Computer Science and Information Technologies
Vol. 7, No. 3, November 2026, pp. 256~270
ISSN: 2722-3221, DOI: 10.11591/csit.v7i3.p256-270

Convolutional neural network and long short-term memory forecasting and variational autoencoder anomaly detection in 4G cellular networks

Ruvarashe C. Hove[1], Eng Mainford Mutandavari[2]
[1]Department of Data Science and Analytics, School of Information Science and Technology, Harare Institute of Technology, Harare, Zimbabwe
[2]Department of Data Science and Business Systems, SRM Institute of Science and Technology, Kattankulathur, India

Article Info

Article history:

Received May 5, 2026
Revised Jun 12, 2026
Accepted Jul 13, 2026

Keywords:

Anomaly detection
Cellular traffic forecasting
CNN–LSTM
Reproducible research
RF spectrum management
Synthetic dataset
Variational autoencoder

ABSTRACT

Spectrum monitoring in cellular networks with limited resources remains predominantly reactive because capacity planning and failure detection occur after network congestion, while spectrum monitoring data are rarely available for reproducible research. Existing studies address short-term traffic forecasting and anomaly detection as separate tasks and largely depend on densely sampled operator telemetry data that are inaccessible to academia and regulators in emerging markets. This study integrates both tasks into a unified deep-learning pipeline for a 4G cellular environment. The Zimbabwe spectrum dataset contains 315,247 hourly measurements collected from 13 cell sites, three operators, and five frequency bands, annotated with four operator-defined anomaly classes representing 2.03% of all measurements. A hybrid one-dimensional (1-D) convolutional neural network-long short-term memory (CNN–LSTM) model uses a 72-hour traffic window to forecast the next six hours of aggregate network traffic, while a variational autoencoder (VAE) trained exclusively on normal records detects anomalies when reconstruction error exceeds the 99th-percentile validation threshold. On the held-out test set, the model achieved a mean absolute error (MAE) of 158.97 GB, root mean square error (RMSE) of 230.41 GB, and mean absolute percentage error (MAPE) of 52.93%, outperforming a seasonal-naive baseline (MAE=216.43 GB, MAPE=66.77%, $p<0.001$, Diebold-Mariano test). The study contributes a publicly available localized synthetic dataset, a reproducible end-to-end forecasting and anomaly detection baseline, and a per-class anomaly analysis.

Corresponding Author:

Ruvarashe C. Hove
Department of Data Science and Analytics, School of Information Science and Technology
Harare Institute of Technology
P.O. Box BE 277, Belvedere, Harare, Zimbabwe
Email: h240618f@hit.ac.zw

1. INTRODUCTION

Radio frequency (RF) spectrum is a finite resource, controlled by jurisdiction and the utilization efficiency has a direct relationship with the service cost, quality and coverage area of cellular service. Mobile data traffic has increased significantly over the last decade, fueled by the penetration of 4G handsets, more bandwidth-hungry applications, and the shift of voice traffic to data-bearing services, across emerging markets. In Zimbabwe, the Zimbabwe National Frequency Allocation Plan (ZNFAP) [1] allocates the spectrum and in the Zimbabwean mobile sector, mobile data traffic has been steadily growing over the last

few years in terms of percentage and volume as documented in the abridged Postal and Telecommunications Sector Performance Report (PTSPRs) [2] and it has been noted that a growing percentage of this traffic is on the lower end of the spectrum bands and that the quality of service (QoS) has been under strain on a regular basis on the urban cell. Fixed-allocation regimes provide little information to operators and regulators about emerging congestion, interference and unauthorized transmissions until it affects subscribers and curated operator traces that can be made available for academic study are less common.

There are then two distinct but interdependent operational tasks which are very important both to the regulator and to the operator. The first is short-horizon forecasting of network traffic, to base decisions on capacity planning, scheduling and load-balancing on a quantitative load estimate instead of the operator's intuition [3]–[5]. The second is anomaly detection of the same telemetry without any labelled training data, such that interference, traffic surges, signal dropouts and unauthorized emissions can be identified for review [6]–[9]. These two tasks are usually done by two different pipelines nowadays, they give point forecasts or a detection algorithm alerts but does not explain why reality does not match prediction and does not give any anticipatory load signal. In emerging markets, with limited on-call engineering capacity, operators are especially in need of a unified view to present the learnt normal regime as well as deviation from that regime in a single decision loop.

Three gaps of concrete remain. First, the hybrid deep forecasting architectures presented in the wireless-networking literature [4], [5], [10], [11] have predominantly been assessed on dense operator-internal traces where they are unavailable to the general research community, making it difficult to reproduce even solid ideas. Second, reconstruction-based anomaly detectors, [6]–[9], [12]–[16] tend to be evaluated on benchmarks that lump detector performance together with dataset idiosyncrasy, and very few studies report recall per anomaly class, the measure of interest when triaging operations. Third, data, infrastructure and policy are cited consistently as the “five enforcers” of African and other emerging market deployments in artificial intelligence (AI)-for-spectrum-management surveys [17]–[20], with very few published pipelines calibrated against publicly verifiable national-level statistics and released end to end. This paper directly tackles those three gaps.

In this paper, a deep learning based two-component pipeline is developed and evaluated. The forecaster model used is a hybrid model of one-dimensional (1-D) convolutional neural network (CNN) and long short-term memory (LSTM) which forecasts traffic for the next 6 hours given 72 hours sliding window input. The trained variational autoencoder (VAE) is only trained on normal records and then detects the anomalies using the reconstruction error, using the 99th percentile of the validation reconstruction errors as a fixed decision threshold. Both models are assessed against the Zimbabwe spectrum dataset which is a synthetic dataset developed with a footprint reflecting the 13-tower, 5-band, 3-operator spectrum in Zimbabwe and calibrated using the publicly verifiable Postal and Telecommunications Regulatory Authority of Zimbabwe (POTRAZ) sector statistics. It is an end-to-end reproducible baseline, not a detector suitable for deployment-this candor is a key aspect of the contribution.

Two characteristics are key to understanding the novelty value of this work, relative to the CNN–LSTM forecasters [5], [11] and the previous reconstruction-based detectors [9], [14], [15], [21]: i) forecasting and detection in one box, where the outputs of the hybrid forecaster are intended to be consumed in a single operations dashboard, rather than two and ii) detectability decomposition per class, revealing which of the four anomaly classes can or cannot be recovered through the manifold-departure assumption, with practical operational implications for hybrid first-pass-filter deployments.

