A Hybrid AI-Based Framework for Robust Sleeping Cell Detection in 5G Networks

Le Nhu Quynh [1], Truong Duc Tai [2,*], Dinh Thi Phuong [2], Nguyen Anh Tu [2], and Tran Van Tung [2]

[1] Faculty of Information Technology, Post and Telecommunications Institute of Technology, Hanoi, 100000, Vietnam
[2] Broadband Wireless Center, Viettel High Technology Industries Corporation, Viettel Group, Hanoi, 100000, Vietnam
Email: quynhln@ptit.edu.vn (L.N.Q.); taitd5@viettel.com.vn (T.D.T.); phuongdt9@viettel.com.vn (D.T.P.); tuna8@viettel.com.vn (N.A.T.); tungtv8@viettel.com.vn (T.V.T)
*Corresponding author

***Abstract*—The rapid evolution of 5G networks introduces complex operational challenges, particularly the "sleeping cell" phenomenon, where base stations appear functional in management systems yet fail to serve users. Although, traditional monitoring systems rely on static thresholds that struggle to adapt to the non-stationary, multi-periodic nature of 5G traffic, leading to excessive false alarms and missed degradations. This paper proposes a Hybrid Artificial Intelligence (AI)-Based Framework for robust sleeping cell detection that reconciles domain expertise with neural sensitivity. The framework integrates 3 components: a deterministic rule-based engine for high-confidence anomaly identification, a multi-task self-supervised temporal backbone to learn representations of normal behavior, and a supervised ensemble trained on domain-aware synthetic anomalies to address label scarcity. A key contribution is the introduction of a rule-priority fusion mechanism that ensures critical service-loss alarms take precedence while machine learning modules identify subtle, non-linear degradations. Evaluated on 126,801 hourly Key Performance Indicator (KPI) records from 194 live 5G cells, the proposed approach achieves an F1-Score of 0.728 and a The Receiver Operating Characteristic-Area Under the Curve (ROC-AUC) of 0.856, with 98.7% agreement with expert rules and the discovery of 11,375 previously undetected degradations. The results demonstrate a practical and scalable solution for proactive fault management in real-world 5G networks.**

***Keywords*—5G networks, sleeping cell detection, time-series anomaly detection, self-supervised learning, Temporal Convolutional Attention Network (TCAN), ensemble learning**

I. Introduction

The Radio Access Network (RAN) is a critical component of modern cellular systems, responsible for connecting end users to the core network and ensuring Quality of Service (QoS) and user experience Quality of Earnings (QoE). Modern 5G deployments leverage ultra-dense gNodeB configurations to support diverse service requirements, generating high dimensional telemetry through hundreds of Key Performance Indicators (KPIs) such as Session Success Rate (SSR), Call Drop Rate (CDR), throughput, active users, and traffic volume. The management of these complex environments increasingly aligns with global frameworks such as ISO/IEC TS 25011 [1] for service quality measurement and ISO/IEC 27004 [2] for systematic performance monitoring. However, these KPI streams exhibit non-stationary behaviors driven by diurnal patterns, mobility, and configuration shifts, creating intricate temporal and spatial dependencies. Such volatility fundamentally challenges traditional monitoring paradigms, necessitating a shift from rigid thresholding toward the intelligent, multi-objective frameworks explored in this work [3].

Anomalies in such complex KPI environments are diverse, subtle, and often interdependent, ranging from sudden spikes and level shifts to gradual degradations and temporary instabilities. Among these, the sleeping cell is one of the most critical types of anomalies. A sleeping cell refers to a base station cell that appears operational in the network management system but fails partially or completely to serve users [4]. Unlike catastrophic hardware failures that trigger immediate alarms, sleeping cells exhibit gradual degradation patterns that evade conventional monitoring systems, creating undetected coverage gaps and degraded QoE. Detecting such anomalies requires models capable of capturing multi-scale temporal patterns, cross-KPI dependencies, and neighboring cell correlations [5, 6].

Traditional rule-based methods, which rely on static thresholds, remain a common approach in production telecom and IT environments [7, 8]. However, these rigid systems produce excessive false alarms, with operational data revealing that network operations teams handle thousands of alerts daily, and studies indicating that approximately 99% of security operation center alarms are false positives [9, 10]. This "alarm fatigue" forces time-consuming manual triage workflows, with some teams spending up to three hours daily on alert review [11, 12], causing critical anomalies to be deprioritized or overlooked. Early interventions in time series anomaly detection relied on statistical foundations

Manuscript received November 23, 2025; revised February 4, 2026; accepted February 28, 2026; published June 29, 2026.

such as Z-score analysis [13, 14] and Auto Regressive Integrated Moving Average (ARIMA) models [15], which utilized pointwise deviations and linear forecasting to identify irregularities. However, these methods often fail to capture the non-linear dynamics of complex sequences. Subsequent deep learning approaches adopted Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) [16–18] to model temporal dependencies, yet they suffered from recurrent bottlenecks and gradient vanishing over long sequences. To meet the real-time constraints of 5G environments, research has evolved toward parallel attention architectures including Transformers [19, 20] which capture global trends with higher computational efficiency. Currently, the paradigm is shifting toward Multi-Task Self-Supervised Learning (SSL) [21]. By synthesizing reconstruction, forecasting, and contrastive objectives, these modern frameworks) [22, 23] learn robust representations of normality that significantly outperform traditional single-objective approaches. In parallel, 5G-specific frameworks [24, 25] and comprehensive benchmarks [26, 27] demonstrate the practical effectiveness of AI-driven fault management in production networks. Despite these advancements, 3 critical research gaps persist. First, most Time Series Anomaly Detection (TSAD) models rely on a single anomaly criterion either reconstruction error or forecasting residuals, failing to triangulate subtle faults through complementary embedding perspectives. While multi-task SSL shows promise [22], existing 5G Radio Access Network (RAN) monitoring systems have not exploited joint optimization of reconstruction, forecasting, and contrastive objectives for robust anomaly detection. Second, recent hybrid frameworks [28, 29] integrate rule-based logic with deep learning but largely rely on symmetric or additive fusion strategies. As a result, domain-critical zero-tolerance rules are treated on par with uncertain Machine Learning (ML) probabilities, allowing high-confidence normal predictions to statistically dilute rule-triggered service-loss alarms. The asymmetric importance of domain knowledge in telecom fault management therefore remains insufficiently addressed. Third, global benchmarks [26, 30] remain domain-agnostic, lacking the parametric failure signatures unique to 5G infrastructure such as gradual Social Security Rulings-Continuing Disability Review (SSR-CDR) coupled drifts and oscillatory instabilities characteristic of sleeping cells. This domain gap limits the practical applicability of models trained on generic datasets.

Motivated by these gaps, we propose a Hybrid AI-Based Framework that first detects clear anomalies using expert-defined domain rules and then leverages multi-task self-supervised learning and supervised ensemble models to capture complex, hidden patterns.

Our main contributions are as follows:

- Triple-Objective MultiTask: The Center for Arts in Natick (TCAN) self-supervised learning: We enhance the TCAN architecture with joint optimization for forecasting, reconstruction, and contrastive learning tasks to learn robust, multi-scale temporal representations of KPI behavior [31].
- Domain-Aware Synthetic Tuning: We address label scarcity by generating synthetic anomalies (dips, drifts, shifts) modeled after actual 5G failure curves to train a Smart Supervised Ensemble.
- Rule-Priority Logic-Based Merging: We introduce an asymmetric fusion mechanism where expert-defined rules act as a high-priority gatekeeper, ensuring critical outages take precedence while ML modules capture subtle degradations, thereby balancing sensitivity, precision, and interpretability.
- Comprehensive evaluation: We validate the framework on a realistic synthetic dataset modeled from live 5G telemetry records. Our results demonstrate that the hybrid pipeline effectively reconciles domain expertise with neural sensitivity, outperforming traditional rule-only and standalone self-supervised baselines.

The rest of this paper is organized as follows: Section II reviews related work, Section III presents the proposed methodology and describes the experimental setup, Section IV discusses the results, and Section VI concludes the paper.

