state model of the input data can be created automatically by a variational auto-encoder (VAE) [20]. Using VAE, Gregor et al. [21] created the state model of the image data and used this model to create similar image data. Tran et al. [22] used VAE to obtain the features of the human face image and create a state model of the features. They finished reconstructing the photos of the human face using this approach. Wang [23] applied deep VAE to finish the unsupervised dimension reduction and visualization of Single-cell RNA sequencing data. Na [24] used VAE to achieve data reduction and feature extraction and then combined VAE with DNN to create high-precision gas dispersion models.

VAE and several of its derivation models are extensively employed in image, voice, vision, and language due to their powerful statistical distribution generation capabilities for complicated data [25-26], but they have not been applied to the anomaly detection of equipment functioning status. In this paper, we propose an early fault alarm approach for smart meters based on VAE. As input data, the variational feature state space matrix is employed, which is made up of several key indicators retrieved from the power consumption data. From the variational feature state space matrix, the approach automatically generates the state model using VAE. By meticulously calculating the disparities between predetermined and changing state model, and juxtaposing it against an automatically generated metric, smart meters possess the capability to provide early warning indicators [27].

2 FMEA of Smart Meter

A. Physical Structure of Smart Meter

The typical Smart meter is composed of a current sampling transformer, voltage sampling transformer, and

measurement integrated circuit. The data processing unit is composed of an MCU (Microcontroller Unit), data internal cards, power-off detection, and calendar clocks. The input and output system is composed of LCD (Liquid Crystal Display) components, meter pulse, keyboards, external card connectors, and auxiliary terminals. The power supply unit is composed of a transformer, a linear voltage source, and a battery. Figure 3 displays the smart meter’s building block diagram.

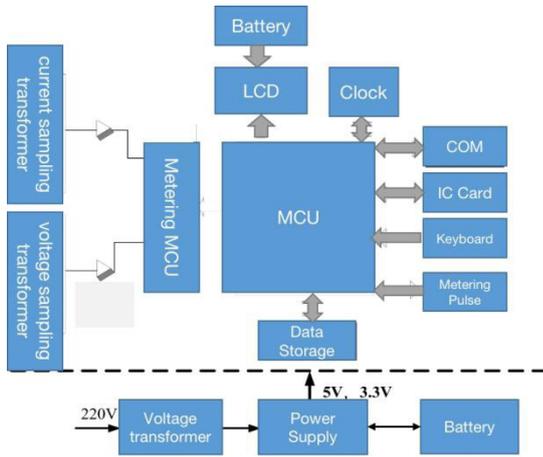


Figure 3. A typical smart meter structure

Its basic operation entails sampling the user’s supply voltage and current in real-time, processing that signal using a smart meter’s special integrated circuit to produce a pulse output proportional to the electric energy, and then processing that output under the control of a single chip computer to display the pulse as power consumption.

B. Performance Index of Smart Meter Component

The key characteristic of smart meters encompass measurement errors, daily timing errors, meter constant, communication functionality, and LCD. The key characteristic of smart meters mainly include measurement errors, daily timing errors, meter constant, communication function, and LCD because these are the key factors that affect the accuracy and reliability of the data collected by the smart meters. Measurement error refers to the difference between the actual value and the estimated value of energy consumption, which can be caused by aging, faults, and other factors. Daily timing errors refer to the time deviation of measurement data due to inaccurate timing of smart meters². The meter constant refers to the ratio of the energy passing through the meter to the number of revolutions of the meter disk. Communication function refers to the ability of the smart meter to communicate with other devices in the network. LCD refers to the display screen on the smart meter that shows information such as energy consumption and billing information. These performance indicators are important for ensuring that smart meters provide accurate and reliable data for energy management and billing purposes. The faults mentioned in this paper mainly involve the above several indicators. When one of these indicators has an abnormal value, the electric energy meter is judged to be faulty.

1) Measurement error

The deviation between the measurement pulses of the reference standard meter and the tested meter at a specific load point is used to quantify the measurement error of a smart meter. Then it is compared to the reference measurement pulse as follows

$$E_m = \frac{P_{ref} - P_{test}}{P_{ref}} \times 100\% \tag{1}$$

where $E_m \geq 1\%$ is the failure criterion of the normal smart meter for normal users (class 1).

2) Daily timing error

The hardware clock circuit for the smart meter has a temperature correction feature, and its terminal output frequency is 1Hz. Within the temperature range of $-25\text{ }^\circ\text{C}$ to $+60\text{ }^\circ\text{C}$, the clock precision is less than $\pm 1\text{s/d}$. At the reference temperature of $23\text{ }^\circ\text{C}$, the clock accuracy is less than $\pm 0.5\text{s/d}$.

3) Measurement constant

The smart meter constant is a number that describes the relationship between the electric energy value recorded by the smart meter and the matching test output value. The number of pulses the A/D converter sends when the intelligent meter senses one degree of electricity or pulse constant can also be used to express it. The test output of the watt-hour meter and reading indication must be related in a way that is consistent with the nameplate, which can be calculated as shown in (2).

$$\Delta E = \left| \frac{n}{C} - E \right| < 1 \times 10^{-\alpha} \tag{2}$$

Where E is the reading value of the meter, n is the pulse constant of the meter, $*$ is the cumulative value of the meter and $*$ is the decimal place of the meter.