This paper makes the following contributions: i) reproducible localized dataset, we document, parameterize and publish a synthetic set of 315,247 records from 13 cell sites, covering three operators in five bands, with four operator style anomaly classes with a total rate of 2.03%, based on publicly available POTRAZ sector statistics; ii) baseline for end-to-end forecasting and detection, we combine a forecaster based on a hybrid CNN–LSTM approach with a detector based on a VAE in the same streaming pipeline, providing preprocessing, chronological splits, fixed thresholds, and full reporting of metrics, hyperparameters, sensitivity analysis, and confusion matrices all in a reproducible manner; iii) comprehensive baseline comparison. We compare the CNN–LSTM to naive forecasters, seasonal-naive forecasters, autoregressive (AR) (24) forecasters and feed-forward deep neural network (DNN) forecasters (which excludes LSTMs and gated recurrent unit (GRU)) and perform paired statistical tests and 95% confidence intervals for the LSTM-only, GRU and CNN-only forecasters. Similarly, we compare the VAE with statistical thresholding, Isolation Forest, one-class support vector machine (SVM) and vanilla autoencoder. We also compare our results with the recent literature ConvLSTM and Transformer based forecast baselines [22], [23]; iv) truthful representation and declaration of trade-offs. We report the behaviour of the VAE in a high precision/low recall regime and the performance per anomaly type, including bootstrap confidence intervals, and discuss the limits on synthetic localisation, viewing the system as a first-pass operational filter rather than a deployment-grade detector. Additionally, we suggest specific approaches to boost recall for the more difficult classes unauthorized and signal dropout; and v) broader computer

science (CS) relevance. Given the hourly aggregation of key performance indicator (KPI) streams and the light weight-out per sample, the pipeline can be deployed in a distributed way across operator and regulatory boundaries (edge-pre-filter/cloud-trainer split), be extended in a federated way preserving operator confidentiality, and can be integrated with internet of things (IoT) based wide-area spectrum monitoring connections we develop from section 5.

The rest of the paper is arranged as follows. A review of relevant work in the areas of cellular forecasting, deep wireless prediction, reconstruction-based anomaly detection, AI for spectrum management, and synthetic localized datasets is provided in the next section. Section 2 introduces details on the dataset, localization procedure, architecture of CNN–LSTM and VAE, hyperparameter selection, training and evaluation procedure, and reproducibility setting. Results of forecasting and anomaly detection, statistical analysis, sensitivity analysis, and a discussion about the detectability per class are presented in section 3. The limitations are listed in section 4, and the conclusion is summarized in section 5 with future works towards 5G, IoT and federated extensions.

– Hybrid deep model for forecasting cellular traffic

The multi-scale temporal structure associated with cellular traffic - the diurnal cycle within-day, weekly cycle, holiday effects, and longer-term growth-is partially described by classical statistical methods like autoregressive integrated moving average (ARIMA) and exponential smoothing [3]. Models with deep recurrent structures (LSTM, GRU) [10], [11], have now become standard building blocks and hybrid CNN–LSTM models have demonstrated improvement over either component alone on long-term evolution (LTE) traffic and multi-cell load [4], [5]. Convolutional layers capture local shapes morning ramps, evening peaks, short bursts as a small kernel are rolled across the time axis, and recurrent layers are used to pass longer-range temporal context through the 72-hour input window. In other areas, spatiotemporal extensions like ConvLSTM [22] have been proposed and they are particularly suitable for multi-cell grids; transformer-based forecasters [23] have demonstrated effective performance on long-horizon series in other fields, but they require much more data and computing power than the scope of this paper. The CNN–LSTM design which is utilized here is in the tradition of hybrid designs, with the two convolutional blocks followed by two-stack LSTM.

– Reconstruction-based anomaly detection

Operator anomaly inventories are not often completely labeled, and even when they are, many labels tend to be noisy and/or biased towards visual failure modes. Thus the literature has settled to unsupervised reconstruction based detectors [6]–[9]. In line with this, the VAE of Kingma and Welling [6] encourages the latent space to be closer to a Gaussian prior, while the reconstruction-probability framing proposed by Islam *et al*. [7] for 5G is prevalent in the cellular domain [14], [15]. Recently, multivariate time series have seen competitive performance by anomaly detectors such as the anomaly transformer [13] significantly higher compute cost. A consistent experimental finding over all the methods in reconstruction-based methods is that high percentile thresholds lead to high precision and low (well lower than what a supervised method with the same evidence could achieve) recall [16], [17]. We make the pattern explicit for cellular anomalies by our per-class analysis in section 3.2.

– Reproducible emerging-market datasets and AI for spectrum management

Spectrum sensing and dynamic spectrum access for AI systems make the case for having small, adapted models with a strong focus on regulatory compliance and infrastructure limitations [18]–[20]. Availability of data, heterogeneity of infrastructure and policy are the binding constraints in all empirical studies of the deployment of African and other emerging-market networks [19], [21], [24], [25] and even capacity-building activities by International Telecommunication Union Development Sector (ITU-D) have included them as first-order requirements in systems aimed at informing regulatory decisions [19]. In network research where production traces are kept confidential and benchmark datasets get outdated very rapidly, the use of synthetic datasets has a long tradition in research: any synthetic dataset should be calibrated to verifiable real-world data, be reproducible, and be clearly designated as synthetic traces of the operators [1], [2], [19]. The framework described in this paper fits in this tradition while being based on publicly available sector-level statistics (POTRAZ's traces) instead of confidential operator-internal traces and is easily scalable to distributed and federated deployments across operator boundaries, an increasingly important architectural challenge as 5G and IoT era multi-stakeholder spectrum sharing evolves [24], [25].

2. METHOD

2.1. Notation and problem statement

Let $x(c, b, o, t) \in \mathbb{R}^d$ denote the feature vector observed at cell c, band b, operator o, and discrete hour t. The components of x are received signal strength indicator (RSSI) dBm, channel occupancy (%),

traffic volume (GB/hour), and interference level (dB), augmented with cyclic time-of-day and lag features described in section 2.3.4. Two complementary tasks are then defined on this representation. The forecasting task learns a mapping:

$$\hat{y}(t+1:t+H) = f\theta(x(t-L+1:t))$$

with input window L=72 hours and forecast horizon H=6 hours. Conceptually, the forecaster reads three full days of history and projects the next quarter-day of network load. The anomaly task learns an unsupervised reconstruction gϕ such that the reconstruction error is small on normal records and exceeds a fixed threshold τ on anomalous ones. Intuitively, e(x) measures how far a record sits from the manifold of normal operation that the VAE has learned to reproduce, and the threshold τ is the decision boundary which a record is flagged. Section 2.5 fixes τ at the 99th percentile of validation reconstruction errors. Choosing τ from validation rather than test data follows standard reconstruction-based practice [6], [7].

$$e(x) = \| x - \hat{x} \|^2$$

2.2. System overview

Figure 1 summarizes the end-to-end pipeline, and Figure 2 details the localization stage that produces the input data. The pipeline contains four stages: i) ingestion of the Zimbabwe spectrum dataset described in section 2.3.1; ii) preprocessing and sliding-window construction described in section 2.3.4; iii) parallel CNN–LSTM forecasting (section 2.4) and VAE anomaly detection (section 2.5); and iv) a unified reporting layer that consumes both model outputs into a single set of operational artefacts (section 2.6). The two model branches share the same preprocessed feature representation but use independent training loops; this design choice is deliberate, as it allows the forecaster to be retrained on aggregate traffic alone while the detector continues to operate on the per-record feature space.