II. Literature Review

A. Evolution of Temporal Modeling in RAN Anomaly Detection

Mobile networks exhibit diverse anomaly types, including sudden KPI spikes, level shifts, gradual degradations, temporary instabilities, and sleeping cells-base stations that appear operational yet fail to serve users [3]. Early RAN monitoring relied on rule- or threshold-based KPI tracking (e.g., SSR, CDR, and throughput). While interpretable, these heuristic methods struggle with the non-stationary, multi-periodic nature of 5G traffic [5, 6]. Statistical approaches like Z-score analysis [13, 14], ARIMA [15] improved robustness but lacked the capacity to model complex cross metric dependencies. To address this, unsupervised reconstruction methods were introduced; however, they are sensitive seasonal or diurnal fluctuations when faced with non-stationary, leading to high false-positive rates [32–36]. Conversely, supervised models offer precision but are hindered by the extreme scarcity of labeled anomalies in live 5G environments and poor generalization under domain shifts [24, 37, 38]. Initial deep learning attempts utilized RNNs and LSTM units [17]. While revolutionary, these architectures suffer from “recurrent bottlenecks” which is the inability to process long sequences in parallel and the tendency for gradients to vanish over extended horizons [16]. This is particularly problematic for 5G RAN, where high-frequency telemetry requires capturing dependencies over thousands of time steps.

Recent work has shifted toward parallel temporal architectures. The Temporal Convolutional Network (TCN) introduced dilated causal convolutions for efficient long-range modeling [39]. Transformer-based approaches capture global temporal dependencies with reduced

computational complexity through sparse attention mechanisms [19, 40]. Decomposition-based models isolate seasonal components [41], while State Space Model–based approaches enable efficient long-range dependency modeling [42]. However, these architectures often lack sensitivity to local point anomalies or introduce computational overhead unsuitable for real-time 5G inference.

Building upon the Temporal Convolutional Attention Network (TCAN) proposed by Lin *et al.* [31], which integrates dilated temporal convolutions with multi-head attention, this work extends the TCAN paradigm through multi-task self-supervised learning tailored for 5G RAN anomaly detection. Motivated by hybrid Thriving Communities Network (TCN)–attention [43, 44], the proposed architecture captures multi-scale temporal dynamics via dilated convolutions while leveraging attention mechanisms to emphasize salient temporal and cross-metric dependencies. This hybrid design enables robust and precise detection of subtle performance degradations under unlabeled and non-stationary 5G RAN conditions, while maintaining practical inference efficiency.

B. Progression Toward Multi-Task Self-Supervised Learning

To address 5G label scarcity, research has shifted toward Self-Supervised Learning (SSL). Zhang *et al.* [21] categorize these methods into generative, contrastive, and adversarial paradigms, driving a transition from isolated single-objective tasks toward integrated multi-task frameworks that capture complex temporal dynamics.

1) Single-objective SSL approaches

Reconstruction and Forecasting: Traditional SSL methods rely on autoencoders to detect anomalies via reconstruction error, or on predictive models that identify deviations through forecasting residuals [33, 34].

Contrastive Representation Learning: More recent advances, such as Carla [45], Softclt [46], leverage contrastive learning to separate normal temporal neighborhoods from anomalous outliers within a latent representation space.

2) Multi-task SSL integration

However, existing research typically treats these objectives in isolation. Recent work by Choi and Kang [22] demonstrates that multi-task self-supervised learning, which combines contextual, temporal, and transformation consistency with uncertainty weighting, significantly outperforms single-objective approaches across multiple downstream tasks, including anomaly detection. Similarly, Zeng *et al.* [47] proposed universal time series representation learning by jointly optimizing reconstruction and contrastive objectives to improve generalization.

Building on this paradigm shift, this work synthesizes reconstruction, forecasting, and contrastive learning into a Triple-Objective Multi-Task TCAN architecture. By jointly optimizing these 3 SSL heads, the proposed framework provides a more robust triangulation of anomalies compared to single-objective models. This integrated approach is particularly effective for 5G KPI time series, which often exhibit heavy-tailed noise that can induce false positives in simpler reconstruction-based SSL methods.

C. Knowledge-Augmented Frameworks and Rule-priority Fusion

The latest frontier in telecommunications AI involves bridging the gap between black box neural weights and the established operational logic of network providers. Recent industry trends show a movement toward Knowledge-Augmented Neural Networks that honor domain-specific recovery logic. Frameworks have introduced multi-stage pipelines that combine rule engines with semi-supervised classifiers [28, 29]. However, these often use simple ensemble voting or averaging, which can dilute the importance of critical service-loss rules. This paper advances the "Hybrid Fusion" paradigm through a Rule-Priority Logic Based Merging mechanism. Unlike standard ensembles, this structure ensures that expert-defined rules-critical for zero-tolerance outages take precedence as a priority filter, while the TCAN module identifies subtle, hidden degradations that rules may miss. This modular "plug-and-play" design allows for independent updates to the rule engine or the ML modules, ensuring the pipeline remains scalable and operationally resilient for processing vast hourly record datasets.

D. Benchmarking Standards and Data Challenges

The standardization of anomaly detection has been propelled by benchmarks like TSB-UAD [26] and TSB-AutoAD [30], which provide diverse testing environments. For evaluating model generalization. Despite these global standards, the specific task of 5G sleeping cell detection suffers from a lack of publicly available, high-quality labeled datasets. Most State-of-the-Art (SOTA) models are evaluated on synthetic or generic IoT data, which does not reflect the complex spatial-temporal behavioral clusters of a real RAN. This study addresses this gap by training a supervised ensemble on domain-aware synthetic anomalies, ensuring the model is specifically tuned to the failure modes (e.g., gradual traffic drift vs. abrupt service loss) observed in live 5G infrastructure.

III. Materials and Methods

A. Method

This section presents the proposed Hybrid AI-Based Framework for robust sleeping cell detection in 5G networks, together with the dataset, preprocessing pipeline, baseline models, evaluation metrics, and implementation details used to validate its effectiveness. The framework integrates rule-based detection, self-supervised temporal modeling, synthetic data augmentation with a supervised ensemble, and a rule-priority fusion mechanism to achieve high accuracy, interpretability, and adaptability in complex 5G network dynamics.

1) Problem formulation and overview of hybrid pipeline

a) Formal task definition

Let $\mathcal{X} = \{x_t\}_{t=1}^{T} \in R^{T\times K}$ denote a multivariate KPI time series collected from a 5G cell, where K represents the number of monitored metrics (SSR, CDR, traffic, etc.) and T is the temporal length. The objective is to learn a decision function $f: \mathcal{X} \to \{0,1\}$ that reliably identifies sleeping cells—base stations that remain operational in network management systems yet fail to effectively serve users, while minimizing both false positives (alarm fatigue) and false negatives (missed degradations). Detecting such conditions requires capturing subtle, non-linear performance degradations that typically evade traditional threshold-based alarm mechanisms.

b) Architectural overview

- The proposed framework decomposes anomaly detection into three complementary pillars (Fig. 1): (1) a Deterministic Rule Engine R that enforces zero-tolerance domain constraints, (2) a Multi-Task Self-Supervised TCAN Φ_{SSL} that learns temporal-spatial representations of normal behavior, and (3) a Supervised Ensemble Ψ_{Ens} trained on domain-specific synthetic failures. These components are unified through a rule-priority fusion mechanism that adaptively reconciles expert knowledge with learned patterns.

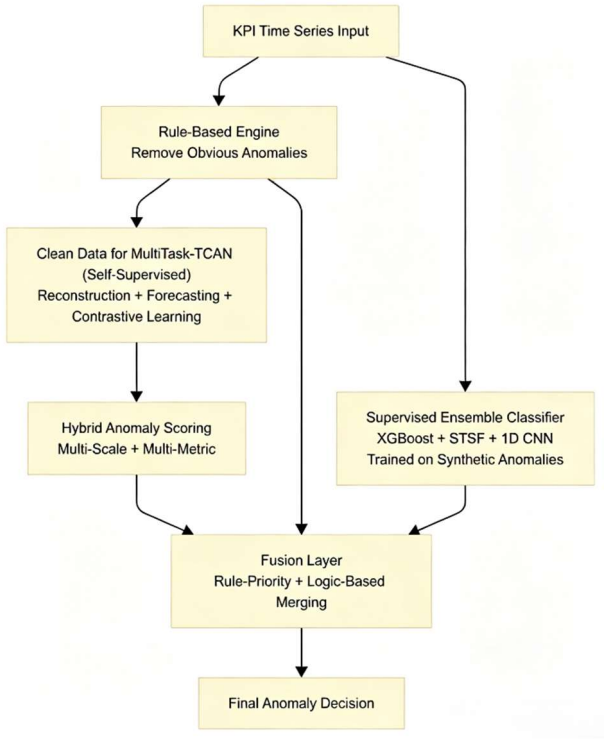

Fig. 1. The proposed framework is designed as a multi-stage pipeline that sequentially processes network KPI data. The workflow proceeds through rule-based anomaly filtering, self-supervised representation learning using MultiTask-TCAN, supervised classification.