4) Communication

Infrared communication, carrier communication, and RS-485 communication are all part of the smart meter’s communication capabilities. A 485 communication defect can be used to verify whether a smart meter’s 485 communication is functioning normally by reading the smart meter’s address information through 485 communication on the test bench.

5) LCD

The LCD of smart meters can instantly determine whether something is normal to the unaided eye. The upgraded pre-test results showed three potential scenarios for liquid crystal displays operating at high temperatures: the brightness of the liquid crystal dims slightly, the reading is only dimly visible when the liquid crystal darkens, and as the LCD lacks a display characters cannot be seen on it.

C. Effect of Stress on Smart Meters

According to the analysis of the actual use environment, the environmental stress of the smart meter includes temperature, humidity, vibration, magnetic interference, environmental pressure, and working electrical stress (Table 1). The impact of the environment and workplace stress on smart meters will next be examined, along with the process by which temperature, humidity, electrical stress, and other elements affect smart meters.

1) Influence of temperature on smart meter

The a sampling circuit is divided into voltage sampling and current sampling. The current sampling is mainly composed of magnet shunt. Defined that ρ_0 is essentially a metal conductor with a small temperature coefficient of resistance value. According to the change characteristics $\rho = \rho_0(1+\alpha t)$ of the resistance of the metal material with temperature (where t is the temperature in Celsius, the resistance at 0 °C, and the temperature coefficient of resistance). It is clear that as the temperature rises, the resistance of the magnet resistance increases, increasing both the resistance value and the current measured by the shunt, or the current value sampled by the meter. The current transformer uses more energy than the manganin shunt and is an electromagnetic component. The used copper material has a high-temperature coefficient, and the temperature has a significant impact on the characteristics. They both will result in sampling error, which will then have an impact on measurement error.

One of the main factors that affect the time accuracy of the real-time clock chip is the accuracy and stability of the clock crystal oscillator. The temperature drift of the crystal oscillator is the primary cause of its accuracy deviation. The frequency drift rate of a crystal oscillator, also known as the aging rate, along with the temperature characteristics of the crystal oscillator affects the momentum of the oscillator's frequency change, which invariably results in the daily timing error.

The rated accuracy index can only be guaranteed within a specific working range since the temperature will impact the weighted resistance network of the operational amplifier. ADC accuracy is significantly impacted by offset and gain errors that are brought on by variations in ambient temperature. Make the meter consistent out of the ordinary. Temperature will also impact the zero bias of the circuit, cause the operating of the amplifier point to wander, and in extreme situations, render the amplifier useless. The temperature sensor used by the measuring chip is essentially a temperature diode, whose output voltage varies as a function of temperature. Its output voltage roughly follows a linear relationship with temperature. It is simple to create interchangeability errors and linearity calibration errors when the temperature field varies significantly, which leads to sensor failure and measurement error. HTN-type liquid crystal is primarily the liquid crystal material used in single-phase smart power meters. The Arrhenius reaction rate theoretical model predicts the impact of temperature stress on liquid crystal failure.

2) Influence of humidity on smart meters

The principal way that humidity damages semiconductor goods is by allowing moisture to enter the integrated circuit (IC) through the pin gaps and other openings in the plastic packaging of the IC. The high-temperature environment produces water vapor, which leads to pressure that causes the IC resin packaging to split and oxidize the metal inside the IC device, which leads to product failure.

3) Influence of electrical stress on smart meters

The thermal imbalance of the working environment of the meters is the primary indicator of how electrical stress affects smart meters. The current sampling circuit of the meters is

what has the most impact. The coil of the current sampling circuit produces heat. To accurately measure electricity, the meters must be in a thermally stable state. The internal thermal stability of the meter will become out of balance under excessive electrical stress, which will lead to inaccurate measurement.

Table 1. Environmental stress factors of smart meter failure

Performance indicators affected	Corresponding main external environmental factors
Measurement error	Temperature
	Electrical stress
Daily timing error	Temperature
	Humidity
Measurement constant	Temperature
Communication	Temperature
Lcd	Temperature

3 State Model

A. State Model

VAE (Variational Autoencoder) is an algorithm used for data generation. It can extract features from input data and create a probability distribution model that reflects the original form of the input data. This model can generate both original and new data that resemble the input data. The principle theory of VAE is shown in Figure 4. y The feature from input x and θ is represented as Φ . θ represents the generative parameter, which is included in the $p_\theta(y)$ and $p_\theta(x|y)$. The ϕ variational parameter of the approximate posterior $q_\phi(y|x)$ is represented as ψ . The desired state model is achieved through the posterior $p_\theta(y|x)$, which involves the feature extraction process. Finally $p_\theta(y)p_\theta(x|y)$, the generation model, represented as ϕ , is used to produce new data. This completes the new data generation process.

