

Article

A Data-Driven Method for Typical Load Profile Extraction in Electricity Market User Profiling

Jing Yang ¹, Chao Pang ², Xin Luo ³, Yifan Lv ⁴, Jingjiao Li ^{4,*} and Ke Xu ⁴¹ State Grid Customer Service Center, Beijing 100062, China; zjuyangjing@126.com² State Grid Tianjin Electric Power Research Institute, Tianjin 300384, China; pc_tj_du@163.com³ Beijing Tsingsoft Technology Co., Ltd., Beijing 100085, China; xluo@tsingsoft.com.cn⁴ School of Electric Power Engineering, School of Shen Guorong, Nanjing Institute of Technology, Nanjing 211167, China; x00206191126@njit.edu.cn (Y.L.); xukesam@njit.edu.cn (K.X.)

* Correspondence: jjli@njit.edu.cn

Abstract

Accurate extraction of typical load curves (TLCs) is essential for electricity market trading, demand-side management, and optimal design of energy storage systems. However, conventional methods are highly sensitive to anomalous consumption days caused by equipment failures or maintenance, which can distort normal electricity consumption patterns. To address this issue, this paper proposes a two-stage unsupervised framework that integrates a deep sequence model with an anomaly detection algorithm for robust TLC extraction. First, a Transformer-based autoencoder is employed to learn complex temporal dependencies and intrinsic patterns from historical daily load data, extracting robust periodic features by reconstructing the input load sequences. Subsequently, the reconstruction error of each daily load curve is computed as an anomaly assessment metric. These reconstruction error features are then fed into an Isolation Forest algorithm to identify anomaly loads that significantly deviate from the learned normal patterns, without requiring predefined thresholds or labeled data. Validation using real-world commercial and industrial electricity consumption data demonstrates that the proposed method effectively filters out various anomalies (e.g., spikes, troughs, and shape distortions) that conventional methods fail to exclude. The extracted TLCs exhibit improved robustness and representativeness. Further case studies indicate that adopting purified TLCs to guide electricity procurement in market trading facilitates more scientific trading strategies and avoids increased electricity costs caused by distorted load patterns. In summary, the proposed Transformer-Isolation Forest hybrid framework provides an effective data-driven solution for robust TLC extraction. The resulting TLCs can be directly used to guide day-ahead market bidding, optimize power purchase contract decomposition, and assess user demand response potential.

Keywords: typical load profile; load data cleaning; Transformer; Isolation Forest; electricity market trading

Academic Editor: Juan-José González de la Rosa

Received: 15 May 2026

Revised: 19 June 2026

Accepted: 26 June 2026

Published: 28 June 2026

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

1. Introduction

The global transition towards digitalized electricity markets and the widespread deployment of Advanced Metering Infrastructure (AMI) have made high-temporal-resolution load data widely available. These data are essential for understanding

customer behavior, assessing demand-side flexibility, and supporting market operations. In this context, extracting stable, interpretable, and representative typical load curves from historical data has become a fundamental task for applications such as load management, demand response, distributed resource optimization, and customer segmentation [1–3]. In addition, representative typical load curves provide a crucial data foundation for short- and medium-term load forecasting. Recent studies indicate that AI-based forecasting models, including deep neural networks and Transformer architectures, can achieve high forecasting accuracy; however, they remain highly dependent on the quality and representativeness of historical training data [4]. Furthermore, abnormal operating conditions—such as equipment failures, maintenance cycles, and metering errors—can distort load patterns, introducing bias into forecasting models, thereby reducing forecasting performance and compromising model robustness [5]. Therefore, identifying and eliminating days with abnormal loads prior to extracting typical load curves not only benefits customer analysis but also helps improve the reliability of downstream forecasting applications.

However, the inherent complexity of industrial and commercial load patterns—driven by production rhythms, shift changes, weekend effects, seasonal variations, and external conditions—poses significant challenges due to their non-stationarity, multi-scale periodicity, and morphological drift [6,7]. Recent work on Transformer-based load analysis further confirms that spatiotemporal variations in load sequences affect model generalization, while industrial power time series exhibit strong periodicity alongside non-stationary behavior [8].

The current research landscape is dominated by classical clustering methods, particularly K-means and K-shape, owing to their simplicity, computational efficiency, and interpretability. Nevertheless, these methods face substantial limitations when applied to highly volatile or complex load patterns. They struggle to capture temporal variations and non-linear similarities, and their outputs are highly sensitive to distance metrics, exogenous variable embeddings, and evaluation criteria. Consequently, centroid shift, shape distortion, and lack of robustness are common [9–11]. To overcome these issues, deep clustering approaches based on autoencoders (AEs) and variational autoencoders (VAEs) have been proposed. These methods improve representation quality by learning nonlinear feature compressions in a latent space [12,13]. However, their optimization objectives remain cluster-oriented, making it difficult to exclude the influence of atypical days—such as equipment failures, or load peaks/troughs—on the final typical curves [14]. Moreover, static reconstruction or single-stage clustering frameworks show limited robustness when load states evolve over time [13,14].

In recent years, advancements in the fields of load analysis and time series anomaly detection have increasingly relied on deep learning models. Early research primarily utilized recurrent neural networks (RNNs), long short-term memory networks (LSTMs), and gated recurrent units (GRUs) to capture temporal dependencies in load sequences. Although these models have demonstrated satisfactory predictive performance, their sequence processing mechanisms limit their ability to capture long-range dependencies and complex multi-periodic patterns in high-resolution load data [15]. Recent studies have shown that Transformer-based architectures outperform traditional methods such as RNNs in many power system applications. This is attributed to their self-attention mechanism, which efficiently models long-term temporal relationships and supports parallel computation [16], and can be combined with mechanisms such as autoencoders and density modeling to enhance the representation of complex power time series [17,18]. However, in scenarios characterized by strong periodicity, non-stationarity, and diverse anomaly patterns, single-stage Transformers or general-purpose autoencoders still

struggle to balance the fidelity of normal fluctuations with the sensitivity of anomaly detection [19].

On the other hand, unsupervised statistical methods such as Isolation Forest and the 3σ rule are widely used for outlier screening due to their simplicity and label-free nature [20,21]. Yet, when applied directly to raw high-dimensional load curves, they may incorrectly flag physically reasonable high-volatility days (e.g., routine production adjustments) as anomalies. A more robust alternative is to first characterize the normal evolutionary trend of loads and then detect deviations in the residual space [22].

Despite these advances, existing research has focused either on anomaly detection itself or on load forecasting performance. Relatively little attention has been paid to how anomaly removal affects the quality and representativeness of the extracted typical load curves. Furthermore, although hybrid frameworks combining deep representation learning and anomaly detection algorithms have been studied, their application in robust typical load curve extraction for power market user analysis remains limited. Recent reviews on demand-side anomaly detection in the power sector have highlighted the urgent need to develop hybrid frameworks capable of simultaneously addressing representation learning, unsupervised applicability, engineering deployment feasibility, and operational explainability [23,24]. Therefore, there remains a need for a dedicated framework that integrates temporal representation learning, anomaly cleansing, and typical load curve extraction.

Based on this, this paper proposes a two-stage “purification-first, extraction-second” framework for robust typical load curve (TLC) extraction. The principal conclusions are as follows: First, a Transformer-based autoencoder learns the normal evolutionary trend of daily load sequences, and reconstruction residuals are fed into an Isolation Forest model to automatically identify and filter out atypical days (equipment failure days, and peak/trough anomaly days). Second, following a “purify first, extract later” principle, robust typical load curves are derived from the purified sample set. Using real-world industrial and commercial load data, we systematically evaluate anomaly detection performance (Precision, Recall, F1-score) and compare against USAD, AST, standalone Isolation Forest and the 3σ rule. Experimental results demonstrate that the proposed framework significantly enhances the robustness of typical load curve extraction, thereby validating its practical value for electricity market customer analysis.