2) Rule-based engine

The rule engine $R(X) \to \{0,1\}$ encodes telecommunications domain expertise to identify high-confidence anomalies that violate fundamental service guarantees. Rather than exhaustive threshold checking, R implements a stratified detection hierarchy where each rule category targets distinct failure modes observable in 5G RAN operations (Table I).

Each rule r_i is assigned a severity weight based on its operational impact. This priority encoding directly influences the final fusion mechanism, ensuring that "Zero-Tolerance" failures take precedence as a high-priority gatekeeper. By functioning as both an interpretable first line of defense and a noise filter, R complements the downstream Multi-Task TCAN (Φ) and Supervised Ensemble (Ψ_{Ens}), which focus on subtler or previously unseen degradations.

To address scalability and maintainability concerns, rule parameters (e.g., persistence duration, deviation thresholds) are automatically calibrated using statistical baselines computed from rolling historical data. This enables the system to adapt to evolving traffic dynamics without manual reconfiguration. Within this architecture, the rule-based layer functions as both a noise filter and an interpretable first line of defense, complementing downstream learning-based detectors that handle subtler or previously unseen anomalies.

3) MultiTask TCAN

a) Architectural foundation

The proposed self-supervised pipeline builds upon the TCAN architecture [31], denoted as $\Phi_{SSL}: R^{T\times K} \to R^{d}$, where T represents the temporal window length, K denotes the number of monitored KPIs, and d is the embedding dimension. The architecture utilizes a hierarchy of L enhanced temporal blocks. Each block $\mathcal{B}$ is defined by the triplet of dilated causal convolutions for multi-scale receptive fields, Multi-Head Attention (MHA) for cross-metric correlation, and residual normalization for convergence stability.

While the base TCAN architecture demonstrated strong performance in time series forecasting, we enhance it with multi-task self-supervised learning objectives (reconstruction, forecasting, and contrastive learning) specifically designed for 5G anomaly detection (Fig. 2) [31].

b) Joint multi-task optimization

Unlike single-objective Self-Supervised Learning (SSL) approaches that rely solely on reconstruction or forecasting, the proposed framework jointly optimizes three complementary SSL objectives to triangulate anomalies from distinct representation perspectives: Reconstruction Loss (L_{recon}) captures fundamental structural dependencies by minimizing $\mathcal{L} = |X - \hat{X}|_2^2$ where $\hat{X}$ denotes the decoder output.

Forecasting Loss ($\mathcal{L}_{\text{frcs}}$) models' short-term predictive dynamics over a 24-hour horizon (h = 24) using $L_{\text{forecast}} = |X_{t+1:t+h} - \widehat{X_{t+1:t+h}}|_2^2$ Contrastive Loss (L_{contrast}) enforces discriminative latent representations by maintaining invariance to minor perturbations while ensuring separability from anomalous outliers.

TABLE I. DOMAIN KNOWLEDGE MAPPING: RULE CATEGORIES AND TRIGGER LOGIC

Anomaly Category	KPI Trigger Logic	Priority	Formal Constraint
Service Outage	Simultaneous SSR, traffic, and user count collapse	Critical (P0)	$(SSR_t < \tau_s) \wedge (T_t < \tau_{tr}) \wedge (U_t < \tau_u)$
Gradual Drift	Sustained polynomia-l trend deviation	High (P1)	$\lvert\text{polyfit}(X_{t-w:t})\rvert > \theta_{\text{drift}}$
Temporal Anomaly	Weekend/night deviation from baseline	Medium (P2)	$\lvert x_t - \mu_{\text{context}}(t)\rvert > k \cdot \sigma_{\text{context}}$
Spatial Imbalance	Inter-cell KPI divergence within cluster	Medium (P2)	$\max_{i,j \in \text{cluster}} \lvert x_i - x_j\rvert > \tau_{\text{spatial}}$
Oscillatory Instability	Repetitive on-off patterns	High (P1)	$\max(\text{PSD}(X)) > \theta_{osc}$

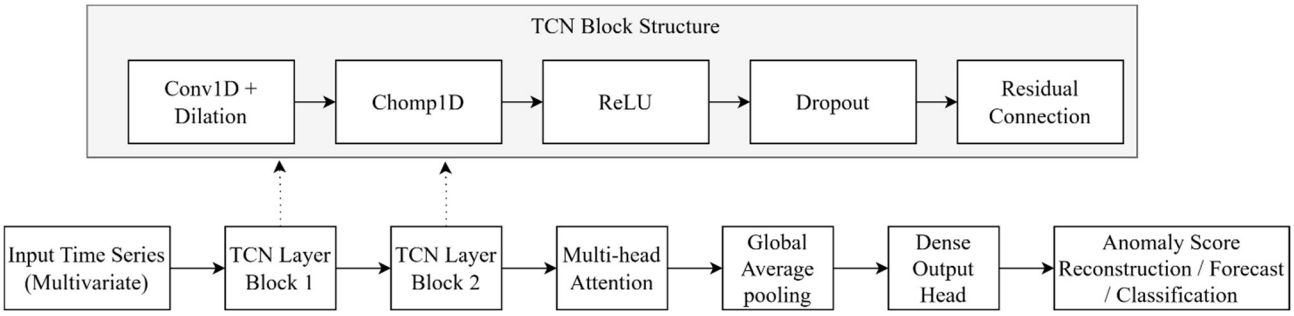

Fig. 2. Proposed TCAN architecture. Internal structure of the TCAN block with stacked temporal layers, integrating a specialized multi-head attention mechanism for multi-dimensional KPI dependency extraction.

The overall SSL objective aggregates the 3 losses using empirically calibrated weights, formulated as Eq. (1):

$$\mathcal{L}_{TCAN} = \alpha\mathcal{L}_{recon} + \beta\mathcal{L}_{forecast} + \gamma\mathcal{L}_{constrast} \quad (1)$$

where $\alpha = 0.4$, $\beta = 0.4$ and $\gamma = 0.2$.

These coefficients were empirically determined through ablation studies to achieve a balanced optimization among reconstruction, forecasting, and contrastive objectives. The reconstruction and forecasting tasks are assigned equal weights ($\alpha = \beta = 0.4$) to jointly capture the underlying temporal structure and short-term predictive dynamics of KPI sequences. A smaller contrastive weight ($\gamma = 0.2$) ensures that representation learning enhances feature separability without overpowering the reconstruction–forecasting balance or distorting the latent space.

Adjusting these coefficients influences both convergence behavior and downstream anomaly detection performance. Increasing γ strengthens embedding discrimination but may slightly degrade reconstruction accuracy, while emphasizing either α or β biases the model toward structural fidelity or predictive sharpness, respectively. Nonetheless, within a moderate range (±0.1), the framework remains robust, showing less than 3% variation in F1-Score. In practice, the chosen ratio (0.4: 0.4: 0.2) provided the most stable convergence and consistent anomaly detection performance across multiple RAN KPI datasets.

c) Multi-scale temporal modeling

To enable multi-scale anomaly detection, 3 TCAN models are trained on different window lengths:

- 72 h (3 days) for short-term patterns,
- 144 h (6 days) for medium-term patterns,
- 288 h (12 days) for long-term patterns.

For each window, the final SSL anomaly score is a weighted combination of reconstruction error, forecast error, and embedding deviation, aggregated across all scales:

$$S_{TCAN} = w_r E_{recon} + w_f \mathcal{L}_{forecast} + w_e \|z\| \quad (2)$$

where $w_r = 1.0$, $w_f = 0.5$, and $w_e = 0.3$.