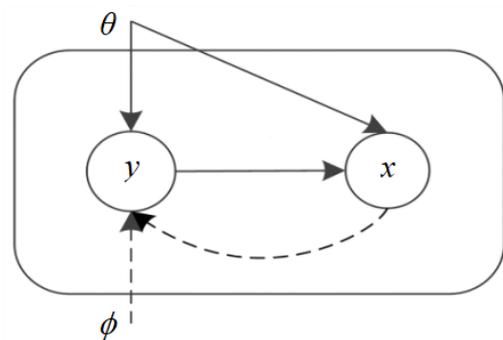


Figure 4. Working principle of VAE

The state model should be modified as

$$p_\theta(y|x) = \sum_{k=1}^K \alpha_k p_{\eta_k}(y_k) \tag{3}$$

It is evident that the model comprises K sub-models, where α_k and η_k are the weights and the parameters, respectively.

B. Generalized State Model

The state model $p_\theta(y|x)$ can be:

$$p_\theta(y|x) = \frac{p_\theta(x|y)p_\theta(y)}{p_\theta(x)} \quad (4)$$

The computational complexity of (4) is prohibitively high due to the difficulty in estimating $p_\theta(x)$. Consequently, a variational approximation $q_\varphi(y|x)$ is utilized to evaluating $p_\theta(y|x)$. The D_{KL} is Kullback-Leibler divergences frequently defied by [26]. $D[q_\varphi(y|x)||p_\theta(y|x)]$ can be formulated as follows:

$$D[q_\varphi(y|x)||p_\theta(y|x)] = E_{q_\varphi(y|x)} \log \frac{q_\varphi(y|x)}{p_\theta(y|x)} \quad (5)$$

And $\log p_\theta(x)$ can be written as:

$$\log p_\theta(x) = D(q_\varphi(y|x)||p_\theta(y|x)) + C(\theta, \varphi; x) \quad (6)$$

where $D[q_\varphi(y|x)||p_\theta(y|x)] \geq 0$.

And (4) should be written

$$\log p_\theta(x) \geq C(\theta; \varphi; x) \quad (7)$$

where

$$C(\theta, \varphi; x) = -D(q_\varphi(y|x)||p_\theta(y)) + E_{q_\varphi(y|x)} \log p_\theta(x|y) \quad (8)$$

The $C(\theta; \varphi; x)$ is stochastic gradient variational Bayes.

Set C_1 and C_2 as follows:

$$C_1' = \frac{1}{2} \sum_{m=1}^M \{1 + \log[(\sigma_m)^2] - (\mu_m)^2 - (\sigma_m)^2\} \quad (9)$$

where σ_m is mean and μ_m is the standard deviation.

$$C_2' = \frac{1}{C} \sum_{c=1}^C \log P_\theta(x|y^{(c)}) \quad (10)$$

$$y^{(c)} = g_\varphi(x, \varepsilon^{(c)}) \quad (11)$$

where $\varepsilon^{(c)} \sim N(0, I)$.

Therefore (8) can be obtained by:

$$C(\theta; \varphi; x)' = \frac{R}{T} \sum_{t=1}^T (C_1' + C_2') \quad (12)$$

where R and T represent data points from X and in the sample drawn from x .

4 Method

Since most users' electricity usage follows a regular pattern, especially when a large number of users are aggregated into one substation load, the substation load curve exhibits a certain degree of regularity. Furthermore, the power consumption data at a specific stage remains relatively consistent throughout the normal operation of smart meters. This consistency may suggest that the change data follows a specific distribution. A state model is created based on the distribution of change data across different phases. Any changes in power usage caused by a smart meter malfunction will alter this state model. By utilizing the power consumption change data from various periods as input, the modeling capabilities of VAEs enable us to efficiently construct the state model for smart meters.

The early alarm method's principle is depicted in Figure 5 as below.

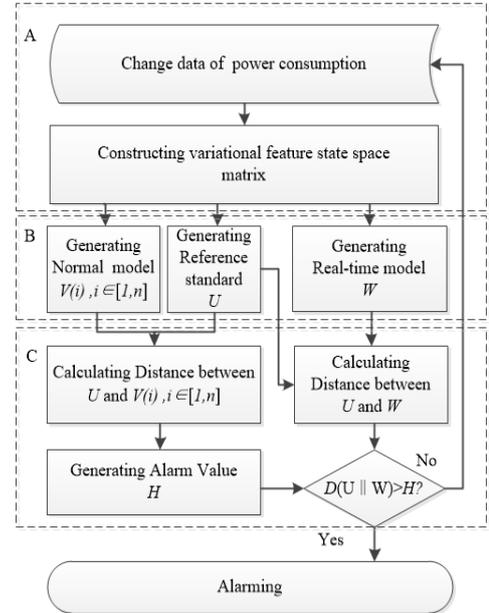


Figure 5. Working flow of early alarm method

A. Preprocessing of Input Data

To build the state model, we utilize data sets b derived from the original data, with each data set q containing multiple groups of data. We derived the feature p and obtain the state space matrix $S(a)$, $a \in [1, b]$ as:

$$S(a) = \begin{bmatrix} S_{1,1} & \cdots & S_{1,p} \\ \vdots & \ddots & \vdots \\ S_{q,1} & \cdots & S_{q,p} \end{bmatrix}_{p \times q}, a \in [1, b] \quad (13)$$

The resulting state model can then be used to analyze the system's behavior, make predictions about future states, and identify any potential issues or trends.