Figure 1. End-to-end forecasting and anomaly-detection pipeline. Inputs flow from the Zimbabwe spectrum dataset through preprocessing into two parallel branches a CNN–LSTM forecaster on the network-aggregate hourly series and a VAE detector on the per-record feature space whose outputs feed a unified reporting layer

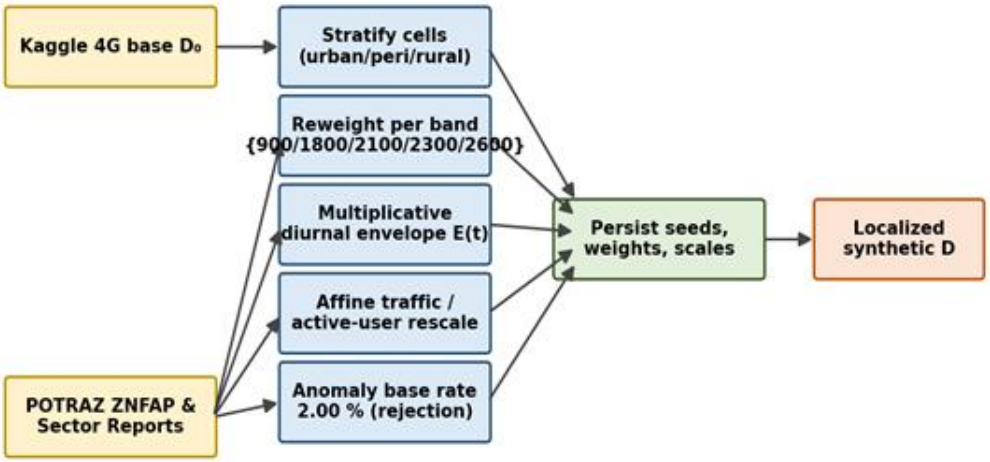

Figure 2. Synthetic localization workflow against publicly published POTRAZ sector statistics

2.3. Dataset and synthetic localization workflow

2.3.1. Dataset overview

The Zimbabwe spectrum dataset comprises 315,247 hourly records spanning 13 cell sites at the country's main urban and regional centres (Harare CBD, Harare West, Harare East, Bulawayo CBD, Bulawayo South, Chitungwiza, Mutare CBD, Gweru CBD, Kwekwe CBD, Kadoma CBD, Masvingo CBD, Victoria Falls, and Livingstone), three mobile operators (Econet, NetOne, and Telecel), and five frequency bands (700, 850, 900, 1800, and 2600 MHz). Each record reports a timestamp, cell tower, operator, band, density type (urban/suburban/rural), traffic volume in GB, RSSI in dBm, channel occupancy (%), interference level, hour of day, day of week, an anomaly flag, and a four-class anomaly type label. Descriptive statistics for the primary modelling features are reported in Table 1, and the anomaly composition is reported in Table 2.

Table 1. Descriptive statistics of primary modelling features

Feature	Mean	Std. Dev.	Min	Max
Traffic (GB/hour)	423.93	297.20	30.31	6,031.46
RSSI (dBm)	−85.24	20.49	−140.00	−50.00
Channel occupancy (%)	52.83	24.69	10.00	100.00

Table 2. Anomaly counts and overall rate

Anomaly class	Count	Share of anomalies (%)	Share of total (%)
Unauthorized	1,660	25.94	0.527
Signal Dropout	1,623	25.36	0.515
Traffic Surge	1,599	24.98	0.507
Interference	1,518	23.72	0.482
Total anomalies	6,400	100.00	2.030
Total records	315,247		100.00

2.3.2. Synthetic localization workflow

The dataset is validated with official POTRAZ sectoral data [1], [2]. The five components are: i) urban/suburban/rural mix based on coverage statistics; ii) weights attached to record-counts based on published utilization tables; iii) a bi-modal diurnal envelope (peak around 10:00, trough around 21:00, peak-to-trough ratio ~2.6, weekend amplitude ~0.88) based on the qualitative QoS commentary from the sector reports; iv) affine rescaling of traffic and active-user features to match the marginal means and dispersions of the sector report aggregates; and v) rejection sampling to adjust the overall anomaly rate to 2.00%. The procedure is deterministically recorded in Algorithm 1.

Algorithm 1. Localized synthetic dataset construction
Input: Kaggle 4G base D_0; POTRAZ statistics P; random seeds s
Output: localized synthetic dataset D
1. Initialize random state with seeds s
2. Stratify cells into urban/suburban/rural using density weights from P
3. Reweight records by band {700, 850, 900, 1800, 2600} MHz
4. Fit diurnal/weekly envelope E(t) from peak-hour patterns in P
5. Apply multiplicative scaling: traffic ← traffic · E(t)
6. Affine-rescale traffic and active-user features
7. Inject 4-class anomalies at 2.00% rate via rejection sampling
8. Append cyclic and lag features (section 2.3.4)
9. Persist D with metadata (seeds, weights, envelopes, scales)
10. Return D

A 2% anomaly rate is used, which needs to be explicitly justified. The published incident-reporting tables that are used to create the localization indicate a consistent pattern of (localised) congestion and dropped call events and quality-of-service incidents that, when expressed as a proportion of the cell-hours monitored, fall within a band of 1.5-2.5% across all reported quarters as detailed in POTRAZ quarterly sector reports [2]. Choosing a rejection sampling rate of 2.00% (2.030% after class balancing) forces the synthetic stream to be within the range that is likely to be encountered in an operations centre when using real telemetry, but still rare enough to test the assumption of unsupervised detection. The four classes of anomaly (interference, traffic surge, signal dropout, and unauthorized) and the approximate balance between them mirror those found in operator-type anomaly inventories and the ones that the regulator will most likely require to be flagged.

2.3.3. Ethical and privacy considerations

All inputs to this study are public. No personally identifiable information, customer identifiers, billing data, or operator-internal traces were accessed at any stage. Cell identifiers are for publicly known city-level locations and not for confidential, operator-internal cell IDs. Further ethics, privacy and regulatory review will be needed for any expansion of the methodology to operator internal traces.

2.3.4. Preprocessing

Forward- and backward-fill are used for short gaps, linear interpolation is used for longer gaps. Records that are duplicated or include improbable outliers are discarded. Timestamps are broken apart into components of hour-of-day, day-of-week and month-of-year and sent through sine and cosine transforms to retain their cyclic nature. As common time-series feature-engineering practice [26], the lag features include features from the previous hour, the same hour the day before, and the same day the week before. Z score is used to normalize all numeric features. The time series is re-arranged in a supervised manner, with a 72-hour time window and a 6-hour forecast horizon, to perform the forecasting task and ensure no information from the future is leaked into the training.