4) Supervised ensemble with domain-aware synthetic training

While SSL is effective at detecting previously unseen anomalies, it may produce false positives for recurring and well-characterized failure patterns. To address this limitation, we train a supervised ensemble model, denoted as Ψ_{Ens}, on synthetically generated anomalies that explicitly encode 5G domain-specific degradation signatures.

a) Synthetic anomaly injection

We construct 5 parametric anomaly types: dips, drifts, level shifts, oscillations, and sleeping cells using distribution preserving transformations that preserve the statistical characteristics of real KPI time series (see Section IV).

b) Ensemble architecture

The supervised ensemble Ψ_{Ens} integrates 3 complementary classifiers: XGBoost (gradient-boosted decision trees), Supervised Time Series Forest (STSF), and a lightweight One-Dimensional Convolutional Neural Network (1D-CNN). Model outputs are aggregated via soft voting to balance robustness and sensitivity. Key hyperparameters, including tree depth, number of estimators, and convolutional filter sizes, are optimized using grid search on a held-out synthetic validation set.

5) Rule-priority fusion

a) Fusion formulation

The fusion layer reconciles deterministic rule-based outputs, $S_{rule} \in \{0,1\}$, with probabilistic machine learning scores, S_{SSL}, $S_{\text{Ens}} \in [0,1]$, through an asymmetric priority-aware gating function:

$$S_{final} = \lambda_r S_{rule} + (1-\lambda_r)(\eta S_{SSL} + (1-\eta) S_{ensemble}) \quad (3)$$

where $S_{rule} \in \{0,1\}$ is a binary rule flag; S_{SSL}, $S_{\text{Ens}} \in [0,1]$ are normalized anomaly scores; $\lambda_r \in [0.5,1.0]$ is the rule weight (Table I); and $\eta \in [0,1]$ balances model contributions. Priority Encoding

For critical failure conditions (e.g., service outages), $\lambda_r \to 1.0$ enforces strict rule precedence, ensuring alarms are triggered regardless of model uncertainty. For medium-priority rules or borderline rule activations accompanied by strong model evidence ($S_{\text{SSL}} > 0.8$), the fusion mechanism adaptively increases the influence of learning-based components. This asymmetric design guarantees that critical failures are never suppressed by voting effects, while still enabling ML models to identify subtle degradations that evade static thresholds.

b) Confidence stratification

The final fusion score is stratified into 3 operational tiers: (i) High (rule–model consensus on critical outages); (ii) Medium (sleeping cells and gradual drifts identified by SSL and Ensemble modules); and (iii) Low (transient noise suppression). This modular design enables independent rule updates without retraining, supporting robust and flexible NOC deployment.

B. Synthetic Dataset Constructuion and Validation

1) Motivation and design principles

Due to the scarcity of labeled anomalies in real 5G KPI datasets, we developed a domain-aware synthetic dataset that closely replicates realistic network behavior while enabling controlled injection of anomaly patterns. The design principles ensure preservation of: (1) statistical distributions and cross-KPI correlations, (2) temporal dependencies and daily seasonality, (3) domain-specific KPI interactions, and (4) anomaly characteristics observed in operational networks.

2) Baseline construction

The dataset was constructed from 126,801 real 5G KPI records, including ENDC Secondary Serving Radio (ENDC SSR), ENDC Charging Data Record (ENDC CDR), NR Range Ambiguous Signal Ratio (NR RASR), Max Users, and Traffic. A normal baseline was generated by applying a low-pass filter to remove high-frequency noise while preserving mean, variance, and periodicity. Clean windows of length 72 h, 144 h, and 288 h were extracted as candidate segments for anomaly injection.

3) Anomaly injection algorithms

Anomaly patterns were generated using parametric formulations designed to preserve statistical fidelity. 5 representative Algorithms 1−5 are provided below.

Algorithm 1: Distribution-Preserving Drift Generation

Require: Original KPI series X, duration T, intensity α, baseline standard deviation σ, target value v_t.
Ensure: Modified series $\bar{X}$ with drift anomaly
$c \leftarrow linspace(0, 0.3\alpha, T)$ ▷ Generate transition curve
$n \sim N(0, 0.02\sigma)$ of length T ▷ Sample Gaussian noise
$\bar{X} \leftarrow X \cdot (1-c) + v_t \times c + n$ ▷ Apply drift anomaly

Algorithm 2: Percentile-Based Target Generation

Require: KPI series X, percentile p.
Ensure: Target value v_t for anomaly generation anomaly
$X' \leftarrow X \setminus \{NaN\}$ ▷ Remove missing values
$v_t \leftarrow Percentile(X',\ p)$

Algorithm 3: Controlled Increase/Decrease Anomaly Generation

Require: Original KPI series X, duration T, intensity α, maximum multiplier m_{max}
Ensure: Modified series $\overline{X}$ with increase or decrease anomaly
Decrease:
$c_d \leftarrow linspace(1.0,\ 1-0.4\alpha,\ T)$
$n \sim N(1.0, 0.01)$ of length T
$\overline{X} \leftarrow X \times c_d \times n$
Increase:
$m_c \leftarrow 1 + (m_{max} - 1) \times 0.5\alpha$
$c_i \leftarrow linspace(1.0,\ m_c,\ T)$
$\overline{X} \leftarrow X \times c_i \times n$

Algorithm 4: Oscillatory Fluctuation Anomaly Generation

Require: Original KPI series X, duration T, intensity α
Ensure: Modified series $\overline{X}$ with oscillatory anomaly
$f \leftarrow 0.2 + U(0,\ 0.3)$ ▷ Sample frequency
$t \leftarrow linspace(0.2\pi,\ T)$
$o \leftarrow \sin(t \times f) \times 0.3\alpha$
$n \sim N(0, 0.005)$ of length T
$m \leftarrow 1 + o + n$
$\overline{X} \leftarrow X \odot m$

Algorithm 5: Sleeping Cell Anomaly Generation (Network-Specific)

Require: Original SSR series X_{SSR}, original CDR series X_{CDR}, duration T
Ensure: Modified SSR series $\overline{X}_{SSR}$ series, CDR series $\overline{X}_{CDR}$
$\overline{X}_{SSR} \leftarrow X_{SSR} \times linspace(1.0,\ 0.85\alpha,\ T)$
$c \leftarrow 1 + \left(1 - \overline{X}_{SSR}/X_{SSR},\right) \times 0.3$
$\overline{X}_{CDR} \leftarrow X_{CDR} \times c$

4) Preservation constraints

To ensure anomaly realism, the following constraints were enforced:

- Statistical moment preservation: Mean and standard deviation changes constrained to < 5%.
- Effect size control: Cohen's d maintained between 0.5 and 1.2 for detectability.
- Anomaly spacing: Minimum of 100-time steps between anomaly events.
- Smooth transitions: Gradual curves used instead of abrupt changes.
- Noise realism: Gaussian perturbations ($\sigma = 0.01 - 0.02 \times std$) added.

5) Validation results

To verify that the synthetic dataset accurately represents live 5G network behavior, we compared injected sequences to real KPI telemetry using statistical and signal-based metrics. As shown in the Validation Profile (Fig. 3), the synthetic data exhibits high fidelity across key dimensions.

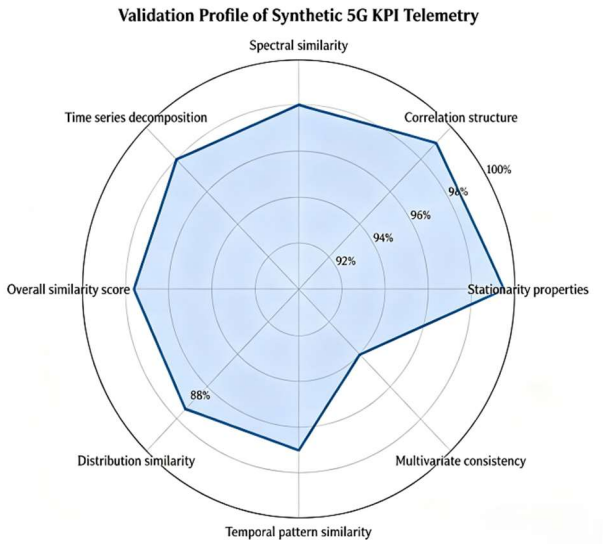

Fig. 3. High-fidelity validation of synthetic 5G KPI data. Comparative analysis showing strong alignment with real-world distributions across spectral, temporal, and correlation dimensions, with a mean similarity index of > 98%.