Three distinct features, namely variance, mean value, and peak value, are extracted to properly reflect the changes in

the input data. Therefore, when a smart meter malfunctions, the input data it measures will change, and the data features will also change accordingly. This change will be directly reflected in the variational feature state space matrix, thus achieving smart meter problem detection.

B. Generation of State Model

$S(a)$, $\alpha \in [1, b]$ is the training data. All b datasets cover the entire deterioration process of the smart meters from normal to failure. The first data set during the normal operation stage is used to train the reference standard U , while the subsequent n sets of normal operation stage datasets are used to train V . As a reference standard, the state model created from the usual power consumption change data is defined as (14). The reference standard is contrasted with the real-time state model in (15).

$$U = \sum_{i=1}^N \alpha_i p_{\eta_{U_i}}(E_1) \quad (14)$$

$$W = \sum_{j=1}^N \beta_j p_{\eta_{W_j}}(E_2) \quad (15)$$

where E_1 and E_2 are the normal and real-time input data respectively. To optimized θ_U of the state model U , the stochastic gradient variational Bayes is applied, which contains α_i the η_{U_i} weight of each sub-model and the parameter of each sub-model.

C. Determining Alarm Threshold

Considering the normal defect of smart meter, the state model of some sub-distributions will deviate from predetermined standard. We utilize the notation D_{KL} to assess the discrepancies between the predetermined standard and the real-time state model based on their specific characteristics. A failure of the equipment may be inferred if the calculated result exceeds a pre-determined threshold. The difference between the predetermined standard U and the real-time model W is defined as per Equation (16).

$$D(U \| W) = \sum_{k=1}^K \alpha_k [KL(U_k \| W_{l(k)}) + \log \frac{\alpha_k}{\beta_{l(k)}}] \quad (16)$$

where:

$$l(k) = \arg \min [D(U_k \| W_l) - \log \beta_l] \quad (17)$$

The thresholds for most monitoring techniques are often determined by practical expertise. However, given the wide range of potential issues that can impact smart meters, which manual adjustment of alert settings based on past experiences is not feasible. To address this challenge, we present a technique that automatically determines warning levels using historical data set. This approach offers broader applicability and ensures that each smart meter's alert value is tailored to its unique status.

To establish these warning levels, we first assume that the quantized values of the normal state models for smart meters follow a normal distribution. We apply the alarm value according to the 3delta criterion as follows:

$$H = \mu_D + 3\sigma_D \quad (18)$$

If the quantized values of the N normal state models follow a normal distribution, and the distance between each model and the predetermined standard has been quantified, we can calculate the alarm value using Eq. (18). Here, μ_D and σ_D are the average value and the standard deviation.

5 Experiments

To ensure the effectiveness of VAE method, an experiment was conducted using 12 groups of historical fault case data from various smart meters, the details of these data are shown in Table 2. Each set of data is extracted from the monitoring data of the faulty smart meter in reality. Taking one of the faulty smart meters as an example, the time domain data of the positive active energy is shown in Figure 6.

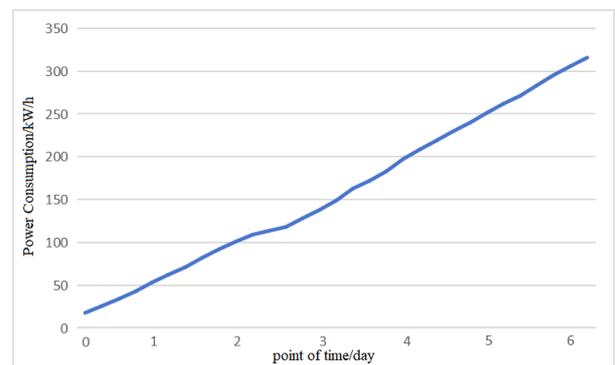


Figure 6. Observed positive active energy

A. Fault Case Data

One of the smart meters used in this test is shown in Figure 7. Each fault case contains the same number of data sets, and each data set contains 10 different types of data. The new sparse fault case data is obtained by sampling from each type of data of the original fault case. To balance computational efficiency and alarm accuracy in experiments, the sampling period is set to 1 minute, and one value is taken for each sampling. The sampling period is also applicable for practical use. The fault sample data can be obtained by calculating the change in the power consumption per minute of each type of data in the newly generated fault case data. the sampling interval is set to 10 minutes.

The experimental data is divided into numerous datasets at regular intervals, each containing 240 data points. This approach enhances the model generation efficiency, leveraging the effectiveness of the suggested method. To establish the normal state model, we choose the first 10 datasets with the initial dataset as the benchmark.