2. The Proposed Transformer-IF Model

2.1. Transformer-Based Trend Reconstruction of Load Sequences

Industrial and commercial user load sequences exhibit significant non-stationarity and multi-periodicity characteristics (e.g., daily and weekly cycles). To accurately reconstruct normal electricity consumption patterns, this paper employs the encoder-decoder architecture of the Transformer for feature learning.

2.1.1. Input Embedding and Positional Encoding

First, the raw daily load sequence X is mapped into a high-dimensional feature space, and positional encoding is introduced to preserve the sequential order information between time steps. The input data is a time series of dimension (t, d) , where t is the number of time steps and d is the number of features. A linear transformation is first applied as follows:

$$E = W_E X + b_E \quad (1)$$

where $W_E \in R^{d_{\text{model}} \times d}$ is the weight matrix and b_E is the bias term.

Positional encoding is then introduced to retain temporal order information:

$$PE(t, 2i) = \sin\left(\frac{t}{10000^{2i/d_{\text{model}}}}\right), \quad PE(t, 2i+1) = \cos\left(\frac{t}{10000^{2i/d_{\text{model}}}}\right) \quad (2)$$

Finally, the input representation is given by:

$$Z = E + PE \quad (3)$$

2.1.2. Encoder and Self-Attention Mechanism

The encoder captures global dependencies between different time steps in the input sequence through a multi-head self-attention mechanism, extracting cross-temporal coupling features. The attention mechanism is formulated as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

where Q , K , and V denote the Query, Key, and Value matrices, respectively, and the softmax operation preceded by scaling (division by $\sqrt{d_k}$) to stabilize the gradient yields the normalization factor for the attention head dimension. The computation of multiple attention heads is performed via linear transformations followed by concatenation and fusion. The feed-forward network (FFN) within the encoder consists of two fully connected layers, computed as follows:

$$\text{FFN}(X) = \max(0, XW_1 + b_1)W_2 + b_2 \quad (5)$$

where W_1 and W_2 are learnable weight matrices, b_1 and b_2 are bias terms, and the ReLU activation function is employed to enhance non-linear expressiveness. The entire encoder incorporates residual connections and layer normalization to ensure stable gradient propagation and improve model training effectiveness.

After reconstruction through the encoder and decoder, the model outputs an approximate representation of the normal evolutionary trend of load sequences. Since the model primarily learns stable patterns that frequently appear in historical samples, it achieves good fitting capability for normal periodic fluctuations, gradual ramping, and regular peak-valley variations. Conversely, abnormal load mutation data, which do not conform to the dominant distribution pattern, are typically difficult for the model to reconstruct accurately. Consequently, the residuals formed between the original and reconstructed sequences can effectively amplify anomalous information, providing a purer and more sensitive input space for the subsequent iForest (Isolation Forest) anomaly identification.

2.1.3. Decoder and Trend Generation

Decoder: As illustrated in Figure 1, the decoder architecture is largely consistent with that of the encoder, consisting of multi-head self-attention (MHSA) layers, encoder–decoder attention layers, and a feedforward neural network. The goal of the Transformer autoencoder is to reconstruct the complete input payload sequence, enabling the model to utilize contextual information from the entire sequence and improve the quality of reconstruction of normal payload evolution patterns. In the encoder–decoder attention module, the query is generated by the decoder, while the key and value are derived from the encoder’s output. Finally, the output is obtained through a fully connected layer and softmax normalization.

Output Layer: A linear transformation is applied to the decoded features to recover the time series:

$$\hat{Y} = Z_{\text{dec}} W_{\text{out}} + b_{\text{out}} \quad (6)$$

where W_{out} and b_{out} are learnable parameters, and $\hat{Y}$ represents the reconstructed (predicted) values [13].

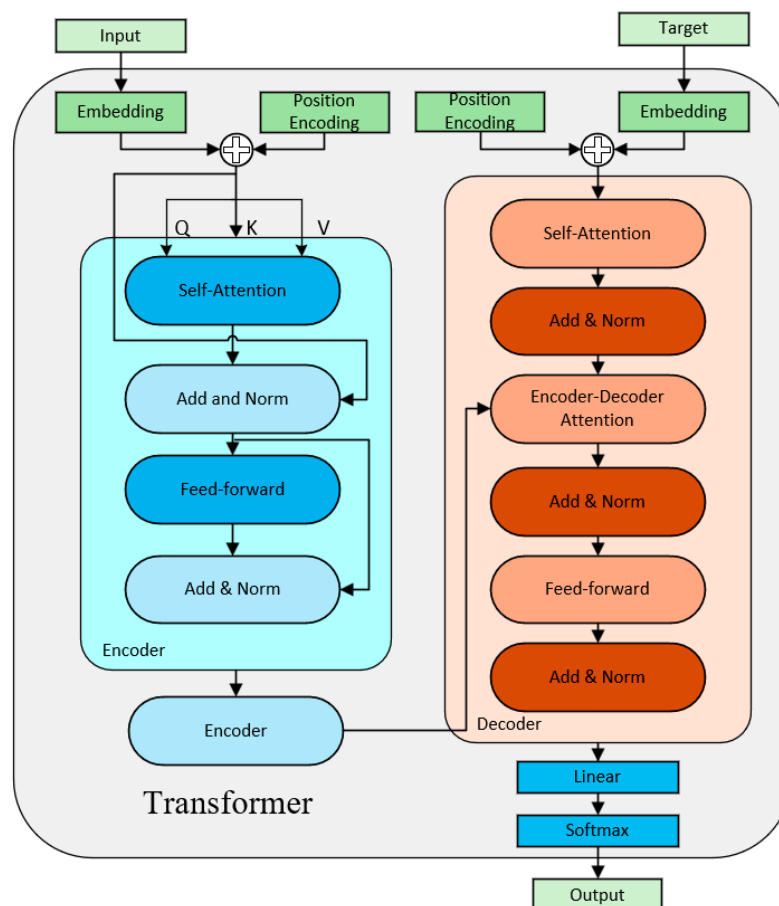


Figure 1. Transformer architecture.

2.2. Residual Outlier Detection Based on an Improved Isolation Forest

After obtaining the reconstructed sequence $\hat{X}$, the residual sequence E between the original sequence and the reconstructed sequence is calculated as follows:

$$E = |X - \hat{X}| [cite, tart] \quad (7)$$

Residuals of normal samples typically exhibit small fluctuations around zero, while anomalies manifest significant sparsity and isolation in the residual space.

The standardized residual features are fed into the Isolation Forest model. This algorithm constructs multiple iTrees (Isolation Trees) by randomly partitioning the feature space, and defines the anomaly score $s(x, n)$ using the average path length $E(h(x))$ of a sample in the trees:

$$s(x, n) = 2^{-E(h(x))/c(n)} \quad (8)$$

When $s(x, n)$ is close to 1, the average path length of the data point in the Isolation Forest is shorter, indicating that it is more sparsely distributed in the data space and thus more likely to be an anomaly. Conversely, when $s(x, n)$ approaches around 0.5, the data point requires more splitting steps to be isolated, indicating that it is a normal point [16]. This study incorporates the reconstruction error computed by the Transformer:

$$Residual = Y - \hat{Y} \quad (9)$$

The normalized result $Z_{Residual} = (Residual - \mu) / \sigma$ is used as the input to the Isolation Forest. Finally, anomaly identification is performed based on a pre-defined anomaly score threshold (e.g., the 95th percentile). Compared with traditional anomaly detection methods, the Isolation Forest has lower computational complexity in high-dimensional feature spaces, effectively identifies anomalous patterns in power load data, and automates load anomaly detection and identification.

2.3. Extraction Process of Typical Load Curve

After the Transformer encoder–decoder completes the reconstruction of the load sequence, the discrepancy between the original load sequence and the reconstructed result can effectively amplify the characteristics of anomalous days, providing a sensitive input for anomaly detection. Based on this property, this study nests the encoder–decoder with the Isolation Forest model to achieve robust extraction of typical load curves.