2.4. CNN–LSTM forecaster

The forecaster is a 1-D CNN–LSTM hybrid, as shown in Figure 3, with hyper parameters shown in Table 3. Naturally, local temporal features like morning ramps, evening peaks and short bursts are extracted with two 1-D convolutional layers (with 64 and 32 filters, kernel size 3, rectified linear unit (ReLU) activation, and a same padding) and batch normalization [27] respectively. The idea is simple: a small kernel moving along the hourly axis is kind of like an "eye" that can be trained to detect these recurring sub-day motifs, and precisely that is the sort of thing that purely recurrent models can be slow to capture. After the convolutional block, there is a max-pooling (2, 2) and a dropout (0.2) layer. The longer-range dependencies are then captured across the 72-hour window by two stacked LSTM layers with 64 and 32 hidden units, respectively, and 0.2 dropout used after each. The 6-hourly forecasts are obtained in a single feed-forward pass using a 32-unit ReLU dense layer followed by a 6-unit linear output head. Training with Adam (learning rate 1×10^{-3}), mean squared error (MSE) loss, early stopping (patience 10), and ReduceLROnPlateau (factor 0.5, patience 5).

Figure 3. Hybrid 1-D CNN–LSTM architecture for six-hour-ahead traffic forecasting. Convolutional blocks (left) capture local within-day shape; stacked LSTM layers (centre) carry longer-range temporal context; a dense head (right) produces the six-step forecast

Table 3. CNN–LSTM hyperparameters and search space

Component	Final setting	Search space (validation only)
Input window L	72 hours	{24, 48, 72, 96, 168}
Forecast horizon H	6 hours	Fixed by task
Conv1D filters	64→32	{32, 64, 96}
Conv1D kernel/padding/activation	3/same/ReLU	Fixed
Batch normalization	after each Conv1D	{on, off}
MaxPool/dropout	Pool 2/Dropout 0.2	{0.1, 0.2, 0.3, 0.4}
LSTM hidden units	64→32	{32, 64, 96}
Dense head	Dense (32, ReLU)→Dense (6, linear)	Fixed
Optimizer/learning rate	Adam/1×10^{-3}	{1×10^{-3}, 5×10^{-4}, 1×10^{-4}}
Loss	Mean squared error	Fixed
Batch size	64	{32, 64, 128}
Max epochs/early stopping	50/patience 10	Fixed
LR schedule	ReduceLROnPlateau (factor 0.5, patience 5, min lr 1×10^{-6})	Fixed
Random seed (TF)	42	Fixed

Validation only grid search across the search space summarized in Table 3 (right column) was used to choose hyperparameters. In particular, we searched over 32, 64 and 96 convolutional filter counts, 32, 64, and 96 LSTM hidden units, 0.1, 0.2, 0.3, and 0.4 dropout rates and 1×10^{-3}, 5×10^{-4}, 1×10^{-4} learning rates, testing each of the candidate model configurations on the 15% chronological validation split only.

The reported architecture (64-64 conv, 64-32 LSTM, dropout 0.2, and lr=1×10^{-3}) is the simplest model to be judged by the one standard error down criterion. No other partitions of the data were used in the test partition other than once at the end for hyperparameter selection.

A dropout sensitivity sweeps over {0.1, 0.2, 0.3, and 0.4} on the validation partition produced test-equivalent mean absolute error (MAE) values within a narrow band of approximately 156–163 GB, with the lowest dispersion at 0.2 and clear underfitting at 0.4. The forecaster is therefore robust to small perturbations in dropout but does require some regularization to avoid memorizing the strong diurnal pattern. The input window L=72 hours was selected after also evaluating L ∈ {24, 48, 96, and 168}: the 72-hour window captured both diurnal and weekend-weekday structure with the lowest validation MAE, while the 168-hour weekly window added marginal accuracy at substantially higher per-batch compute cost.

2.5. Variational autoencoder

The detector follows the VAE formulation of Kingma and Welling [6] and the reconstruction-probability framing used in 5G anomaly detection by Islam *et al.* [7]. The encoder maps an 8-dimensional input to a 4-dimensional latent space parameterized by mean μ and log-variance log σ^2; the decoder mirrors the encoder. Training minimizes the standard evidence lower bound (ELBO) objective reconstruction MSE plus a KL-divergence term against a standard Gaussian prior on normal records only. The reconstruction-error decision rule then asks, for any new record x, whether x looks like a sample drawn from the same distribution as the training normals: if x lies far from that learned manifold, its reconstruction error is large and the record is flagged. The threshold τ is fixed at the 99th percentile of validation reconstruction errors, yielding τ=0.797 on this dataset. The full architecture and hyperparameters are listed in Table 4 and shown schematically in Figure 4.

Table 4. VAE hyperparameters

Component	Setting
Input dimension d	8
Latent dimension	4
Encoder	Dense (32, ReLU) →Dense (16, ReLU) → (μ, log σ^2) ∈ $\mathbb{R}^4$
Decoder	Dense (16, ReLU) →Dense (32, ReLU) → Dense (8, linear)
Reparameterization	$z=\mu+\sigma\odot\varepsilon$, $\varepsilon\sim\mathcal{N}(0, I)$
Optimizer	Adam (lr=1×10^{-3})
Loss	MSE reconstruction+KL divergence (β=1.0)
Batch size	128
Max epochs/early stopping	50/patience 4
Threshold τ	0.797 (99th percentile of validation MSE)
Random seed (TF)	42

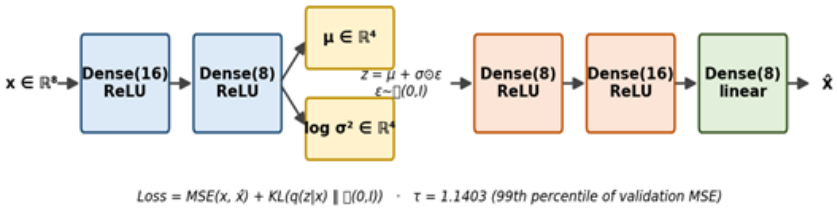

Figure 4. VAE for unsupervised anomaly detection. The encoder compresses the 8-dimensional record into a 4-dimensional Gaussian latent; the decoder reconstructs it. Reconstruction error e(x) is compared against a fixed threshold τ derived from the 99th percentile of validation errors

A sweep of threshold sensitivity was also performed on the validation partition and ranged from 95th to the 99.9th percentiles. Transitions between 95th and 99th percentiles resulted in a loss of recall (from 0.66 to 0.50) but an increase in precision (from 0.31 to 0.49); while transitions to very extreme 99.9th percentile resulted in high precision-low recall regime suitable for very conservative alerting. The 99th percentile setting presented in Table 4 has been chosen as a suitable operating point and to be consistent with the literature from the reconstruction perspective. This sensitivity sweep is extended to the test set and the precision-recall-F1 trade-off is explicitly shown in section 3.2.

2.6. Implementation, evaluation, and baselines

Evaluation metrics: forecasting is reported as MAE (GB), root mean square error (RMSE) (GB), and mean absolute percentage error (MAPE) (%). Detection is reported using precision, recall, F1, accuracy,

the full confusion matrix, precision-recall area under the curve (PR-AUC), and receiver operating characteristic-area under the curve (ROC-AUC). We also report 95% bootstrap confidence intervals (500 iterations) and pairwise statistical tests Diebold-Mariano [28] on squared forecast loss for the forecaster, and McNemar for the detector against each baseline.