Figure 7. Smart meter for test

B. Data Collecting System

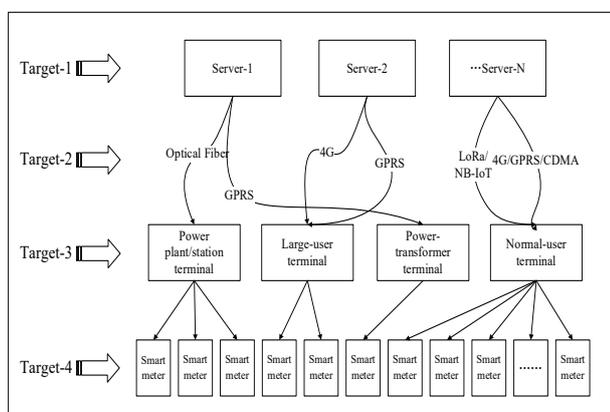


Figure 8. AMI system for data collecting

The AMI system is taken into account in this paper. The system, which consists of a main station, a communication network, data terminators, and smart meters (Figure 8), covers over 20 million smart meters of end users. The information of users, including power consumption, voltage, current, and special issues, will be recorded and transmitted periodically. It is necessary to frequently capture end-user data, such as power usage, current, and voltage, for instance, every 15 minutes. A transmission procedure from smart meters will be gathered by various types of data terminals due to the trade-off in data communication costs. The data terminal communicates with the main station using a variety of communication techniques to set the time.

Table 2. Details of fault case data

Fault type	Number of cases
Overload burning	4
Programming error	3
Electrical and mechanical failure	5

In Table 2, programming errors in the context of power consumption measurement can refer to errors in the software or firmware that is responsible for measuring and reporting power consumption. These errors can cause incorrect measurements or reporting of power consumption values. Depending on the nature of the error, it could be present all the time or only at specific points in time. For example, if

there is an error in the algorithm used to calculate power consumption, it could result in consistently incorrect measurements. On the other hand, if there is an error that only occurs under certain conditions, such as when a specific sensor fails, it could result in incorrect measurements only when those conditions are met. Overloading conditions or mechanical failures can also affect the measured value of power consumption. Overloading refers to a situation where a component or system is subjected to a load that exceeds its rated capacity. This can cause the component or system to draw more power than it normally would, resulting in higher power consumption measurements. Mechanical failures, such as a seized bearing or a broken gear, can also cause changes in power consumption. For example, if a bearing seizes, it can cause increased friction and resistance, which can result in higher power consumption as the motor works harder to overcome the increased resistance.

C. Results and Analysis

The results obtained via the state model-based method (SM-based method) are summarized in Table 3. It is evident from the data presented in this table that 11 cases can be early alarmed, representing 91.67% of the total number of cases. This finding supports the conclusion that the SM-based method exhibits a high level of early alarm effectiveness for detecting failures in smart meters. To compare its performance with the SM-based method, the early alarm approach described in reference (HDFS-based method) was evaluated using the same failure case data. The outcomes indicate that the HDFS-based strategy is only 50% effective and achieves early defect alarm only half of the time. Furthermore, Table 4 reveals that the SM-based method constructs models more rapidly than the HDFS-based method, indicating that the proposed method may be more computationally efficient. Table 5 compares the average advance time scales for three types of failures by both the SM-based approach and the HDFS-based method. The SM-based method exhibits better performance in all time scale for all failures.

Table 6 shows the experimental results of the HDFS-based method for various types of faults. The method performs well in detecting overload burning. However, it faces challenges in implementing effective early alarms for different kinds of failures. In contrast, the SM-based method demonstrates high early warning effectiveness for all kinds of failures. This suggests a wide range of potential applications for the SM-based method. To further improve the early alarm effectiveness of the SM-based method, we will investigate reducing the sampling interval in our next stage of research, as one of the fault cases was missed by the alarm.

Consider the instance of excess burning. Figure 9 displays the results of condition monitoring from the two overload burning methods. The solid curves indicate the real-time distance values, while the horizontal dotted lines represent the alarm thresholds automatically established by each method. It is evident from Figure 9 that both approaches effectively distinguish between normal and abnormal data. However, the SM-based approach raises an alarm 21 minutes before the actual failure, while the HDFS-based approach provides a 10-minute advance. Consequently, the SM-

based strategy outperforms the HDFS-based method. The robustness of the alarm threshold ensures that when the distance value surpasses the alarm level, it signifies system recognition of a fault rather than a false alarm. Notably, all 11 early alarms generated by this method in the experiment were valid.

The number of data points would be 120 and 480, respectively, to figure out the relationship between size of the training data set and the average advance time in both approaches. The test results for newly divided data presented in Table 7 indicate that both early failure alert systems will experience a decline in effectiveness and average advance time scale when the dataset contains inappropriate data points. A lack of sufficient points can destabilize the state model, while an overabundance of points can render it insensitive to small changes.

The aforementioned test findings indicate that for common smart meter issues, the SM-based method outperforms the HDFS-based method. Considering the model generation effectiveness, the SM-based method are superior than the HDFS-based method. The KPCA-based method’s performance is significantly limited because the HDFS-based method should not be generative. Hence, it is challenged to detect minor changes from historical data sets. Additionally, the efficacy of HDFS-based method is greatly influenced by the choice of kernel function. The state model created by the SM-based method aligns with the dataset, enhancing its ability to capture minute changes. Furthermore, the SM-based method’s predetermined alert value exhibits a high level of adaptability.