Specifically, each daily load sequence is first fed into the encoder to extract high-dimensional latent features, capturing long-term dependencies and multi-scale periodic variations in the sequence. Subsequently, the decoder reconstructs the load sequence based on these latent features, preserving regular intra-day and intra-week fluctuations while maintaining low sensitivity to abnormal peaks or troughs. By calculating the residuals between the original and reconstructed sequences, a feature sequence that reflects anomalous deviations is formed. After feeding this residual sequence into the Isolation Forest, the model can automatically identify anomalous days based on the isolation degree of samples in the feature space, thereby effectively removing abnormal load data.

After the removal of anomalous days, the remaining purified sample set retains normal periodicities and reasonable local fluctuation characteristics, from which typical load curves can be further extracted via mean calculation or clustering analysis. This nested structure realizes a closed-loop process of “trend learning—anomaly removal—typical extraction,” effectively improving the representativeness and robustness of typical load curves, and providing reliable data support for downstream electricity market trading strategies and load management.

The final representative load curve is obtained by calculating the element-wise averages of all purification day sequences. Specifically, after removing outlier days, the remaining N purification day sequences are averaged at each time point t :

$$TLC(t) = \frac{1}{N} \times \sum_{i=1}^N purified_sequence_i(t) \quad (10)$$

Here, $TCL(t)$ represents the typical load value at time t , N is the number of valid days remaining after outliers are removed, and $purified_sequence_i(t)$ represents the load value at time t for the i -th purified daily sequence. This averaging operation provides a statistically robust representative curve because the purified sample set, after removing outlier days, is highly homogeneous; the averaging operation effectively captures the most frequently occurring normal electricity consumption patterns within the evaluation period.

The flowchart of the proposed Transformer-IF model is illustrated in Figure 2. In Figure 2, the color-coded shapes are as follows: light green parallelograms for inputs/outputs, light pink rectangles for computational steps, light peach diamonds for logical decisions, and orange rectangles for exception handling.

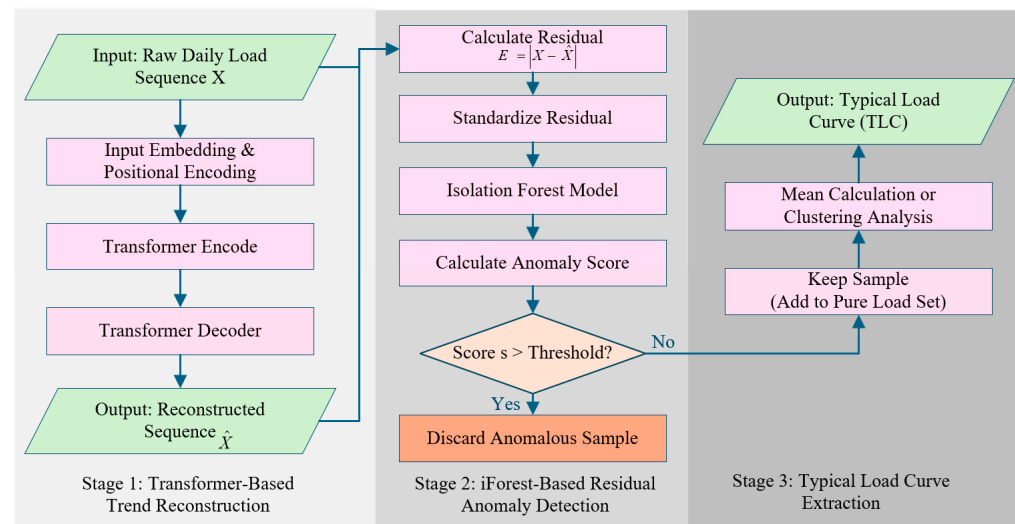


Figure 2. Flow chart.

3. Performance Evaluation Index

To assess the accuracy of the proposed model in detecting and predicting daily user load patterns, this study employs precision, recall, and the F1 score as evaluation metrics. These indicators are defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

In the above formulations, TP (True Positive) denotes the number of samples correctly identified as anomalous, i.e., samples for which both the predicted anomaly label and the ground-truth label are 1. TN (True Negative) represents the number of samples correctly identified as normal, i.e., samples with both predicted and ground-truth labels equal to 0. FP (False Positive) refers to the number of falsely detected samples, specifically normal samples that are misclassified as anomalous. FN (False Negative) indicates the number of missed detections, i.e., anomalous samples that fail to be identified.

Precision measures the accuracy of the model in detecting positive samples, with higher values indicating better performance. Recall evaluates the model's ability to comprehensively identify positive samples, where larger values are preferable. Since precision and recall are mutually constrained, the F1 score is introduced as a comprehensive indicator that balances both metrics; a higher F1 score reflects superior overall detection performance.

The calculation of electricity procurement costs is based on a time-of-use pricing mechanism, and the specific formula is as follows:

$$\text{Cost} = \sum_{t=1}^T (P_t \times L_t) + \text{CapacityCharge} \quad (14)$$

Here, P_t represents the time-of-use electricity price for time period t , L_t represents the typical load value for the corresponding time period, and *CapacityCharge* represents the fixed capacity charge calculated based on transformer capacity or maximum demand.

In the day-ahead market, generators submit their generation volumes and bids for each time period, while consumers submit only their electricity demand without specifying a price, accepting the system's cleared price; the cleared price is calculated as follows:

$$P = \sum_{cost}^{gen} / Q_{total} \quad (15)$$

This is calculated by dividing the total generation cost by the total load. This simplified model does not account for the impact of transmission line congestion, and the clearing price reflects the average marginal generation cost across the entire grid.

4. Case Analysis

4.1. Test Platform and Data Sources

This study uses an Intel Core i7-12800HX CPU (2.60 GHz), 16 GB RAM, and an NVIDIA RTX 4070 (4 GB VRAM), with software built on PyTorch 2.1.0, NumPy 1.24.3, Pandas 2.0.2, Scikit-learn 1.3.0, and Matplotlib 3.7.1. To validate the proposed framework for Typical Load Curve (TLC) extraction, four representative industrial and commercial power users—covering chemicals, steel, and electrical equipment manufacturing—are selected. The dataset contains 15-min load data (96 daily samples) from 2021.01.01 to 2023.12.31, covering seasonal, holiday, and production patterns, with typical non-physical abrupt rises and drops. Ground-truth labels are generated via rule-based labeling plus manual verification, targeting missing, duplicate, and mutation anomalies, to assess the model's anomaly cleaning capability.

4.2. Model Parameters and Experimental Configuration

The accuracy of TLC extraction is directly affected by the parameters of the Transformer reconstruction model and the iForest detector. Training uses the Adam optimizer ($\text{lr} = 0.001$) and MSELoss to minimize reconstruction error. Input sequences are constructed via sliding windows; see Table 1 for parameter configurations.

Table 1. Hyperparameter configuration of the Transformer-iForest framework.

Module	Parameter	Value	Description
Transformer	Sliding window length	30	Employs a 30-point sliding window to capture local dependencies via self-attention, balancing local context for anomaly detection without over-smoothing. Sliding 96 times reconstructs the full 96-point daily load curve.
	Embedding dimension d_{model}	64	Linear transformation into a 64-dimensional space
	Number of encoder/decoder layers	3	3-layer Encoder–Decoder structure
	Number of attention heads	4	Employs Multi-head attention for capturing long-term dependencies
iForest	Contamination rate	0.05	Assumes 5% of anomalous days in raw load data
	Number of trees	100	Default value for $n_{\text{estimators}}$ for isolation trees
	Random seed	42	Ensures experimental reproducibility

4.3. Performance Evaluation and Comparison of Typical Load Curve Extraction

4.3.1. Threshold Sensitivity Analysis and Adaptive Mechanism

Figure 3 illustrates the effect of different threshold quantiles on anomaly removal for four types of industrial and commercial users. The threshold quantile directly affects the detection performance of mutation anomalies. As the threshold quantile increases, the reference anomaly set, constructed from the reconstruction errors of the Transformer

autoencoder, decreases in size, leading to improved recall of mutation anomalies until saturation. However, higher thresholds may misclassify some normal days as anomalies, causing a decline in precision in the high-threshold range. When the threshold is excessively high, deviations between model outputs and ground-truth anomaly labels occur, resulting in noticeable fluctuations in overall anomaly detection performance. The F1 score exhibits a unimodal characteristic, indicating a stable optimal trade-off between precision and recall.