- Statistical and Spectral Integrity: 99.5% similarity in stationarity and 98.0% in spectral characteristics, preserving 5G traffic frequency-domain traits.
- Structural and Temporal Consistency: 99.0% correlation structure and 98.0% time series decomposition similarity, maintaining KPI interdependencies (e.g., traffic volume and active users).
- Overall Behavioral Fidelity: 97.6% overall score and 97.0% temporal pat tern similarity, providing reliable ground truth for unsupervised and rule-priority evaluations.

This chart indicates that even the lowest metric (multivariate consistency at 94.0%) remains within acceptable industrial ranges. This comprehensive validation ensures that the performance gains observed in Section VI are representative of real-world network dynamics rather than artifacts of the synthetic generation process.

6) Integration with hybrid framework

Within the hybrid anomaly detection framework, the TCAN model is trained on clean operational data using self-supervised learning, enabling multi-scale representation of normal behavior. Evaluation is conducted by injecting controlled synthetic failures into the same baseline data: the TCAN component captures unseen deviations, while the supervised ensemble, trained on labeled synthetic anomalies, identifies known failure types. This dual approach provides both unsupervised adaptability and supervised precision, while also offering ground truth for the evaluation of rule-based detection and fusion mechanisms.

C. Experiment

1) Dataset

The dataset comprises hourly KPI measurements collected from a live 5G network over a 30-day period (May 14, 2025–June 13, 2025). It includes 126,801 records from 66 gNodeBs, each with three cells, yielding 194 unique (gNodeB, Cell_ID) combinations. Five basic key performance indicators (KPIs) were analyzed:

- ENDC SSR (%)—session setup success ratio
- ENDC CDR (%)—call drop ratio
- NR RASR (%)—Random access success ratio
- NSA PS Traffic (GBytes)—total user plane traffic
- Max RRC Connected NR ENDC User—maximum active NR users

Basic profiling revealed a healthy network with occasional extreme fluctuations. For instance, ENDC SSR averaged 97.58% ($\sigma = 9.35$), but exhibited extreme outliers (0% to 200%), indicative of transient failures or data artifacts. These sporadic fluctuations underscore the need for an anomaly detection framework robust to such deviations. NR RASR averaged 78.93%, below the > 90% target, making it the most concerning KPI. Traffic and max user counts were low on average (0.56 GB, 5.73 UEs), reflecting underutilized cells in less populated areas. Approximately 4.3% of values were missing, concentrated in specific cells.

Stationarity tests Azure Data Factory (ADF) across all cells showed that 57.8% of KPI time series were stationary, while 36.7% were inconclusive and 5.4% non-stationary, indicating trend shifts and seasonality must be explicitly modeled.

To illustrate typical temporal dynamics, Signal Temporal Logic (STL) decomposition was performed for one representative gNodeB (gNB00002) across its three cells Fig. 4.

Each KPI displayed clear daily seasonality, visible as periodic oscillations in the seasonal component, while the trend component showed medium-term drifts and transient spikes (e.g., Cell 1 around May 20). Residuals remained highly variable, reflecting non-Gaussian noise and localized anomaly-like events.

Aggregated STL analysis across all cells showed similar patterns. User and traffic KPIs exhibited clear daily seasonality, while control-plane KPIs (SSR, RASR) showed weaker periodicity. Some cells displayed slow trend drifts preceding performance degradation. Residuals were heavy-tailed and non-Gaussian, indicating irregular localized fluctuations beyond simple Gaussian assumptions.

Further multivariate analysis (Fig. 5) confirmed strong KPI interdependencies: traffic volume and active users were highly correlated ($r = 0.715$), and SSR degradation coincided with traffic dips. Principal Component Analysis (PCA) on 25 cell-level features explained 60.2% of the variance in the first three components, with PC1 dominated by load-related KPIs and PC2-PC3 capturing reliability metrics. The PCA scatter plot revealed behavioral clusters (e.g., high- vs. low-load cells) and outliers, highlighting network heterogeneity and the need for adaptive multivariate anomaly detection beyond global thresholds. These findings highlight three key challenges for anomaly detection in 5G networks:

- Seasonality and no stationarity: static thresholds are unreliable.

- Cross-KPI dependencies: multivariate modeling is required.
- Cell-specific variability and residual noise: simple models overfit or miss localized failures.

This motivates a robust multivariate Time Series Anomaly Detection (TSAD) framework capable of handling nonstationary, multi-scale patterns with cross-KPI dependencies, which we address using a hybrid self-supervised TCAN model with rule-priority fusion.

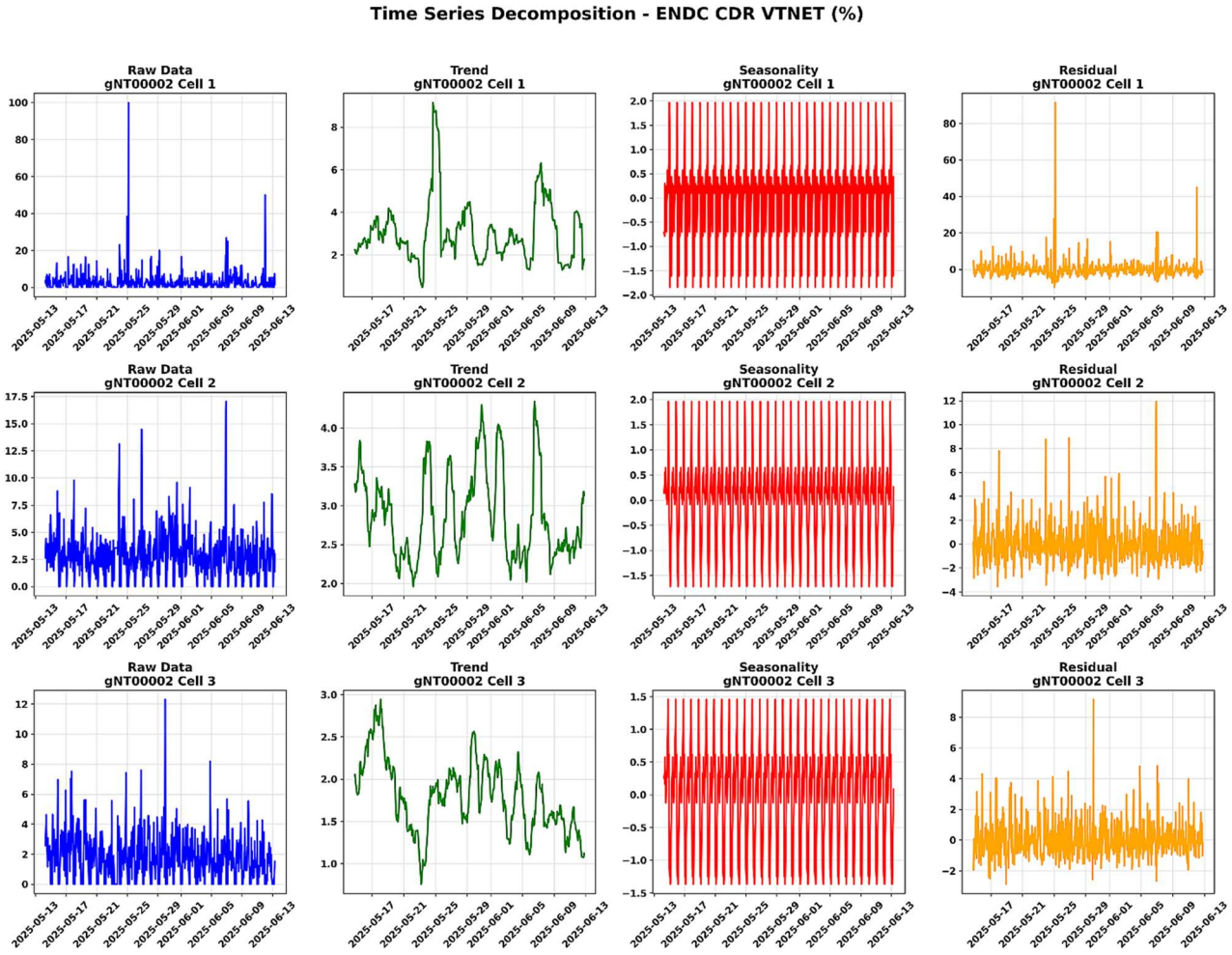

Fig. 4. STL Decomposition of ENDC CDR (%). Multi-cell analysis of gNB00002 isolating raw telemetry into trend, seasonal, and residual components. The overlapping of daily periodicity, medium-term drifts, and stochastic noise validates the requirement for robust, multi-scale temporal anomaly modeling.