Table 3. Effectiveness

Method	Samples	Alarms at early stage	(%)
SM-based method	12	11	91.67
HDFS-based method	12	6	50

Table 4. Efficiency

Method	Time consumption (s)
SM-based method	1.2
HDFS-based method	2.5

Table 5. Average advance time scale

Method	Overload burning (min)	Average early alarm time for Programming error (min)	Average early alarm time for Electrical and mechanical failure (min)
SM-based method	20.1	6.5	22.67
HDFS-based method	8.6	3	9.5

Table 6. Details of early alarm results of HDFS-based method

Fault type	Total number of samples	Number of samples alarmed at early stage
Overload burning	4	3
Programming error	3	1
Electrical and mechanical failure	5	2

Table 7. Training data set effect of both methods

Method	Rate of success (%)		Average advance time scale					
			Burning time (min)		Programming error time (min)		Electrical and mechanical failure time (min)	
	120	480	120	480	120	480	120	480
SM-based	75	83.33	16.8	19.4	5.9	6.35	20.6	21.4
HDFS-based	33.33	33.33	6.6	8.05	0	2	7.8	8.65

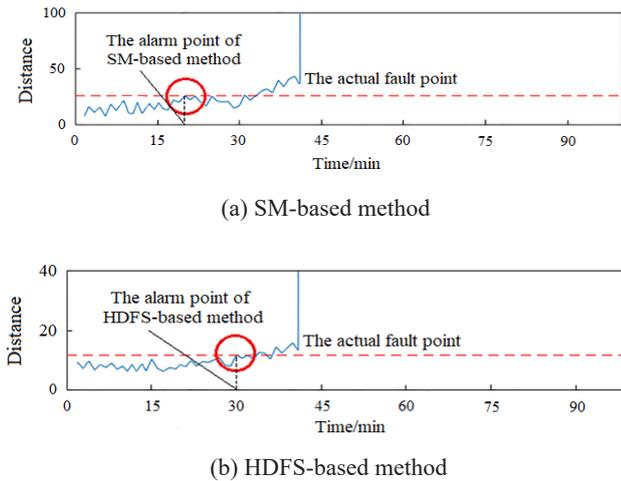


Figure 9. Monitoring results of two methods

D. Reliability Prediction by Manual

Electronic product dependability prediction methods have been the subject of several studies, and many useful standards have been developed. For instance, the mil-hdbk-217f, gjb/z 299c-2006, and PRISM built on this premise are typically used in military electronic equipment. Because military standards are typically conservative, iec-tr 62380, Telcordia sr-332, IEEE Std 1418, and other standards are more commonly utilized for civil electronic goods. The linked database is where these methods diverge most. These methods, when viewed from the perspective of prediction methods, are traditional prediction methods developed from mil-hdbk-217f: that is, they determine the basic failure rate of each type of component through a large number of statistical data, and then they correct the basic failure rate according to the actual working environment by introducing various correction factors to obtain the expected failure rate.

The Telcordia sr-332 served as the foundation for the dependability dynamic prediction method that is provided in this paper. Since Telcordia sr-332 also includes information on the standard deviation of the basic failure rate, this section first describes the reliability prediction method based on this standard. However, for simplicity, we do not take the fundamental failure rate's dispersion into account). Bell Laboratories has proposed Telcordia sr-332, formerly known as Bellcore tr-332, as a commercial standard for electronic device dependability prediction.

In the early fault alarm approach for smart meters, it is possible to consider integrating electronic product dependability prediction methods. For example, the Telcordia sr-332 standard can be used as a basis for dynamic dependability prediction of smart meters. First, FMEA is performed on the smart meter to identify its primary components. Then, a test design is used to calculate the electrical stress factor and the working current is measured using a data acquisition system built on the LabVIEW platform. Finally, the expected value of product dependability is calculated based on the test results and prediction method. This integrated approach can improve the accuracy and reliability of the early fault alarm approach for smart meters.

By dynamically predicting the dependability of smart meters, their status can be better monitored and anomalies can be detected promptly. This approach not only improves the safety and stability of smart meters but also provides an important reference value for engineering applications.

To demonstrate the suggested strategy, this section uses the dependability prediction of an electronic device as an example. Table 3 displays the composition of the product.

FMEA for this product should be done first. The magnetic holding relay is the primary component of the product, per the FMEA results. As a result, this component is created and tested, the amount of electrical stress applied is measured, and the failure rate is calculated using formula (3). According to the guidelines in the literature [8], the effect factor of electrical stress is assumed to be one for other non-critical components. The dependability predicted value of non-critical components is displayed in the last column of Table 3 using formula (4).

The following test design is used to calculate the electrical stress factor: a total of 8 test samples are chosen, and the test is carried out at 40 °C. Measure the working current of the magnetic hold relay after the product has operated steadily. The data acquisition system built on the LabVIEW platform measures working current. Figure 2 depicts the test device.

Table 4 displays the test results. The anticipated value of the estimated failure rate is displayed in Table 4 by formula (3). Using the conventional approach, the anticipated value of product reliability given from equation (1) is 110.3982 fit when dynamic prediction is not taken into account.