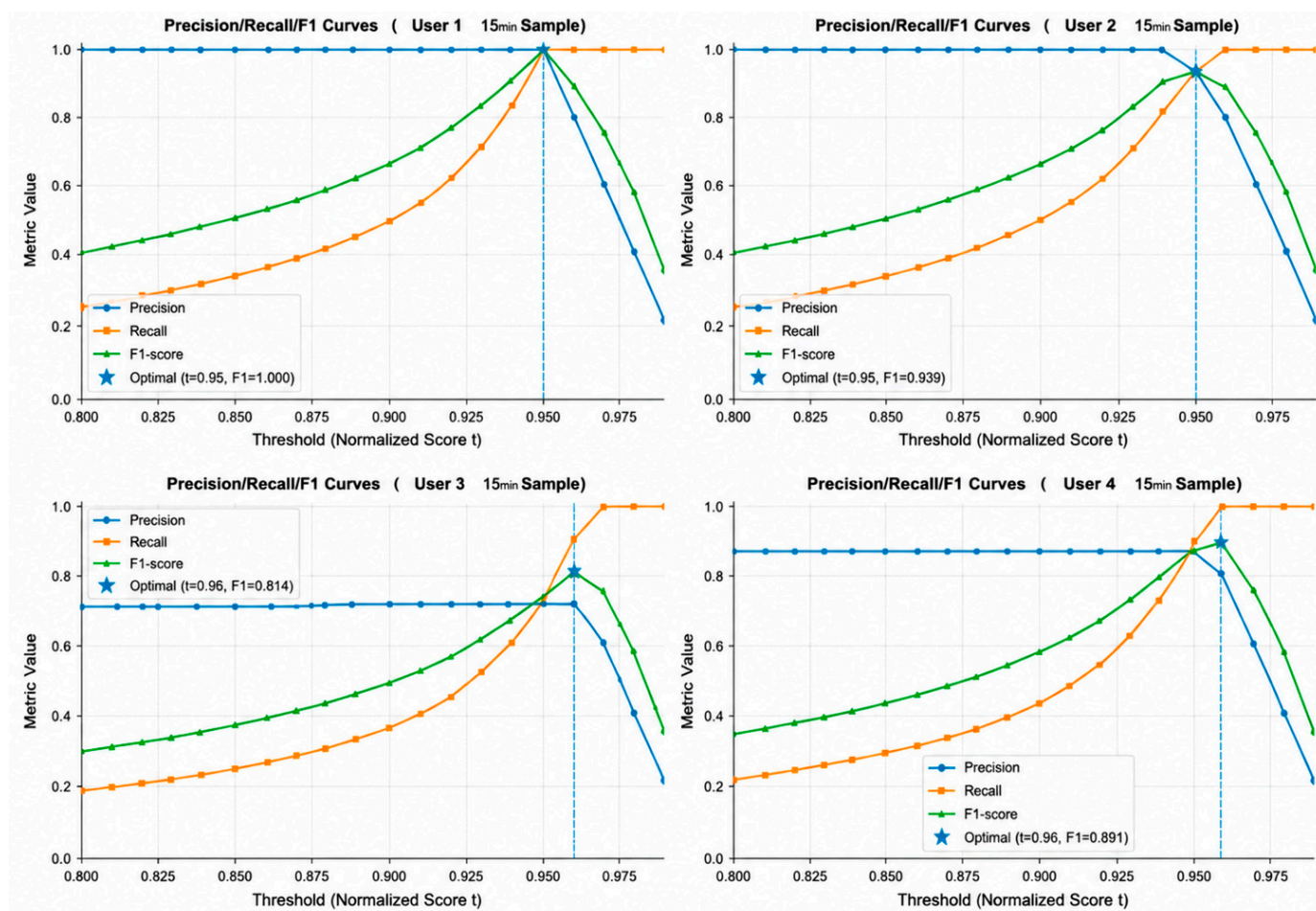


Figure 3. Anomaly removal performance under different threshold quantiles for each user.

Specifically, the optimal thresholds for all four users lie in relatively high quantile ranges. User 1 (chemicals) achieves the best performance with an F1 score close to 1 at approximately the 0.95 quantile. User 2 (another chemical plant) reaches its optimum at around the 0.939 quantile. User 3 (steel) attains an F1 score of approximately 0.814 at the 0.96 quantile. User 4 (electrical equipment) achieves an F1 score of about 0.891 in a similar quantile range.

In response to the phenomenon of varying optimal thresholds across different users, this paper further elucidates the threshold determination mechanism. Although the proposed framework does not require labeled data during the training phase, the anomaly detection in Isolation Forest fundamentally relies on a quantitative criterion—the quantile threshold—which is determined by the statistical distribution characteristics of residual features rather than being manually preset. For new user scenarios, this paper validates the adaptive quantile threshold determination method through threshold sensitivity analysis experiments: in tests across all four different industry users, quantile = 0.95 consistently demonstrated robust detection performance, indicating that this threshold is

universally applicable across users. For the need for fine-tuning for specific users, adaptive threshold determination can be achieved by reconstructing the error distribution characteristics through cross-validation, without the need for additional labeled data.

These experimental results demonstrate that the trend reconstruction features extracted by the Transformer autoencoder, combined with the isolation forest for anomaly detection, can achieve robust mutation anomaly removal under different load levels while maintaining threshold tunability and sensitivity.

To evaluate the performance of the proposed Transformer-IF model in handling anomalous days in industrial and commercial load data, we compared it with USAD (UnSupervised Anomaly Detection on Multivariate Time Series) [25], AST (Adversarial Spatial-Temporal network) [26], isolation forest, and the conventional 3σ rule. To ensure fair comparison: (1) All models were trained, validated, and tested on identical data splits. (2) Hyperparameters were tuned via grid search using the same validation set and evaluation metrics as in Table 1, with optimal configurations listed in Table 2. (3) All baselines were applied to the same raw 96-point load input space. The performance comparison results are presented in Table 3.

Table 2. Hyperparameter configuration of the comparative methods.

Module	Parameter	Value	Description
USAD	encoder/decoder layers	2/2	Number of stacked layers in the encoder and decoder respectively
	hidden_dim	64	Hidden dimension controlling the internal feature representation size
	learning_rate	0.001	Learning rate for parameter updates
AST	attention_heads	4	Number of attention heads for multi-subspace feature extraction
	learning_rate	0.001	Learning rate for parameter updates
isolation forest	Contamination rate	0.05	Assumes 5% of anomalous days in raw load data
	Number of trees	100	Default value for $n_{\text{estimators}}$ for isolation trees
	Random seed	42	Ensures experimental reproducibility
The 3σ rule	Threshold	3σ	Anomaly threshold based on normal distribution, flagging points beyond $\pm 3\sigma$ as anomalous

Table 3. Performance comparison of anomaly identification and cleaning.

Method	Precision	Recall	F1
The proposed Transformer-IF model	0.92	0.87	0.8943
USAD	0.88	0.85	0.8647
AST	0.86	0.83	0.8447
isolation forest	0.7651	0.7237	0.7438
The 3σ rule	0.6234	0.6811	0.6510

Note: Values are averages across four users.