2) *Data preprocessing*

A unified preprocessing pipeline ensured temporal consistency and noise reduction:

- Missing-value handling: About 4.3% of values were missing. Forward/backward filling preserved temporal continuity, and global medians were used for fully missing periods.
- Outlier suppression: Extreme outliers (e.g., SSR > 200%) were clipped using a σ-scaled Median Absolute Deviation (MAD) filter to retain realistic variability.
- Local density filtering: A Gaussian KDE fitted over the last 6 valid samples identified sustained anomalies (density ≤ 0.05) versus isolated noise spikes.
- Normalization and segmentation: KPI values were z-score normalized per cell and segmented into 72 h sliding windows (50% overlap) to capture multi-scale temporal patterns for reconstruction, forecasting, and contrastive tasks.

This process produced standardized and temporally aligned sequences suitable for robust anomaly detection across heterogeneous cells.

3) *Baseline models and evaluation metrics*

The proposed hybrid framework was compared with both operational and deep-learning baselines:

- Rules-Only: Static KPI thresholds (e.g., SSR < 90%, CDR > 5%) commonly used in Network Operations Centers.
- SSL-Only (TCAN): Self-supervised TCAN trained on forecasting, reconstruction, and contrastive objectives.
- Ensemble-Only: Supervised classifier (Random Forest and XGBoost) trained purely on labeled synthetic anomalies.
- Fusion Pipeline (Proposed): Rule-priority fusion combining rule-based, SSL, and ensemble outputs.

Additional comparisons included Bi-LSTM, Autoencoder (AE), Variational Autoencoder (VAE),

Temporal Convolutional Network (TCN), Informer, and Anomaly Transformer models.

Performance was assessed using Precision, Recall, F1-Score, ROC-AUC, and Detection Rate.

Ground Truth for Evaluation:

As real 5G data lacks full anomaly labels, we injected synthetic anomalies into real KPI sequences to create reliable ground truth. These included dips, drifts, abrupt shifts, flatlines, and instability patterns, covering diverse failure modes. Labels were perfectly accurate and used for all metric calculations.

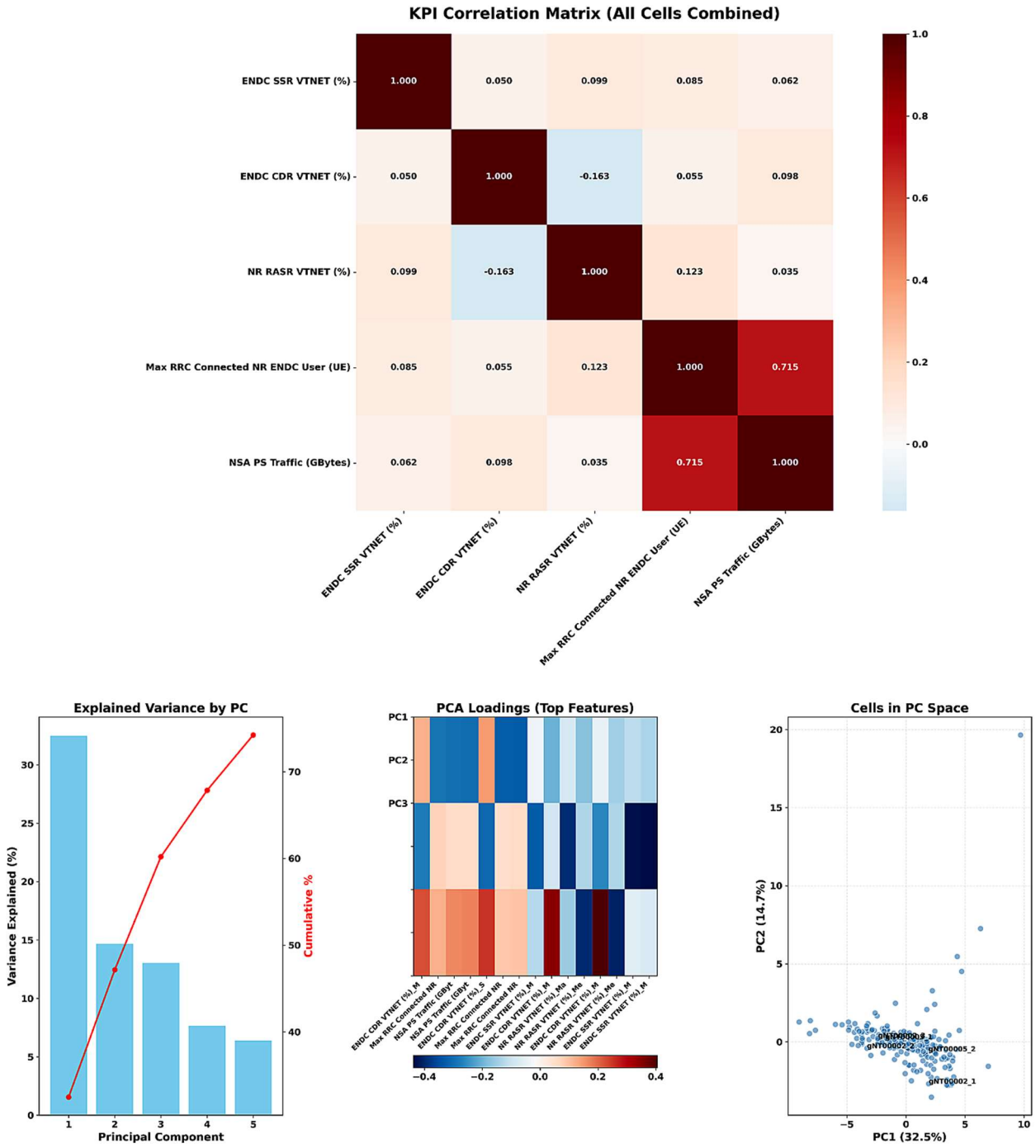

Fig. 5. Multivariate KPI Interdependence and behavioral clustering. (Top) Correlation analysis identifying a high-degree coupling between user density and traffic volume (r = 0.715). (Bottom) PCA results isolating dominant variance components and behavioral clusters, justifying the requirement for adaptive multivariate anomaly detection to handle network heterogeneity beyond global thresholds.

4) Implementation details

All models were implemented in Python, using Pandas for data handling, Scikit-learn for preprocessing and evaluation, PyTorch for the Multi-Task TCAN, and three classifiers XGBoost, Supervised Time Series Forest, and a lightweight 1D CNN) for the supervised ensemble. To capture multi-scale temporal dependencies, KPI sequences

were segmented into 72 h sliding windows with 50% overlap, and three TCAN models were trained with different window lengths: 72 h (short-term), 144 h (medium-term), and 288 h (long-term).

The synthetic dataset for supervised ensemble training was split into 80% training and 20% testing, ensuring unseen synthetic anomalies for evaluation. The Multi-Task TCAN used three dilated convolutional blocks with residual attention, an embedding dimension of 128, and a 24 h forecast horizon. Training employed the AdamW optimizer with learning rate 1×10^{-4}, weight decay 1×10^{-4}, dropout rate 0.3, gradient clipping at norm 1.0, and a CosineAnnealingWarmRestarts scheduler for stable convergence. Anomaly scoring combined reconstruction error (weight 1.0), forecast error (0.5), and embedding deviation (0.3).

For the Smart Supervised Ensemble, hyperparameters such as tree depth, number of estimators, and CNN filter sizes were optimized via grid search. To improve robustness and reduce variance, five bootstrapped ensembles were trained and their predictions aggregated. This setup enabled multi-scale modeling, stable training, and a fair comparison across all baselines and the proposed fusion pipeline.

IV. Result and Discussion

This section presents the experimental results of the proposed anomaly detection pipeline.

A. Quantitative Model Performance

Table II presents the quantitative results of all evaluated models in terms of F1-Score, precision, recall, and Receiver Operating Characteristic-Area Under the Curve (ROC-AUC). The proposed Hybrid TCAN model achieves the highest overall F1-Score of 0.728 and the best ROC-AUC of 0.856, demonstrating its superior capability in identifying anomalies within complex 5G KPI time series.