6 Conclusion

This study presents an advanced early fault detection approach for smart meters, leveraging the power of a Variational Auto-Encoder (VAE). Our approach relies on the extraction of a feature state space matrix derived from fluctuations in power consumption data. This matrix serves as the foundation for the identification of deviations from normal patterns, enabling the automatic generation of alarm values tailored to the unique characteristics of each smart meter. The alarm values are dynamically computed based on historical normal data, ensuring accurate correlation with the current state of each meter. This approach offers significant advantages in terms of adaptability and practicality, making it highly suitable for real-world engineering applications. Furthermore, the method's versatility extends beyond smart meters, enabling its application in various fields, including early warning systems for mechanical equipment. In summary, this study offers a promising new approach to enhancing the reliability and proactive maintenance of smart meters. By combining the power of deep learning and historical data analysis, the proposed method can detect potential faults with high accuracy and advance notice, significantly reducing downtime and operational costs. The approach's adaptability and versatility further enhance its value, opening up new possibilities for enhancing the reliability and efficiency of smart metering systems across various industries. Future research directions could

explore the integration of additional data sources, such as environmental factors or operational conditions, to improve fault detection accuracy even further. Additionally, exploring novel techniques for real-time monitoring and early fault detection could provide valuable insights for maintaining the overall performance and reliability of smart metering systems.

References

- [1] D. Singhal, L. Ahuja, A. Seth, An insight into combating security attacks for smart grid, *International Journal of Performability Engineering*, Vol. 18, No. 7, pp. 512-520, July, 2022.
- [2] M. Ordonez, M. Manbachi, AMI-Based Energy Management for Islanded AC/DC Microgrids Utilizing Energy Conservation and Optimization, *IEEE Transactions on Smart Grid*, Vol. 10, No. 1, pp. 293-304, January, 2019.
- [3] Z. Yang, Y. X. Chen, Y. F. Li, E. Zio, R. Kang, Smart Electricity Meter Reliability Prediction based on Accelerated Degradation Testing and Modeling, *International Journal of Electrical Power & Energy Systems*, Vol. 56, pp. 209-219, March, 2014.
- [4] D. Muyizere, L. K. Letting, B. B. Munyazikwiye, Effect on Transient Stability and Analyses Resulting from a Cyber-Attack on Frequency Relay Device, *International Journal of Performability Engineering*, Vol. 19, No. 1, pp. 20-32, January, 2023.
- [5] W.-H. Zhuang, Y. Zhou, A Survey of Cooperative MAC Protocols for Mobile Communication Networks, *Journal of Internet Technology*, Vol. 14, No. 4, pp. 541-559, July, 2013.
- [6] L. Wen, K. Zhou, S. Yang, L. Li, Compression of smart meter big data: A survey, *Renewable and Sustainable Energy Reviews*, Vol. 91, pp. 59-69, August, 2018.
- [7] J. Yang, M.-Y. Xin, J.-X. Ou, J.-R. Wang, Q. Song, Data accuracy judgment method of metrology automation system based on La Yida criterion, *Power Systems and Big Data*, Vol. 20, No. 11, pp. 74-78, November, 2017.
- [8] Y. Zhang, J. C. Ji, B. Ma, Fault diagnosis of reciprocating compressor using a novel ensemble empirical mode decomposition-convolutional deep belief network, *Measurement*, Vol. 156, Article No. 107619, May, 2020.
- [9] F. Q. Zou, H. F. Zhang, S. T. Sang, X. M. Li, W. Y. He, X. W. Liu, Y. F. Chen, An anti-noise one-dimension convolutional neural network learning model applying on bearing fault diagnosis, *Measurement*, Vol. 186, Article No. 110236, December, 2021.
- [10] Y. Xue, D. Y. Dou, J. G. Yang, Multi-fault diagnosis of rotating machinery based on deep convolution neural network and support vector machine, *Measurement*, Vol. 156, Article No. 107571, May, 2020.
- [11] Y. He, Z. Yang, Low Rank Tensor Approximate Discrete Simulation Method of Smart Meter Reliability Prediction, *6th International Conference on Information Science and Control Engineering (ICISCE)*, Changsha, China, 2019, pp. 360-366.
- [12] X. Gao, X. Diao, J. Liu, M. Zhang, Y. He, A multi-classification method of smart meter fault type based on Model Adaptive Selection Fusion, *Power System Technology*, Vol. 43, No. 6, pp. 1955-1961, June, 2019.
- [13] J. C. Yang, P. Jiang, G. Y. Chen, T. J. Yuan, F. Xue, Study on smart meters fault identification model, *Electrical and Energy Management Technology*, Vol. 58, No. 6, pp. 30-36, June, 2017.
- [14] Z. Q. Li, D. B. Gao, S. Su, J. C. Wang, D. D. Chen, X. J. Zeng, Big data based intrusion detection method of smart meters, *Journal of Electric Power Science and Technology*, Vol. 31, No. 1, pp. 121-126, March, 2016.
- [15] D. Xu, J. He, Z. Yang, Reliability prediction based on Birnbaum-Saunders model and its application to smart meter, *Annals of Operations Research*, Vol. 312, No. 1, pp. 519-532, May, 2022.
- [16] G. Henri, N. Lu, A Multi-Agent Shared Machine Learning Approach for Real-time Battery Operation Mode Prediction and Control, *2018 IEEE Power & Energy Society General Meeting (PESGM)*, Portland, OR, USA, 2018, pp. 1-5.
- [17] K. Song, L. Cui, A common random effect induced bivariate gamma degradation process with application to remaining useful life prediction, *Reliability engineering & system safety*, Vol. 219, Article No. 108200, March, 2022.
- [18] N. Cai, P. Li, Image compression method based on KPCA, *Radio Engineering*, Vol. 48, No. 12, pp 1061-1064, 2018.
- [19] Y. W. Zhao, B. Ma, B. S. Shen, J. J. Zhang, Study on application of state subspace for automatic alarm of reciprocating compressor, *Mechanical Science and Technology for Aerospace Engineering*, Vol. 35, No. 4, pp. 568-572, April, 2016.
- [20] D. P. Kingma, M. Welling, Auto-encoding variational Bayes, *International Conference on Learning Representations*, arXiv, December, 2013. <https://arxiv.org/abs/1312.6114>
- [21] K. Gregor, I. Danihelka, A. Graves, D. J. Rezende, D. Wierstra, DRAW: a recurrent neural network for image generation, *The 32nd International Conference on International Conference on Machine Learning*, Lille, France, 2015, pp. 1462-1471.
- [22] D. L. Tran, R. Walecki, O. Rudovic, S. Eleftheriadis, B. Schuller, M. Pantic, DeepCoder: semi-parametric variational autoencoders for automatic facial action coding, *IEEE International Conference on Computer Vision*, Venice, Italy, 2017, pp. 3209-3218.
- [23] D. F. Wang, J. Gu, VASC: dimension reduction and visualization of single-cell RNA-seq data by deep variational autoencoder, *Genomics, Proteomics and Bioinformatics*, Vol. 16, No. 5, pp. 320-331, October, 2018.
- [24] J. Na, K. Jeon, W. B. Lee, Toxic gas release modeling for real-time analysis using variational autoencoder with convolutional neural networks, *Chemical Engineering Science*, Vol. 181, pp. 68-78, May, 2018.
- [25] C. Du, B. Chen, B. Xu, D. Guo, H. Liu, Factorized discriminative conditional variational auto-encoder for radar HRRP target recognition, *Signal Processing*, Vol.