Table 3 reports the proposed model achieves the highest precision (0.92), recall (0.87), and F1 score (0.8943). This superiority is attributed to the accurate reconstruction of multi-cycle load trends by the Transformer autoencoder and the sensitive detection of residual features by the isolation forest, enabling effective removal of abrupt anomalies while preserving the periodicity and local fluctuations of normal loads. In comparison, USAD and AST yield lower recall under highly volatile or complex load patterns; isolation forest alone suffers from low precision; and the 3σ rule is constrained by its statistical assumptions, exhibiting the weakest overall performance. In summary, the Transformer-IF framework demonstrates higher robustness and representativeness in anomaly

removal and load cleaning, providing a high-quality data foundation for subsequent typical load curve extraction.

As shown in Table 4, the table quantifies the detection performance and computational efficiency of the proposed framework. In terms of reconstruction quality, the average RMSE for users in the chemical industry was only 0.0165, with an MAE of 0.0111, indicating that the proposed method can accurately reconstruct load curve patterns under normal operating conditions. For users in the steel industry, where load fluctuations are significant, the RMSE was 0.2564, MAE of 0.1819. This value is slightly higher than that of other users, primarily due to the large and frequent load fluctuations within the steel industry's production cycle, which is an inherent characteristic of the industry rather than a model flaw. In contrast, the traditional 3σ rule has a false positive rate of over 40% for such normal production fluctuations, whereas the method proposed in this paper, after learning the normal pattern via a Transformer, only classifies days with significant deviations from the normal range as abnormal, effectively distinguishing between normal fluctuations and abnormal conditions. In terms of computational efficiency, the average training time per user is 104.65 s, and the inference time is less than 2 s, fully meeting the computational requirements for real-time market clearing and intraday scheduling scenarios. Although the training time for the chemical industry user is longer, this is due to the larger dataset size and longer evaluation period, requiring the model to learn more complex time-series patterns—a normal training overhead. Overall, the proposed framework achieves a good balance between detection accuracy and computational cost, providing a viable technical solution for real-time applications in the power market.

Table 4. Statistics on reconstruction quality and computational efficiency for each user.

ID	RMSE	MAE	Train_Time (s)	Inf_Time (s)	Anomalies
User 1	0.0165	0.0111	188.69	1.6650	68
User 2	0.2572	0.1861	64.78	1.5758	63
User 3	0.2564	0.1819	69.36	1.6490	63
User 4	0.0162	0.0135	95.77	1.6162	64
Average value	0.1366	0.0981	104.65	1.6265	64.5

4.3.2. Analysis of the Cleaning Effect of Typical User Loads

Figure 4 presents the typical load curve purification results for User 1 (chemical industry) during the evaluation period from 16 March 2021, to 28 March 2023. Subplot (a) shows the original load data sequence, where 79 anomaly points marked with red $\times$ are mainly distributed during March 19–21, exhibiting non-physical load step drops. The zoomed inset in the lower-right corner further illustrates the bi-modal daytime peaks of the normal operating load curve, confirming the user's typical “double-peak daytime, nighttime trough” consumption pattern. Subplot (b) demonstrates the trend reconstruction capability of the Transformer autoencoder. The reconstructed curve exhibits high fitting accuracy for the bi-modal daily pattern under normal operating conditions, while showing significant deviations during anomalous periods, which provides an effective basis for subsequent residual-space anomaly detection. Subplot (c) presents the purified typical load curve after Isolation Forest filtering. The green solid curve completely preserves the normal bi-modal daily periodicity and local smooth fluctuations while effectively eliminating the non-physical anomalies caused by metering failures. The purified typical load curve can serve as a high-quality benchmark input for electricity market trading and load forecasting.

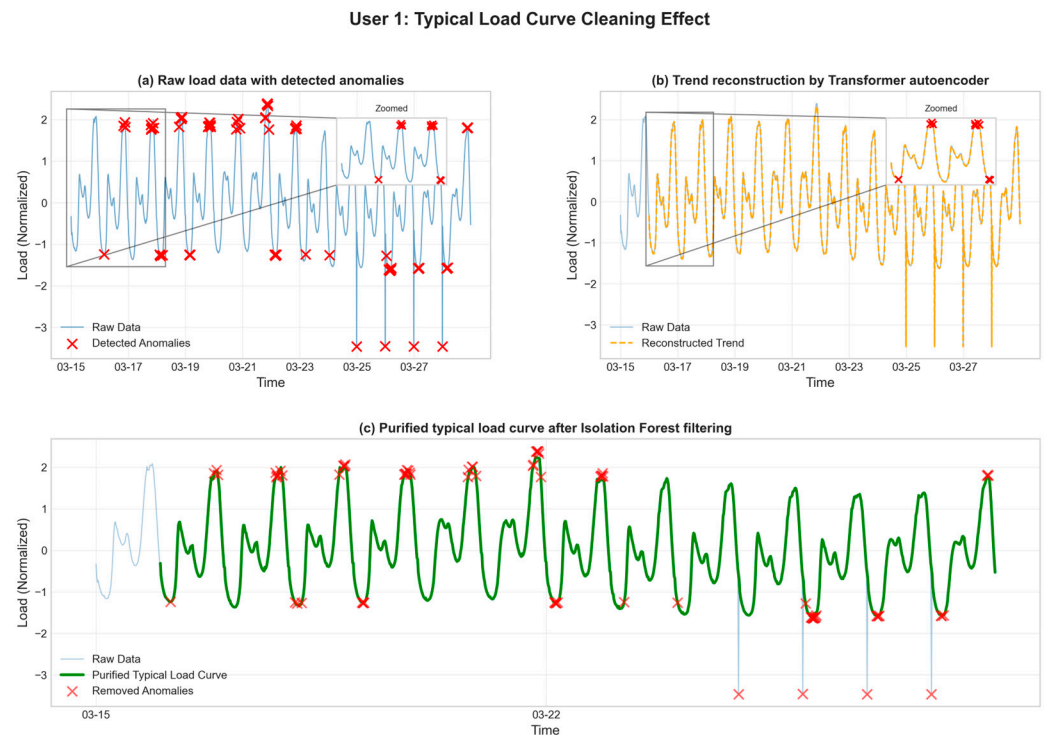


Figure 4. Typical load curve cleaning effect for User 1.

Figure 5 presents the typical load curve purification results for User 2 (also in the chemical industry) during the evaluation period from 16 March to 28 March 2022. Subplot (a) shows the original load data sequence, which exhibits a highly volatile and discontinuous load pattern. The 85 anomaly points marked with red $\times$ are mainly composed of sustained load deviations caused by equipment maintenance and planned production reductions, as well as instantaneous load spikes. The zoomed inset in the lower-right corner illustrates the user's typical uni-modal daily load pattern, with the peak occurring at midday and gradually declining during nighttime hours. Subplot (b) demonstrates the trend reconstruction capability of the Transformer autoencoder. The reconstructed curve exhibits high fitting accuracy for the multi-cycle production rhythm under normal operating conditions, effectively capturing the daily uni-modal fluctuation characteristics, while showing significant deviations during anomalous periods, which provides an effective basis for subsequent residual-space anomaly detection. Subplot (c) presents the purified typical load curve after Isolation Forest filtering. The green solid curve completely preserves the normal intra-day fluctuation characteristics within the production cycle while effectively eliminating the non-physical anomalies caused by equipment maintenance and instantaneous failures.

Figure 6 presents the typical load curve purification results for User 3 (steel industry) during the evaluation period from March 16 to March 28, 2023. Starting from March 18, this user implemented production adjustments with significantly reduced load levels, while daily fluctuations persisted at the lower level. Subplot (a) shows the original load data with 41 anomaly points marked in red $\times$, mainly exhibiting non-physical load drops. The zoomed inset illustrates the typical flattened daily load pattern of this energy-intensive industry. Subplot (b) demonstrates that using the Transformer autoencoder alone successfully captures trend and seasonal components, but tends to overly smooth the pre-adjustment fluctuations, leaving some spikes and troughs incompletely removed. Subplot (c) presents the purified typical load curve (green) after integrating the Transformer autoencoder with Isolation Forest. It effectively eliminates anomalies while preserving reasonable intraday fluctuations. This nested structure significantly improves

anomaly detection accuracy and ensures the continuity and periodicity of the purified curve.