The 11% improvement over the vanilla TCAN demonstrates the effectiveness of the rule-priority fusion strategy in mitigating the conservative bias of LSTM-based models, which miss approximately 78% of anomalies, as well as the false-positive tendency (precision < 0.35) of rule-based and Autoencoder approaches. By addressing these complementary limitations, the proposed architecture provides a robust and reliable detection framework for non-stationary 5G telemetry.

TABLE II. Performance Comparison of Anomaly Detection Models

Model	F1-Score	Precision	Recall	ROC-AUC
Rule-based	0.426	0.344	0.558	0.774
Bi-LSTM [48]	0.356	0.946	0.219	0.706
Deep-LSTM [49]	0.313	0.883	0.193	0.607
TCN [39]	0.133	0.087	0.274	0.476
Autoencoder [33]	0.367	0.243	0.752	0.741
TCAN [31]	0.656	0.715	0.606	0.826
Hybrid TCAN (ours)	**0.728**	**0.816**	**0.657**	**0.856**

B. Detection Behavior and Agreement Analysis

By segmenting the detection results, we observe how the ML components complement traditional methods. While the rule-based engine captures high-priority service outages, the ML modules identify subtle, non-linear degradations.

Due to the lack of labeled real-world failures in operational 5G networks, we mitigate this limitation by employing a Rule-Based Engine based on "Silver Standard" evaluation that leverages expert experiences. This Rule-Based Engine encodes years encodes years of accumulated Network Operation Center (NOC) from real-world deployments, thereby providing a practical approximation of operation ground-truth. Table III addresses concerns about real-world evaluation by comparing our model results to those of the Rule-Based Engine across 126,801 unlabeled records.

The high agreement rate (98.7%) with expert-defined rules, confirming that the proposed model has effectively internalized 5G domain logic and accurately captures established failure modes in production environment. Since the rule engine encodes years of accumulated Network Operation Center (NOC) expertise refined through real-world deployment, this level of concordance demonstrates that the learned representations capture operationally valid anomaly patterns rather than spurious correlations.

TABLE III. Detection Agreement and Discovery Rates

Detection Source	Total Detections	% of Dataset	Agreement with Rules
Rule-Based Engine	1270	1.0%	100%
TCAN Baseline	11,691	9.2%	90.8%
Hybrid Pipeline	1284	1.0%	98.7%

Further analysis reveals the complementary behavior of rule-based and machine learning-based detection mechanisms:

- Consensus detections (Rule + ML): 660 cases corresponding to clear and unambiguous failures.
- Rule-only detections: 610 cases primarily driven by static threshold violations.
- ML-only detections: 11,375 cases capturing subtle and previously undetected degradations.

The 11,375 ML-only detections, which were missed by the rule-based engine, highlight the framework's capability to identify early-stage sleeping cells before they escalate into severe service degradation. These alerts likely correspond to pre-failure states characterized by subtle temporal deviations, cross-KPI coupling effects, or gradual performance drifts that remain below static threshold limits. Consequently, the ML components function as a discovery mechanism for proactive fault detection rather than a direct replacement for expert-defined rules.

C. Final Pipeline Output and Confidence Calibration

The final integrated pipeline produced 1.284 total detections (≈ 1.0% anomaly rate), distributed by confidence level as follows:

- High confidence (> 0.7): 267 detections
- Medium confidence (0.4 to 0.7): 1017 detections
- Low confidence (≤ 0.4): 125,517 detections

The adaptive threshold (0.4) enables fine-grained sensitivity control, effectively filtering low-impact events to reduce operator fatigue while ensuring that critical degradations remain visible.

The evaluation spans 194 cells across 66 gNodeBs, exhibiting diverse operational characteristics, including varying stationarity levels (5.4% non-stationary), traffic ranges of 0.02–5.8 GB, and Server-Side Rendering (SSR) outliers (0–200%). Despite this heterogeneity and 4.3% missing data, the Hybrid TCAN achieves a consistent F1-Score of 0.728 without cell-specific tuning. This setting serves as an implicit stress test, demonstrating the robustness and generalization capability of the proposed framework under real-world conditions beyond controlled synthetic scenarios.

D. Discussion

The experimental results demonstrate that the Hybrid TCAN framework provides a superior balance between high-recall anomaly discovery and high-precision operational filtering.

1) Architectural superiority: TCAN vs. RNNs

The Hybrid TCAN (F1: 0.728) outperforms LSTM-based models (F1: 0.313–0.356) due to its handling of 5G temporal dependencies. Unlike LSTMs, which face recurrent bottlenecks and gradient vanishing over long sequences (72–288 h), TCAN's dilated causal convolutions enable a vast receptive field capturing 24 h and weekly cycles without compression loss. The Multi-Head Attention further focuses on salient shifts and spikes, ignoring heavy-tailed noise, and excels at detecting gradual drifts over 48–96 h that LSTMs miss.

2) Multi-objective learning and robustness

The hybrid approach combines reconstruction, forecasting, and contrastive learning to overcome single-criterion limitations. Autoencoders yield high recall (0.752) but low precision (0.243), while our triple-objective loss flags anomalies only upon simultaneous violations: forecasting detects abrupt shifts, reconstruction captures structural breaks, and contrastive filters noise via embedding consistency. This reduces false positives while preserving recall. For non-stationary KPIs (36.7% inconclusive), the contrastive head ensures invariance to seasonal distortions, mitigating false alarms in heavy-tailed residuals.

3) Specialized attention and failure mode capture

TCAN's four specialized attention heads decompose 5G behavior interpretably, detecting rule-missed failures like oscillatory instability. The Temporal Dependency Head flags abnormal autocorrelation by concentrating attention on oscillatory subsequences, even when individual KPIs remain within thresholds.

4) Rule-priority fusion: Balancing safety and discovery

The asymmetric fusion prioritizes rules with $\lambda r \in [0.5, 1.0]$ based on severity, ensuring critical alerts dominate. This yields 610 rule-only detections for threshold violations, 11,375 ML-only for pre-failure states, and 660 consensuses for unambiguous failures.

5) Bridging the synthetic-real gap

Synthetic injections preserve 97–99% statistical similarity to real data, ensuring label reliability. On 126,801 real records, 98.7% rule agreement validates operational readiness as a reliable assistant. The 11,375 ML-only detections uncover subtle degradations like early "sleeping cells," extending beyond heuristics.

Despite notable advancements, the absence of fully labeled real-world anomalous datasets remains a persistent challenge in operational environments. Further work will incorporate real-world failures confirmed by operators to validate the accuracy of the detection under genuine operational conditions.

V. Conclusion

In this work, we presented an enhanced anomaly detection pipeline for 5G KPI time series, integrating self-supervised temporal attention convolutional networks (TCAN) with rule-based heuristics and ensemble fusion. Experimental results demonstrate that the proposed model achieves a higher F1-Score compared to conventional models such as LSTM, autoencoders, and baseline TCNs. Moreover, the ensemble approach maintains high agreement with existing rule-based methods (≥ 90%) while expanding anomaly coverage through ML-only detections. The final integrated pipeline achieves robust performance, detecting rare anomalies with high confidence while effectively suppressing low-impact events. Stress testing with synthetic anomalies confirms the stability of the ensemble, achieving near-perfect recall (0.991) and precision (1.000). These findings highlight the potential of combining self-supervised temporal modeling with adaptive thresholding to improve early detection of sleeping cells and network degradations in 5G networks.

Despite promising results, several challenges remain. The system depends heavily on historical data, making it sensitive to major network changes such as upgrades or shifting traffic patterns, which may require retraining or online adaptation. Moreover, it currently monitors only five key KPIs (e.g., ENDC SSR, NR RASR), leaving potential blind spots for subtle or localized faults that may appear only in other KPIs or raw counter data.

Future work will address these challenges by expanding KPI coverage to include metrics like latency, Physical Review (PRB) utilization, and handover performance. A more advanced approach involves learning directly from hundreds of raw counters using representation learning to uncover hidden "latent KPIs". In the long term, the pipeline can be integrated into semi-autonomous workflows and eventually evolve into a fully closed-loop

self-healing system that detects, diagnoses, and automatically mitigates network degradations.