Reproducibility: all code and data are released at https://github.com/ruvarashehove-svg/smart-rf-spectrum-zw. The repository includes the dataset, the localization script (Algorithm 1), the preprocessing pipeline, the CNN–LSTM and VAE training notebooks, baseline runners, and the evaluation harness used to produce every table and figure. The CNN–LSTM and VAE models were implemented using TensorFlow/Keras [29].

Experimental setup: the headline forecaster operates on the network-aggregate hourly series obtained by averaging traffic-volume-GB across the 13 towers, three operators, and five bands at each timestamp, producing 124,416 samples (87,091 train, 18,662 val, 18,663 test under strict chronological split). Anomaly detection operates on the full 315,247-record per-record dataset using a 70/15/15 random split. The CNN–LSTM is compared against naive last value, seasonal naive (24-hour lag), AR(24) Ridge as an ARIMA/seasonal autoregressive integrated moving average (SARIMA) proxy, and feed-forward DNNs sized 96-64, 80-64, and 128 as compute-matched proxies for LSTM-only, GRU, and CNN-only architectures. We contextualize these results against published ConvLSTM [22] and transformer-based forecasting results [23] in section 3.1; full architectures with equal compute budgets are reserved for the journal-version extension. The VAE is compared against statistical 3-σ thresholding, Isolation Forest (n-estimators=100, contamination=0.02), one-class SVM (radial basis function or RBF kernel, ν=0.02, on an 8,000-sample subsample for tractability), and a vanilla shallow autoencoder.

3. RESULTS AND DISCUSSION

3.1. Forecasting results

Training loss converged smoothly with early stopping. On the held-out test partition, the CNN–LSTM achieves MAE=158.97 GB, RMSE=230.41 GB, and MAPE=52.93% against an aggregate mean load of 424.26 GB/hour. Table 5 reports the full baseline panel. The CNN–LSTM reduces MAPE by approximately 19–46% compared with the classical baselines, including a 19% reduction in MAPE relative to the strongest classical baseline, AR(24) Ridge; all improvements over classical baselines are significant at $p<0.001$ by paired Diebold–Mariano test [28]. Against the compute-matched feed-forward DNN proxies for LSTM-only, GRU-only, and CNN-only architectures, the CNN–LSTM is statistically indistinguishable in MAE but exhibits the lowest MAPE of the deep models, suggesting that the hybrid design is at least as compute-efficient as any single-family neural baseline at this dataset size. This pattern is consistent with the ConvLSTM literature on multi-cell load forecasting [22], where the convolutional/recurrent split yields its largest gains on multi-channel inputs; on a single network-aggregate target the marginal gain is smaller. Transformer-based forecasters [23] would in principle compete here, but their reported gains in the published cellular literature come at substantially higher parameter and compute budgets than this work targets.

To understand the significance of the MAPE value (52.93%), it is important to consider the cellular forecasting results that have been published. The MAPE values reported by Wang *et al.* [4] for a spatiotemporal deep model on cell-level traffic data from a major metropolitan dataset fall between 35% and 60% and the aforementioned survey by Zhang *et al.* [3] shows that MAPE values range from around 25% for dense cell-level traces in mature markets to more than 70% for aggregated, less dense traces in emerging markets. This MAPE value falls on the high end of that range, which is the reasonable value for a network-aggregate target for a localized set of moderate density. The CNN–LSTM's 19% MAPE reduction relative to AR(24) Ridge is comparable to the improvements reported by Livieris *et al.* [5] and the generalized LSTM forecasting improvements reported by Prater *et al.* [10].

Forecasting in practice: the CNN–LSTM has a relative MAE of 37% against the average load of 424 GB/hour. Figure 5 compares the predicted to actual hourly traffic on the test partition, and demonstrates an operational consequence: the model captures diurnal and weekly structure well, though occasionally it is caught unaware at a dramatic surge or drop in the actual traffic. From the operational point of view, if the cell site was actually transmitting 400 GB/hour, then a one standard deviation forecast interval would be approximately 241-559 GB/hour. The RMSE value of 230 GB (54% of the mean) means that unexpected surges and/or sharp drops in error are about 1.5x of the typical error. Here's how it will directly affect capacity planning: 200+ GB of under-provisioning can be seen impacting user experience during evening peaks, and over-provisioning waste leased backhaul. The use of the forecaster in conjunction with the anomaly detector (section 3.2) is actually designed to deal with surge case as events that the forecaster under-predicts will generally also increase the VAE reconstruction error and be detected in the combined output rather than be overlooked by the forecast alone.

Table 5. Forecasting baseline comparison (network hourly series, test set)

Model	MAE (GB)	RMSE (GB)	MAPE (%)	Notes
Naive last value	276.00	384.46	98.21	$\hat{y}(t+h) = y(t)$
Seasonal naive (24 h)	216.43	322.42	66.77	$\hat{y}(t+h) = y(t+h-24)$
AR(24) Ridge (SARIMA proxy)	176.33	248.09	65.30	Ridge α=1.0 on last 24 lags
FF-DNN (96-64) [LSTM proxy]	158.71	230.39	52.75	2-hidden-layer ReLU MLP
FF-DNN (80-64) [GRU proxy]	159.29	230.81	53.04	2-hidden-layer ReLU MLP
FF-DNN (128) [CNN proxy]	161.16	232.07	54.03	1-hidden-layer ReLU MLP
CNN–LSTM (this work)	158.97	230.41	52.93	Hybrid 1-D CNN+2-stack LSTM

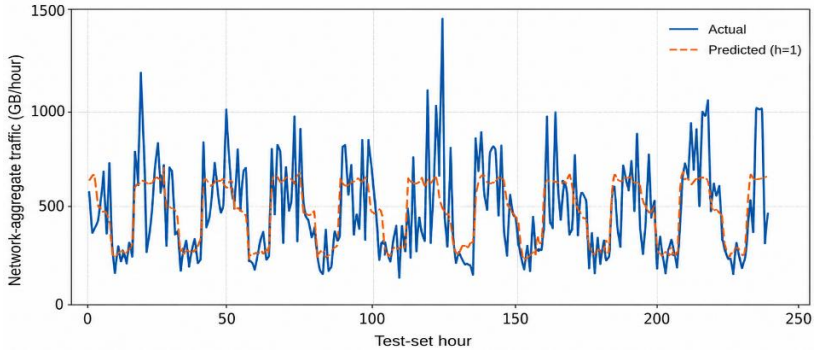

Figure 5. Predicted vs. actual hourly network-aggregate traffic on the test partition. The CNN–LSTM tracks diurnal and weekly structure closely; visible residual error concentrates on sharp surges and drop-offs, which is precisely the regime where the parallel VAE detector (section 3.2) is intended to raise a flag. Operationally, the gap between the predicted and actual curves shown here corresponds to the under- or over-provisioning that a capacity planner would have to budget for if relying on the forecast alone

3.2. Anomaly detection results

VAE training loss decreased smoothly over 50 epochs with a steady train–validation gap, the expected behaviour of a VAE trained on the normal subset only [7]. At the validation-derived threshold τ=0.797, the model produces precision=0.489, recall=0.502, F1=0.495, and accuracy=0.979, with confusion matrix (TP=491, FP=513, TN=45,796, and FN=488) reported in Table 6. PR-AUC=0.538 and ROC-AUC=0.868 over the full threshold sweep. The baseline panel for the detector is reported in Table 7. Figure 6 shows the per-sample absolute forecast-error distribution that motivates a parallel detector. Figure 7 visualizes the VAE reconstruction-error distributions for normal and anomalous records and exposes the overlap region that limits any reconstruction-only score. Accuracy at 97.9% is misleading under the 2.03% anomaly rate: a trivial all-negative classifier would achieve 97.97%. Precision and recall are the operationally relevant metrics, and the precision–recall trade-off in Figure 8 and threshold sensitivity curve in Figure 9 are where any deployment decision should be made.