User 2: Typical Load Curve Cleaning Effect

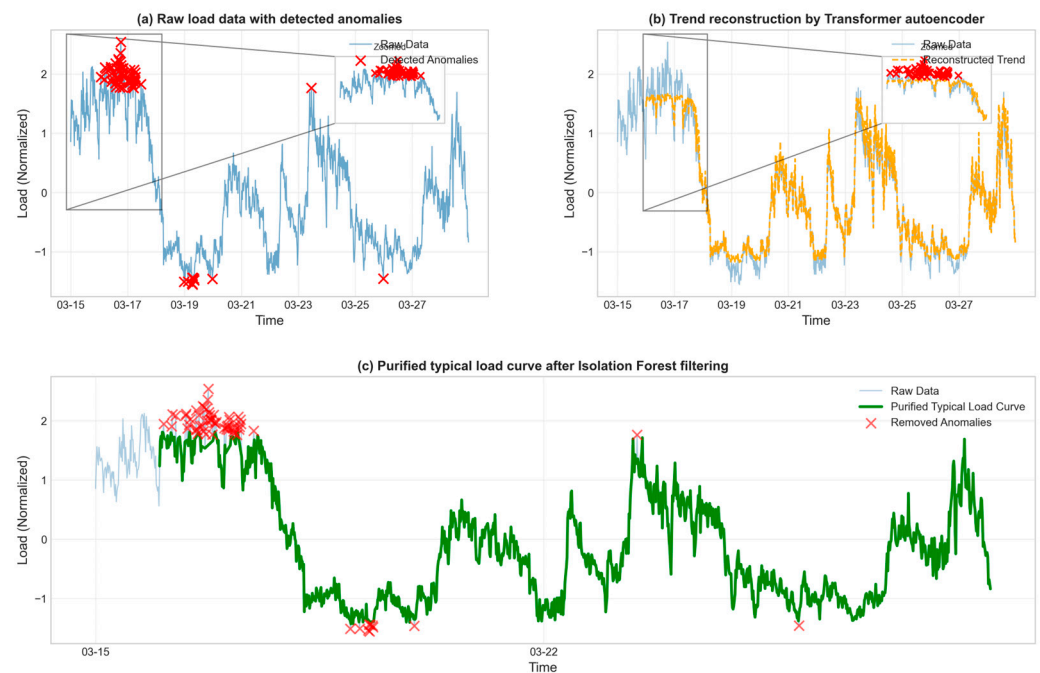


Figure 5. Typical load curve cleaning effect for User 2.

User 3: Typical Load Curve Cleaning Effect

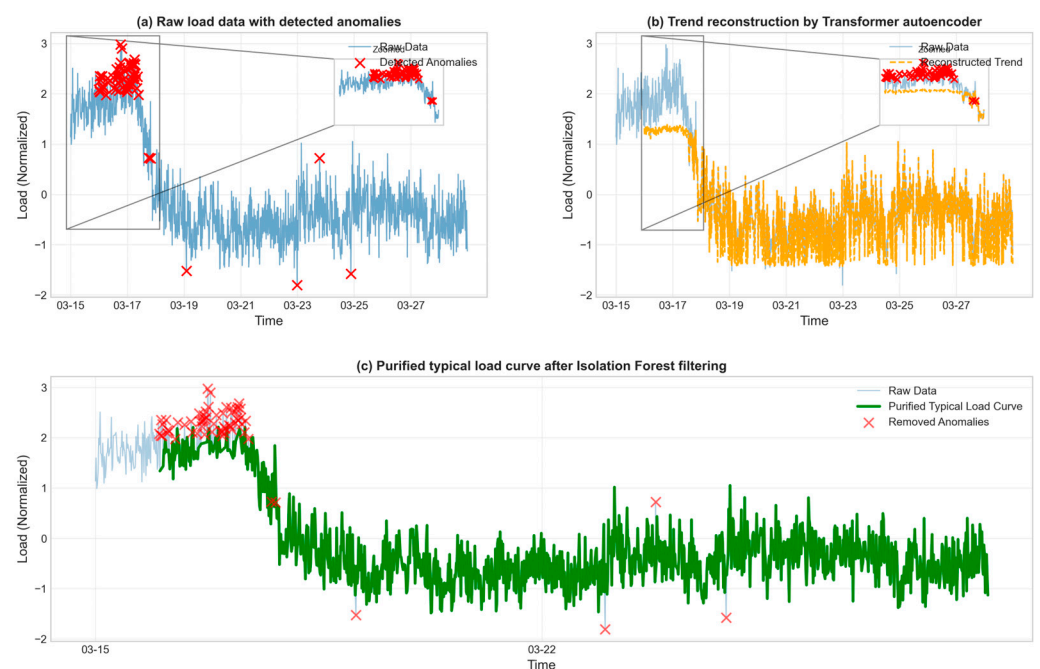


Figure 6. Typical load curve cleaning effect for User 3.

Figure 7 presents the typical load curve purification results for User 4 (electrical equipment manufacturing) during March 16–28, 2023. The load level remained stable, exhibiting a bi-modal daily pattern with considerable intraday fluctuations. Subplot (a) shows the original load data with 52 anomaly points marked in red $\times$, exhibiting local non-

physical steps and low-density outliers. Subplot (b) demonstrates the Transformer's high reconstruction accuracy under normal conditions with significant deviations during anomalies. Subplot (c) presents the purified curve (green) after Isolation Forest filtering, which effectively removes anomalies while preserving reasonable fluctuations. This validates that the Transformer-IF framework achieves stable and interpretable anomaly removal even in high-volatility industrial load scenarios.

User 4: Typical Load Curve Cleaning Effect

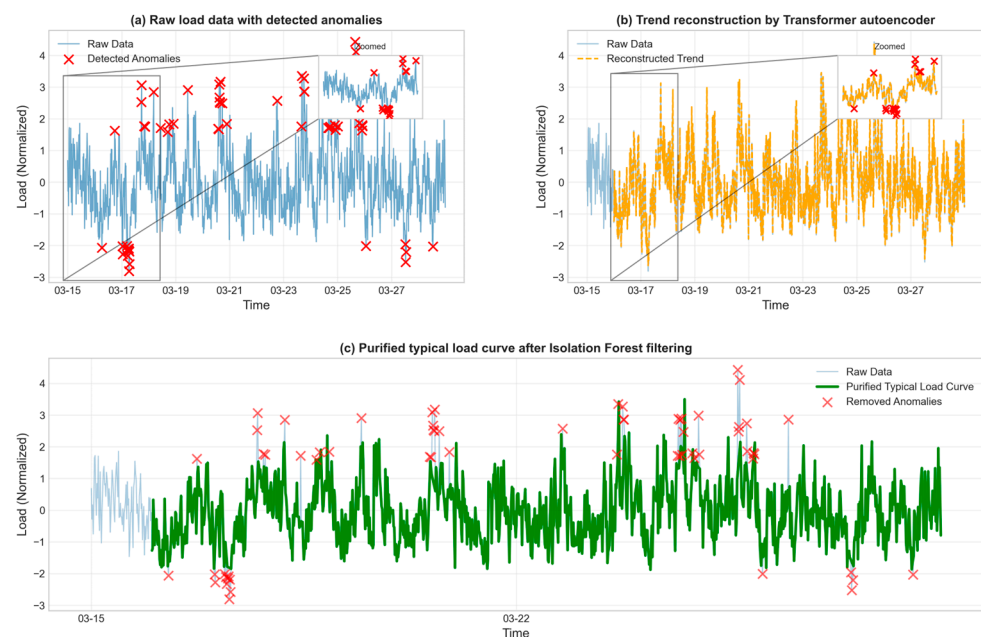


Figure 7. Typical load curve cleaning effect for User 4.