These future directions pave the way toward fully autonomous, self-optimizing 5G and beyond-5G networks, improving operational efficiency, fault resilience, and overall user experience.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Le Nhu Quynh conducted the theoretical research, developed the main ideas, implemented the experimental program, and drafted the manuscript; Truong Duc Tai proposed the initial concept and provided the dataset; Truong Duc Tai, Dinh Thi Phuong, Nguyen Anh Tu, and Tran Van Tung evaluated the results and the manuscript, providing comments and guidance throughout the study; All authors read and approved the final manuscript.

ACKNOWLEDGMENT

The authors wish to thank Broadband Wireless Center, Viettel High Technology Industries Corporation, Viettel Groups for providing the real-world 5G RAN dataset and the experimental environment necessary for this study.

REFERENCES

[1] ISO/IEC. (2017). Information technology—Systems and software Quality Requirements and Evaluation (SQuaRE)—Service quality models. [Online]. Available: https://www.iso.org/standard/35735.html

[2] T. P. S. Documentation, "Information technology security techniques information security management systems requirements," 2005.

[3] F. Wang *et al*., "A survey of deep anomaly detection in multivariate time series: Taxonomy, applications, and directions," *Sensors*, vol. 25, no. 1, 190, 2025.

[4] S. Chernov *et al.*, "Anomaly detection algorithms for the sleeping cell detection," in *Proc. 2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall),* 2015, pp. 1–5.

[5] O. M. Salazar *et al*., "A machine learning framework for sleeping cell detection," *IEEE Access*, vol. 8, pp. 155180–155192, 2020.

[6] H. Zhuv *et al*., "Learning spatial graph structure for multivariate kpi anomaly detection in large-scale cyber-physical systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[7] H. Yu *et al*., "Renyientropy-driven network traffic anomaly detection with dynamic threshold," *Cybersecurity*, vol. 7, no. 1, 2024.

[8] T. Klinsuwan *et al*., "Evaluation of machine learning algorithms for supervised anomaly detection and comparison between static and dynamic thresholds in photovoltaic systems," *Energies*, vol. 16, no. 4, 2023.

[9] B. A. A. Ahmadi *et al*., "99% false positives: A qualitative study of SOC analysts' perspectives on security alarms," in *Proc. 31st USENIXSecurity Symposium (USENIXSecurity 22)*, 2022, pp. 2783–2800.

[10] S. Tariq *et al*., "Alert fatigue in security operations centres: Research challenges and opportunities," *ACM Computing Surveys*, vol. 57, no. 9, 2025.

[11] S. Tariq *et al*., "Towards human-AI teaming to mitigate alert fatigue in security operations centres," *ACM Transactions on Internet Technology*, vol. 24, no. 3, 2024.

[12] D. Sacher, "Finger pointing false positives: How to better integrate continuous improvement into security monitoring," *Digital Threats: Research and Practice*, vol. 1, no. 1, 2020.

[13] A. S. Yaro *et al*., "Outlier detection in time-series receive signal strength observation using Z-score method with Sn scale estimator for indoor localization," *Applied Sciences*, vol. 13, no. 6, 3900, 2023.

[14] V. Chandola *et al*., "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[15] V. Kozitsin *et al*., "Online forecasting and anomaly detection based on the ARIMA model," *Applied Sciences*, vol. 11, no. 7, 3194, 2021.

[16] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," in *Proc. 31st USENIXSecurity Symposium (USENIXSecurity 22)*, 2022, pp. 2783–2800.

[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[18] S. S. Namini *et al*., "A comparison of ARIMA and LSTM in forecasting time series," in *Proc. 17th IEEE International Con ference on Machine Learning and Applications (ICMLA)*, 2018, pp. 1394–1401.

[19] H. Xu *et al*., "Anomaly transformer: Time series anomaly detection with association discrepancy," arXiv preprint arXiv: 2110.02642, 2021.

[20] N. Wu *et al*., "Deep transformer models for time series forecasting: The influenza prevalence case," arXiv preprint arXiv: 2001.08317, 2020.

[21] K. Zhang *et al*., "Self-supervised learning for time series analysis: Taxonomy, progress, and prospects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 10, pp. 6775–6794, 2024.

[22] H. Choi and P. Kang, "Multi-task self-supervised time-series representation learning," *Information Sciences*, vol. 671, 120654 2024.

[23] J. Yue *et al*., "TS2Vec: Towards universal representation of time series," in *Proc. AAAI Conference on Artificial Intelligence*, vol. 36, pp. 8980–8987, 2022.

[24] J. Moysen *et al*., "Big data-driven auto mated anomaly detection and performance forecasting in mobile networks," in *Proc. 2020 IEEE Global Communications Conference (GLOBE COM) Workshops*, 2020, pp. 1–5.

[25] A. Zoha *et al*., "A machine learning framework for detection of sleeping cells in LTE network," in *Proc. Machine Learning and Data Analysis Symposium*, 2014.

[26] J. Paparrizos *et al*., "TSB-UAD: An end-to-end benchmark suite for univariate time-series anomaly detection," in *Proc. VLDB Endowment*, vol. 15, no. 8, pp. 1697–1711, 2022.

[27] S. Schmidl *et al*., "Anomaly detection in time series: A comprehensive evaluation," in *Proc. VLDB Endowment*, 2022, vol. 15, no. 9, 1779.

[28] A. D. Salman *et al*., "Hybrid LLM-assisted fault diagnosis framework for 5G/6G networks using real-world logs," *Computers*, vol. 14, no. 12. 551, 2025.

[29] L. Zhen *et al*., "Anomaly detection model in network security situational awareness based on machine learning: Limitation, techniques, and future trends," *IEEE Access*, vol. 13, pp. 126084–126129, 2025.

[30] Q. Liu *et al*., "TSB-AutoAD: Towards automated solutions for time-series anomaly detection," in *Proc. VLDB Endowment*, vol. 18, no. 11, pp. 4364–4379, 2025.

[31] Y. Lin *et al*.," Temporal convolutional attention neural networks for time series forecasting," in *Proc. 2021 International Joint Conference on Neural Networks (IJCNN)*, 2021, pp. 1–8.

[32] I. T. Jolliffe. (2002). Principal component analysis. [Online]. Available: https://doi.org/10.1007/b98835

[33] G. E. Hinton *et al*., "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, pp. 504–507, 2006.

[34] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," arXiv preprint arXiv: 1312.6114, 2013.

[35] Y. Su *et al*., "Robust anomaly detection for multivariate time series through stochastic recurrent neural network," in *Proc. 25th ACM SIGKDD International Conference on Know Edge Discovery and Data Mining*, 2019, pp. 2828–2837.

[36] A. B. Garcia *et al*., "A review on out lier/anomaly detection in time series data," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–33, 2021.

[37] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[38] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
[39] S. Bai *et al*., "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," arXiv preprint arXiv: 1803.01271, 2018.
[40] H. Zhou *et al*., "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proc. AAAI Conference on Artificial Intelligence*, 2021, vol. 35, pp. 11106–11115.
[41] H. Wu *et al*., "Timesnet: Tem poral 2D-variation modeling for general time series analysis," arXiv preprint arXiv:2210.02186, 2022.
[42] A. Z. Sellam *et al*., "MAAT: Mamba adaptive anomaly transformer with association discrepancy for time series," arXiv preprint arXiv:2502.07858, 2025.
[43] S. Ma *et al*., "Tpad: Temporal-pattern-based neural network model for anomaly detection in multivariate time series," *IEEE Sensors Journal*, vol. 23, no. 24, pp. 30668–30682, 2023.
[44] M. A. Mulia *et al*., "Kbjnet: Kinematic bi-joint temporal convolutional network attention for anomaly detection in multivariate time series data," *Data Science Journal*, vol. 23, no. 12, 2024.
[45] Z. Z. Darban *et al*., "Carla: Self supervised contrastive representation learning for time series anomaly detection," *Pattern Recognition*, vol. 157, 110874, 2025.
[46] S. Lee, T. Park, and K. Lee, "Soft contrastive learning for time series," arXiv preprint arXiv: 2312.16424, 2023.
[47] Z. Zeng *et al*., "Timesurl: Self-supervised contrastive learning for universal time series representation learning," arXiv preprint arXiv: 2312.15709, 2023.
[48] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, vol. 18, pp. 602–610, 2005.
[49] R. Pascanu *et al*., "How to construct deep recur rent neural networks," arXiv preprint arXiv: 1312.6026, 2013.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).