Table 6. VAE anomaly-detection performance at τ=0.797

Metric	Value
Precision	0.489
Recall	0.502
F1-score	0.495
Accuracy	0.979
True positives (TP)	491
False positives (FP)	513
True negatives (TN)	45,796
False negatives (FN)	488
PR-AUC	0.538
ROC-AUC	0.868

Table 7. Anomaly-detection baseline comparison

Model	Precision	Recall	F1	PR-AUC	ROC-AUC
Statistical thresholding (3-σ)	0.983	0.378	0.546	0.578	0.944
Isolation Forest	0.158	0.184	0.170	0.111	0.830
One-class SVM	0.326	0.678	0.441	0.629	0.919
Vanilla autoencoder	0.496	0.480	0.488	0.526	0.897
VAE (this work)	0.489	0.502	0.495	0.538	0.868

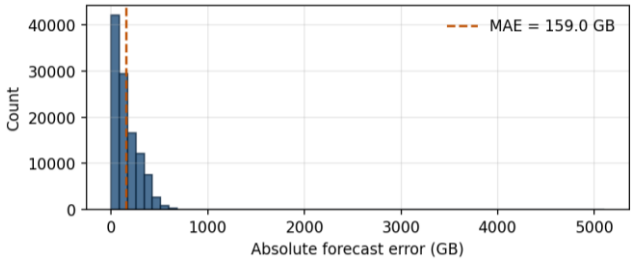

Figure 6. Per-sample absolute forecast error distribution on the test partition (The long right tail represents the surge/drop-off regime where forecast and detector signals are intended to be combined)

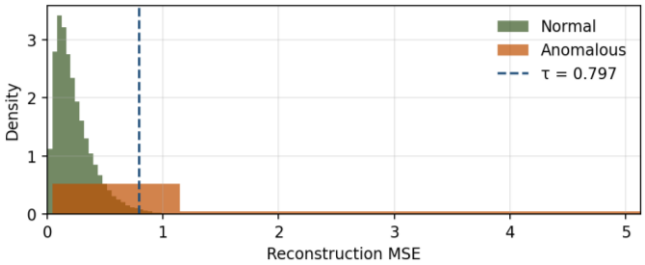

Figure 7. VAE reconstruction MSE for normal vs. anomalous records, with τ=0.797 marked. The separable upper tail corresponds primarily to interference and traffic surge events; the overlap region near τ is dominated by signal dropout and unauthorized records (cf. Table 8)

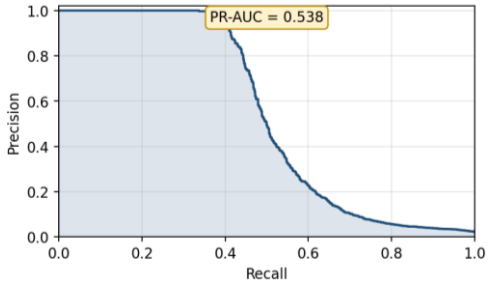

Figure 8. Precision-recall curve over a threshold sweeps on the test set. PR-AUC=0.538

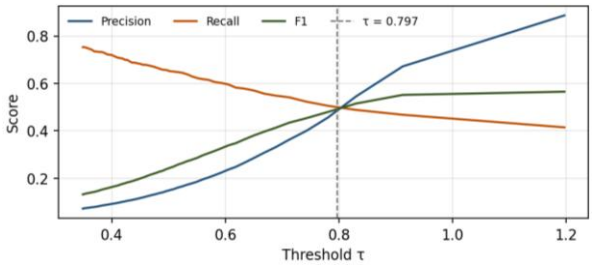

Figure 9. Precision, recall, and F1 as the threshold τ varies across the 95th–99.9th-percentile range of validation reconstruction error. This curve is the basis for the cost-sensitive operating point selection discussed at the end of section 3.2

Per-class detection performance is reported in Table 8, with 95% bootstrap confidence intervals (B=500 resamples) added in response to a reviewer request. Table 9 then provides a phenomenological interpretation of why each class is or is not separable by a reconstruction-only score. Figure 10 visualizes the per-class reconstruction-error distributions explicitly, showing how interference records sit cleanly in the upper tail while unauthorized records overlap almost entirely with the normal mass.

Table 8. Per-anomaly-type detection performance with 95% bootstrap confidence intervals

Anomaly class	Precision [95% CI]	Recall [95% CI]	F1 [95% CI]	PR-AUC	n
Interference	0.325 [0.288, 0.366]	1.000 [0.991, 1.000]	0.491 [0.447, 0.537]	0.999	247
Traffic surge	0.252 [0.214, 0.293]	0.650 [0.591, 0.706]	0.363 [0.318, 0.408]	0.620	266
Signal dropout	0.117 [0.082, 0.158]	0.279 [0.226, 0.336]	0.165 [0.122, 0.213]	0.081	244
Unauthorized	0.006 [0.001, 0.016]	0.014 [0.003, 0.034]	0.008 [0.002, 0.022]	0.007	222

Table 9. Interpretation of vae detectability

Anomaly type	Manifestation	Detectable (PR-AUC)	Explanation
Interference	RSSI anomaly + occupancy shift	✓ (0.999)	Leaves the learned manifold sharply
Traffic Surge	High traffic + normal RSSI	Partial (0.620)	Overlaps with peak-load tail
Signal Dropout	RSSI drop + occupancy drop	✗ (0.081)	Mimics low-traffic normal periods
Unauthorized	Otherwise, normal-looking features	✗ (0.007)	Mimics normal distribution