4.3.3. Comparison of Costs Before and After Correction

Figure 8 presents a comparative analysis of electricity procurement costs in a power market based on an IEEE 14-node system. This experiment simulates an abnormal scenario in which an industrial user's electricity meter malfunctions between 10:00 and 12:00, resulting in the load being overestimated at 1.8 times its actual value. Based on this scenario, the market settlement results for the original load curve were compared with those for the cleaned typical load curve. As shown in Figure 8a, the original load curve exhibited significant deviations during the failure period, whereas the cleaned typical load curve effectively filtered out these outliers. As shown in Figure 8b, during the low-load period at night (00:00–06:00), the clearing price remained around 240–260 yuan/MWh. As system demand increased during daytime operating hours, additional thermal power generation was dispatched, causing the clearing price to rise gradually. The highest price occurred during the peak demand period, reaching approximately 460 yuan/MWh. Figures 8c,d illustrate the economic impact of load data cleansing. Because the raw load curve overestimated electricity demand, this resulted in higher procurement costs and unnecessary market expenditures. After applying the purified TLC, electricity procurement costs were reduced by 54,339 yuan, a decrease of 14.7%.

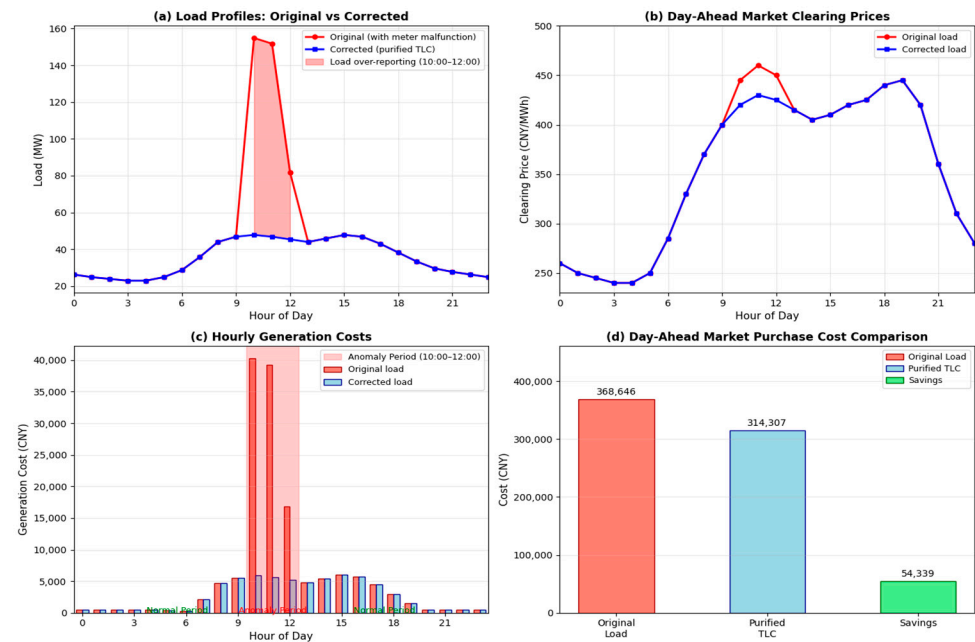


Figure 8. Comparison of electricity procurement costs based on typical load curves before and after purification.

5. Discussion

Experimental results demonstrate that the proposed Transformer-IF framework effectively addresses various types of anomalies in industrial and commercial user load data, significantly improving the robustness of typical load curve (TLC) extraction. Compared with baseline methods such as USAD, AST, standalone Isolation Forest, and the traditional 3σ rule, Transformer-IF achieves superior performance in terms of precision, recall, and F1-score. This advantage primarily stems from the complementary mechanism of its two-stage structure: The Transformer-based autoencoder captures multi-scale temporal dependencies in load data and reconstructs the main evolution trend of daily load, while the subsequent Isolation Forest accurately identifies anomalies based on reconstruction residual features, enabling fine-grained removal.

The proposed framework demonstrates good generalizability across different industrial application scenarios. The four industrial users selected in the experiments cover three typical industries: chemical, steel, and electrical equipment manufacturing. These industries exhibit significant differences in production cycles, electricity consumption patterns, and load fluctuation characteristics. The results show that through a user-specific threshold tuning strategy, the framework can effectively adapt to these inter-industry differences: the optimal threshold is approximately 0.95 for the chemical user, 0.96 for the steel user, and 0.939 for the electrical equipment manufacturing user, all falling within a similar reasonable range. This observation indicates that the core anomaly detection mechanism of the framework possesses universal detection capability across different industry load patterns, while the threshold differences primarily reflect inherent distinctions in load fluctuation characteristics across industries, rather than insufficient generalizability of the framework itself.

Threshold sensitivity analysis reveals that a higher quantile threshold enhances the coverage of abrupt anomalies but may lead to a slight decrease in precision, as some normal micro-fluctuations are misclassified as anomalies. Nevertheless, the optimal threshold intervals across different users are relatively consistent, indicating that the proposed framework exhibits good robustness to parameter settings. Moreover, the purified load sequences retain daily periodicity and fluctuation characteristics, ensuring

that the extracted typical load curves faithfully reflect electricity consumption behavior while avoiding the loss of critical information due to oversmoothing.

To verify the stability of the proposed Transformer-IF framework, this paper conducted a parameter sensitivity analysis. First, tests were performed on the contamination parameter of the Isolation Forest. The results showed that when this parameter varied within the range of 0.01 to 0.10, the change in the F1 score for all four users was less than 5%, confirming the framework's robustness to this key parameter. Second, a sensitivity analysis was performed on the reconstruction error evaluation metrics. The results indicate that the MSE metric, which assigns higher weights to large deviations, is more sensitive to detecting spike anomalies than the MAE. Therefore, the choice of MSE as the primary evaluation metric in this study is justified. Furthermore, time-robustness verification was conducted by training on 2021–2022 data and testing on 2023 data. The results showed that the F1 score decreased by less than 3%, indicating that the proposed framework possesses good cross-period generalization capabilities and can adapt to the time-varying characteristics of load patterns.

Case-specific analyses further validate the practicality of Transformer-IF. For users in the chemical, steel, and electrical equipment manufacturing industries, the framework effectively removes instantaneous anomalies and non-physical abrupt changes while preserving the continuity and periodicity of load trends. The extracted typical curves are smoother in shape and more representative. This not only enhances the interpretability and robustness of TLCs but also provides a reliable data foundation for down-stream applications such as electricity market trading strategy formulation, load fore-casting, and demand response.

Furthermore, experiments demonstrate that the Transformer-IF framework possesses strong generalization capability and stability when handling high-volatility or noise-dense industrial load scenarios. Compared with single-stage autoencoders or directly applying Isolation Forest to raw load sequences, the two-stage processing strategy achieves a favorable balance between anomaly identification and trend preservation, offering an effective solution for the scientific extraction of typical load curves.

6. Conclusions

This paper proposes a two-stage “purification-first, extraction-second” framework based on Transformer-IF, designed to address the challenges of filtering out anomalous daily data and extracting robust typical load curves from high-resolution load data of industrial and commercial users. The model employs a Transformer-based autoencoder to learn from historical daily load data and reconstruct input sequences, effectively capturing daily, weekly, and multi-scale trend variations. During the reconstruction process, the model is able to distinguish between normal load evolution patterns and abnormal mutations, thereby providing high-quality input features for sub-sequent anomaly detection. Furthermore, the framework feeds the reconstruction residuals into an Isolation Forest model to automatically identify anomaly days that deviate from normal trends, including abnormal consumption days as well as spike and trough mutations. Compared with single-stage models or traditional statistical methods, Transformer-IF employs a two-stage processing strategy that removes anomalies while preserving the continuity and fluctuation characteristics of normal load trends, ensuring the physical plausibility and representativeness of the typical load curve.

In its current implementation, Transformer-IF utilizes the temporal features of historical load data for learning and anomaly removal, ensuring unsupervised applicability and strong generalization capability. Nevertheless, the model has room for further optimization: for example, incorporating additional exogenous influencing factors

(e.g., weather, calendar events, production schedules), or improving algorithmic design and computational performance to enhance processing efficiency and accuracy in large-scale industrial load data scenarios.