158, pp. 176-189, May, 2019.

- [26] G. W. Zhao, Y. X. Peng, Semisupervised SAR image change detection based on a siamese variational autoencoder, *Information Processing and Management*, Vol. 59, No. 1, pp. 1-16, January, 2022.
- [27] J. Yang, M. Y. Xin, J. X. Ou, J. R. Wang, Q. Song, Data accuracy judgment method of metrology automation system based on La Yida criterion, *Power Systems and Big Data*, Vol. 20, No. 11, pp. 74-78, November, 2017.

Biographies



Zhou Yang (Member, IEEE) received the B.Sc. and Ph.D. degrees from Beihang University, China, in 2008 and 2014, respectively. During his Ph.D. study, he was a Visiting Scholar with Rutgers University, New Brunswick, NJ, USA, from 2013 to 2014. From 2017 to 2019, he was a Post-Doctoral Fellow with Tsinghua

University. He is currently a Senior Engineer with China Southern Power Grid Company Ltd. (CSG). His research interests include smart metering, readability, and safety in the electricity domain.



Yuxiang Peng was born in Nanning, China, 1974. He received the B.Sc., and Master's degrees from the University of Huazhong University of Science and Technology and Wuhan University respectively. He is currently in CSG. (China Southern Power Grid Company Limited). He's research interests include AMI and Smart Grid

domain.



Wenqian Jiang was born in Hunan, China. She received the B.Sc., and MA.Sc degrees from the University of Hunan University, China in 2007 and 2010 respectively. She is currently pursuing Ph.D at Tianjin University, and a senior engineer in CSG. (China Southern Power Grid Company Limited). Wenqian Jiang's research

interests include AMI and Date analysis.



Juntao Pan was born in Wuzhou, China. He received the B.Sc., and MA.Sc degrees from the University of Hunan University, China in 2008 and 2011 respectively. He is currently a senior engineer in CSG. (China Southern Power Grid Company Limited). Juntao Pan's research interests include smart metering, readability and safety in

electricity domain.



Yanfu Li (Senior Member, IEEE) was a Faculty Member with the Laboratory of Industrial Engineering, CentraleSupélec, University of Paris-Saclay, France, from 2011 to 2016. He is currently a Professor with the Department of Industrial Engineering, Tsinghua University. He has led/participated in several projects

supported by EU, France, and Chinese Governmental funding agencies and various industrial partners. He has coauthored more than 100 publications in international journals, conference proceedings, and books. His current research interests include reliability, availability, maintainability, safety (RAMS) assessment, and optimization with the applications of various industrial systems. He is an Associate Editor of IEEE TRANSACTIONS ON RELIABILITY.