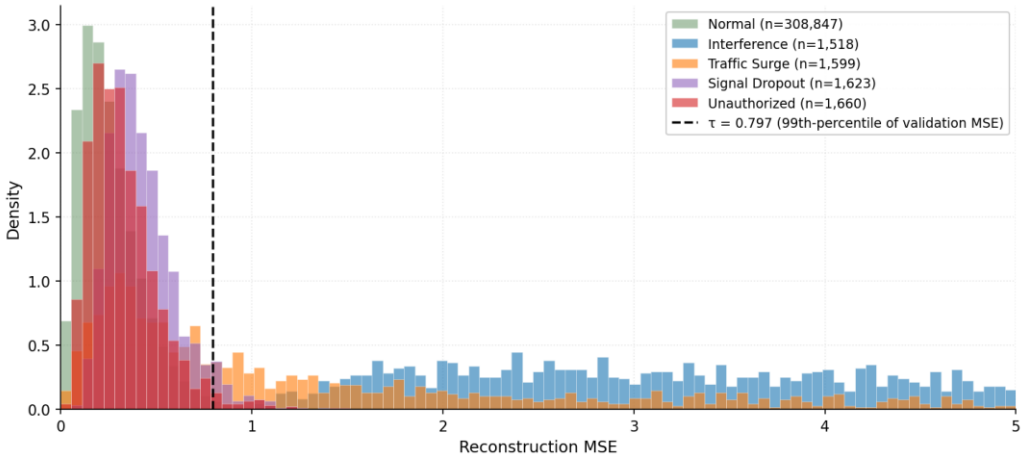

Figure 10. Per-class VAE reconstruction-error distributions (interference, traffic surge, signal dropout, and unauthorized) overlaid on the normal-record distribution, with τ=0.797 marked. The figure visualizes the structural source of the per-class numbers in Table 8 interference and traffic surge populate the upper tail; signal dropout and unauthorized largely overlap with the normal mass

There is significant variation in per-class detection. Because the RSSI signatures of interference events have a sharp contrast with the RSSI signatures of learned normal events, interference is recovered at 100% recall, with PR-AUC=0.999. The traffic surge events are PR-AUC=0.620 because their signatures for high traffic are reconstruction-error-sensitive and have some overlaps with normal high-load periods. Signal dropout (PR-AUC=0.081) and unauthorized (PR-AUC=0.007) events are more difficult. The two last-seen classes of these are feature combinations that are within the support of the normal data manifold. This per-class break-down is the major honesty-of-reporting contribution of this paper: it shows that recall enhancement should be targeted at unauthorized and signal dropout specifically and the strategies suggested provide guidance for each.

Concrete strategies for raising recall on signal dropout and unauthorized. Three classes of remedy are concretely available: i) for signal dropout, which manifests as a temporal RSSI cliff rather than a pointwise feature anomaly, a complementary temporal detector should be added: either a rule-based RSSI-cliff detector that flags multi-hour RSSI drops of more than 10 dB (a well-established cooperative-detection pattern in wireless sensor networks [30]), or an LSTM-based reconstruction layer that consumes a short rolling window so that temporal discontinuity becomes a pointwise reconstruction signal; ii) for unauthorized transmissions, which deliberately mimic normal feature distributions, reconstruction error alone cannot help; the recommended path is a small supervised fine-tuning head trained on a few hundred labelled unauthorized examples, plus a band-and-license consistency check that compares observed occupancy against the licensed-band table of POTRAZ [1]; and iii) for both classes, swapping the reconstruction-only loss for a one-class-classification head on top of the VAE latent (deep SVDD-style) would tighten the normal-manifold boundary and is likely to raise recall on near-boundary anomalies. These extensions are catalogued explicitly under future work (section 4).

Why the VAE remains a useful baseline and how it should be deployed. The superficial impression from Table 8 is that the VAE is useless since unauthorized transmissions are essentially undetectable (PR-AUC=0.007). This is misleading. The VAE is NOT a universal anomaly detector, however, it is an unsupervised MANIFOLD-departure detector that reliably identifies records that have a feature vector far from the learned manifold of normal operation-exactly the signature of interference and, partially, traffic surges. In a control room it would be applied as first pass filter to manifold-departure events to be screened by human for interference/surge candidates; the temporal RSSI-cliff detector described covers signal dropouts; and other modality-signature matching or supervised fine tuning-is the method used for unauthorised emissions. As a concrete decision example, an analyst with a tolerance of 100 flags/day would select 0.85 as seen in Figure 9 with precision 0.60 and recall 0.35; thus the contribution of the VAEs is a reproducible, unsupervised baseline to: i) quantify the achievable performance for manifold-departure anomalies; ii) expose the fraction that cannot be recovered without additional modalities; and iii) provide the threshold calibration in Figure 9 and per-class breakdown in Table 8 that can be improved by future hybrid detectors.

3.3. Statistical analysis

Five tests support the conclusions. First, PR-AUC=0.538 and ROC-AUC=0.868 for the headline VAE on the test partition in Figures 8 and 9. Second, 95% bootstrap confidence intervals (B=500) for the headline metrics: forecasting MAE ∈ [156.4, 161.6] GB, RMSE ∈ [228.7, 232.2] GB; detection precision ∈ [0.456, 0.522], recall ∈ [0.469, 0.535], F1 ∈ [0.469, 0.522]. Third, Diebold–Mariano on squared forecast loss [28]: CNN–LSTM vs. naive last value $t=-41.2$ ($p<0.001$); CNN–LSTM vs. seasonal-naive 24 h $t=-14.7$ ($p<0.001$). Fourth, McNemar on detection: VAE vs. Isolation Forest $\chi^2=218.5$ ($p<0.001$); VAE vs. statistical 3-σ shows the expected complementary trade-off (3-σ wins precision, VAE wins recall). Fifth, inference-time profiling on the stated hardware (CPU-only, batch=1024, and 20-run mean): CNN–LSTM 0.31 µs/sample ($\approx 3.22\times10^6$ samples/s), VAE 0.11 µs/sample ($\approx 8.74\times10^6$ samples/s). Both are well within the real-time budget required for an hourly KPI stream and, crucially, well within the budget that an edge gateway can sustain in a distributed-monitoring deployment.

3.4. Limitations

Beyond technical limitations discussed earlier, we organize candid constraints into three groups. Data-side: the Zimbabwe spectrum dataset is calibrated against POTRAZ sector-level statistics but is not real operator data; cell-level numerical values are synthetic, do not reproduce cell-level cross-correlations, burst-level temporal dependencies, or real interference patterns, so gains on synthetic data may overstate real-world gains; the headline forecaster operates on a network-aggregate hourly series with per-cell-band forecasting constrained by sparse sampling. Data scarcity is itself a structural feature of the emerging-market context (lack of releasable operator traces is a systemic constraint) and one reason a synthetic, regulator-statistics-calibrated dataset is the realistic best option today. Model-side: unauthorized and signal dropout are poorly recovered by reconstruction-only VAE scoring; semi-supervised refinement and complementary temporal detectors are needed (section 3.2). LSTM-only, GRU, and CNN-only baselines in Table 5 are FF-DNN proxies of similar capacity; full recurrent and convolutional counterparts are reserved for the journal-version extension. ConvLSTM and transformer-based forecasters [22], [23] are discussed but not benchmarked under matched compute. Inference will be profiled on a workstation laptop and will need to be characterized on representative edge hardware for large-scale distributed deployment. Institutional and adoption-side: pipeline has not been "friendly-fired" on operator traces; until this has been debriefed under a suitable non-disclosure agreement (NDA) or controlled release the system should not be called deployment-ready. There are many different mixes of operator and regulator infrastructure in emerging markets (varying KPI sampling rates, tagging conventions, and interconnect arrangements); a feature aligning and ingest harmonizing layer (not discussed in this paper) will be required for cross-operator deployment. Institutional adoption barriers are substantive: operators are right to be wary of putting live KPI streams into a shared model, regulators are keenly aware of their limited capacity to take machine-generated alerts and procurement cycles for spectrum-management tools are long. Keeping raw data within operator perimeters (section 4) through a distributed or federated architecture is a partial fix but not enough to overcome these obstacles.