In summary, Transformer-IF provides a robust and interpretable two-stage data-driven method for anomaly removal and typical curve extraction from high-resolution load data, offering reliable data support and technical references for electricity market trading strategies, load forecasting, and integrated energy system planning, while also indicating future research directions for expanding model capabilities and incorporating more influencing factors. As part of future research, we will adopt interval estimation approaches to characterize the uncertainty of the extracted typical load curves, thus enabling a probabilistic framework for electricity market decision support.

Author Contributions: Conceptualization, J. Y.; software, C.P.; validation, X.L.; formal analysis, C. P. and X.L.; writing—original draft, J.Y. and Y.L.; writing—review and editing, J.L. and K.X.; supervision, K.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the science and technology project of State Grid Corporation of China: Research and application of key technologies for electrical energy accurate prediction of user group to support the new business of electrical energy agency purchase (No. 5400-202312226A-1-1-ZN).

Data Availability Statement: The industrial load data that support the findings of this study are proprietary and subject to confidentiality agreements with the data provider, and thus are not publicly available. After the project is completed and the relevant confidentiality restrictions are lifted, the source code, along with a synthetic sample dataset and experimental configuration for the Transformer-IF framework, will be available from the corresponding author for non-commercial academic purposes upon reasonable request.

Conflicts of Interest: Author Xin Luo was employed by Beijing Tsingsoft Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TLC	Typical Load Curves
Transformer-IF	Transformer-Isolation Forest hybrid framework
USAD	Un-Supervised Anomaly Detection on Multivariate Time Series
AST	Adversarial Spatial-Temporal network

References

1. Nystrup, P.; Madsen, H.; Blomgren, E.M.V.; de Zotti, G. Clustering commercial and industrial load patterns for long-term energy planning. *Smart Energy* **2021**, *2*, 100010.
2. Flygare, C.; Nystrand, M.; Eriksson, R.; Castellucci, V. Load profiling using a two-step clustering framework: Capturing electricity usage variability based on temporal and ambient factors. *Int. J. Electr. Power Energy Syst.* **2025**, *171*, 110924.
3. Michalakopoulos, V.; Sarmas, E.; Papias, I.; Skaloumpakas, P.; Marinakis, V.; Doukas, H. A machine learning-based framework for clustering residential electricity load profiles to enhance demand response programs. *Appl. Energy* **2024**, *361*, 122943.
4. Di Piazza, A.; Di Piazza, M.C.; La Tona, G.; Luna, M. An artificial neural network-based forecasting model of energy-related time series for electrical grid management. *Math. Comput. Simul.* **2021**, *184*, 294–305.
5. Morgoeva, A.D.; Morgoev, I.D.; Klyuev, R.V.; Khetagurov, V.N.; Gavrina, O.A. Short-term prediction of energy consumption at concentration factory. *Min. Inf. Anal. Bull.* **2023**, *5*, 157–169.
6. Li, H.; Wang, Z.; Hong, T.; Parker, A.; Neukomm, M. Characterizing patterns and variability of building electric load profiles in time and frequency domains. *Appl. Energy* **2021**, *291*, 116721.

7. Gao, J.; Chen, Y.; Hu, W.; Zhang, D. An adaptive deep-learning load forecasting framework by integrating transformer and domain knowledge. *Adv. Appl. Energy* **2023**, *10*, 100142.
8. Kwon, W.; Jung, M.; Park, J.; Lee, S. FFT-Guided Multi-Window USAD with DTW–Isolation Forest for Reliable Anomaly Detection in Industrial Power Time-Series. *Energies* **2025**, *18*, 6584.
9. Zhang, X.; Ramírez-Mendiola, J.L.; Li, M.; Guo, L. Electricity consumption pattern analysis beyond traditional clustering methods: A novel self-adapting semi-supervised clustering method and application case study. *Appl. Energy* **2022**, *308*, 118335.
10. Jain, M.; Jain, M.; AlSkaif, T.; Dev, S. Which internal validation indices to use while clustering electric load demand profiles? *Sustain. Energy Grids Netw.* **2022**, *32*, 100849.
11. Sandoval Guzmán, B.; Barocio Espejo, E.; Elser, M.; Korba, P.; Segundo Sevilla, F.R. A hybrid clustering approach for electrical load profiles considering weather conditions based on matrix-tensor decomposition. *Sustain. Energy Grids Netw.* **2024**, *38*, 101326.
12. Eskandarnia, E.; Al-Ammal, H.M.; Ksantini, R. An embedded deep-clustering-based load profiling framework. *Sustain. Cities Soc.* **2022**, *78*, 103618.
13. Kumar, A.; Mallipeddi, R. A deep clustering framework for load pattern segmentation. *Sustain. Energy Grids Netw.* **2024**, *38*, 101319.
14. Lin, R.; Chen, S.; He, Z.; Wu, B.; Zou, H.; Zhao, X.; Li, Q. Electricity behavior modeling and anomaly detection services based on a deep variational autoencoder network. *Energies* **2024**, *17*, 3904.
15. Shi, J.; Teh, J. Load forecasting for regional integrated energy system based on complementary ensemble empirical mode decomposition and multi-model fusion. *Appl. Energy* **2024**, *353*, 122146.
16. Huang, Z.; Yi, Y. Short-Term Load Forecasting for Regional Smart Energy Systems Based on Two-Stage Feature Extraction and Hybrid Inverted Transformer. *Sustainability* **2024**, *16*, 7613.
17. Lotfipoor, A.; Patidar, S.; Jenkins, D.P. Transformer network for data imputation in electricity demand data. *Energy Build.* **2023**, *300*, 113675.
18. Wang, S.; Zhao, C.; Liu, X.; Ni, X.; Chen, X.; Gao, X.; Sun, L. Hybrid deep learning framework for anomaly detection in power plant systems. *Algorithms* **2025**, *18*, 704.
19. Zhang, Z.; Qiu, G.; Cheng, Y.; Wang, M. Densely-connected decoder transformer for unsupervised anomaly detection of power electronic systems. *J. Autom. Intell.* **2025**, *4*, 217–226.
20. Patel, V.; Kapoor, A.; Sharma, A.; Chakrabarti, S. Taxonomy of outlier detection methods for power system measurements. *Energy Convers. Econ.* **2023**, *4*, 73–88.
21. Yepmo, V.; Smits, G.; Lesot, M.-J.; Pivert, O. Leveraging an isolation forest to anomaly detection and data clustering. *Data Knowl. Eng.* **2024**, *151*, 102302.
22. Lei, L.; Guo, X.; Zheng, R. An innovative dynamic anomaly detection method based on hybrid data mining technology for building energy consumption. *Energy Build.* **2024**, *319*, 114559.
23. Wang, X.; Wang, H.; Bhandari, B.; Cheng, L. AI-Empowered methods for smart energy consumption: A review of load forecasting, anomaly detection and demand response. *Int. J. Precis. Eng. Manuf.-Green Technol.* **2024**, *11*, 963–993.
24. Cooper, A.; Bretas, A.; Meyn, S. Anomaly detection in power system state estimation: Review and new directions. *Energies* **2023**, *16*, 6678.
25. Audibert, J.; Michiardi, P.; Guyard, F.; Marti, S.; Zuluaga, M.A. USAD: UnSupervised Anomaly Detection on Multivariate Time Series. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20), Virtual Event, 23–27 August 2020; pp. 3395–3404.
26. Wu, S.; Xiao, X.; Ding, Q.; Zhao, P.; Wei, Y.; Huang, J. Adversarial sparse transformer for time series forecasting. In *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*; Curran Associates, Incorporated: Red Hook, NY, USA, 2020. Available online: <https://proceedings.neurips.cc/paper/2020/hash/c6b8c8d762da15fa8dbdbf6baf9e260-Abstract.html> (accessed on 4 March 2026).

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