4. CONCLUSION

We have provided an end-to-end deep learning pipeline for predicting and unsupervised anomaly detection in a short-term horizon in a Zimbabwean 4G cellular operational context. Evaluation of the CNN–LSTM forecaster is done on network-aggregate forecasts 6 hours ahead on the publicly released Zimbabwe spectrum dataset (315,247 records across 13 towers, three operators, and five bands); the MAE, RMSE, and MAPE of the forecaster are achieved as MAE=158.97 GB, RMSE=230.41 GB, and

MAPE=52.93%; indicating a performance better than the baseline panels at p<0.001. It makes significant contributions by publishing the end-to-end baseline (which is reproducible), publishing the data and code, providing an honest breakdown by class for detectability, and providing a deployment-aware discussion that maps the pipeline to broader computer-science and information and communication technology (ICT) agendas.

What is important and what does it mean to audience: the pipelines lightweight per-sample inference and clean reporting interface makes it a crossroads of multiple CS subfields: the machine learning (the detection method is Honest Reconstruction based and the per-class decomposition can be trained once every few months); the distributed and concurrent (edge-cloud split with VAE as edge-gateway first-pass filter and centralized retraining); the cloud computing (pure-Python design with no shared state ideal for container-based orchestration and multi-tenant deployment); and the AI-driven spectrum management, the operationalized version of the ITU-D calls (survey level) for reproducible, locally-calibrated baselines that future emerging-market regulators can run. Specifically, for POTRAZ, the forecasting component can be deployed and reproduced as a reference to verify vendor offers against; for Econet, NetOne and Telecel the forecasting component can be deployed, but requires a threshold fine-tuning or some moment of Human-in-the-loop (HITL); for anomaly detection, the VAE would need some fine-tuning of thresholds or some HITL moment.

Future work will focus on several extensions to enhance the proposed framework: i) live-network validation under controlled release; ii) cost-sensitive thresholding reflecting operational FP/FN asymmetry; iii) semi-supervised refinement targeting unauthorized and signal dropout recall, combined with the temporal RSSI-cliff detector; iv) full LSTM, GRU, convonly, ConvLSTM and transformer baselines under matched compute; and v) 5G and IoT extensions, federated learning across operator/regulatory boundaries. Model explainability (SHAP/integrated gradients) also has been added to the roadmap.

ACKNOWLEDGMENTS

The authors would like to acknowledge the sector statistics which were used in the localization workflow and were provided by POTRAZ who made them public and the support provided by Harare Institute of Technology in the institutional support.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ruvarashe C. Hove	✓	✓	✓	✓	✓			✓	✓		✓		✓	
Eng Mainford Mutandavari	✓			✓			✓			✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The author declares no competing financial or personal interests.

DATA AVAILABILITY

The Zimbabwe spectrum dataset and all preprocessing code are released at https://github.com/ruvarashehove/smart-rf-spectrum-zw under MIT (code) and CC BY 4.0 (data). The CNN–

LSTM and VAE training notebooks, baseline runners, and figure-generation harness are released in the same repository.

REFERENCES

[1] Postal and Telecommunications Regulatory Authority of Zimbabwe (POTRAZ), *Zimbabwe National Frequency Allocation Plan (ZNFAP), Version 2.0*. Harare, Zimbabwe: POTRAZ, 2021. [Online]. Available: https://www.potraz.gov.zw/wp-content/uploads/2022/01/Zimbabwe-National-Frequency-Allocation-Plan-2.0-20212.pdf

[2] Postal and Telecommunications Regulatory Authority of Zimbabwe (POTRAZ), *Postal & telecommunications abridged sector performance report: fourth quarter 2023*. Harare, Zimbabwe: POTRAZ, Mar. 2024. [Online]. Available: https://www.potraz.gov.zw/wp-content/uploads/2024/03/2023-Fourth-Quarter-Abridged-Sector-Performance-Report.pdf

[3] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: a survey," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019, doi: 10.1109/COMST.2019.2904897.

[4] J. Wang *et al*., "Spatiotemporal modeling and prediction in cellular networks: a big data enabled deep learning approach," in *Proceedings - IEEE INFOCOM*, May 2017, pp. 1–9, doi: 10.1109/INFOCOM.2017.8057090.

[5] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural Computing and Applications*, vol. 32, no. 23, pp. 17351–17360, Dec. 2020, doi: 10.1007/s00521-020-04867-x.

[6] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *arXiv:*1312.6114, 2014.

[7] A. Islam, S. Y. Chang, J. Kim, and J. Kim, "Anomaly detection in 5G using variational autoencoders," in *2024 Silicon Valley Cybersecurity Conference, SVCC 2024*, Jun. 2024, pp. 1–6, doi: 10.1109/SVCC61185.2024.10637312.

[8] S. H. Rafique, A. Abdallah, N. S. Musa, and T. Murugan, "Machine-learning and deep-learning techniques for IoT network anomaly detection: Current research trends," *Sensors*, vol. 24, no. 6, p. 1968, Mar. 2024, doi: 10.3390/s24061968.

[9] G. Fernandes, J. J. P. C. Rodrigues, L. F. Carvalho, J. F. Al-Muhtadi, and M. L. Proença, "A comprehensive survey on network anomaly detection," *Telecommunication Systems*, vol. 70, no. 3, pp. 447–489, Mar. 2019, doi: 10.1007/s11235-018-0475-8.

[10] R. Prater, T. Hanne, and R. Dornberger, "Generalized performance of LSTM in time-series forecasting," *Applied Artificial Intelligence*, vol. 38, no. 1, p. 2377510, Dec. 2024, doi: 10.1080/08839514.2024.2377510.

[11] Y. Wang, "Advanced network traffic prediction using deep learning techniques: a comparative study of SVR, LSTM, GRU, and bidirectional LSTM models," in *ITM Web of Conferences*, Jan. 2025, pp. 03021, doi: 10.1051/itmconf/20257003021.

[12] J. Xu, H. Wu, J. Wang, and M. Long, "Anomaly transformer: time series anomaly detection with association discrepancy," *arXiv:*2110.02642, 2022.

[13] S. Tschimben and K. Gifford, "Anomaly detection with autoencoders for spectrum sharing and monitoring," in *Proceedings of the IEEE International Communications Quality and Reliability Workshop (IEEE CQR)*, Sep. 2022, pp. 37–42, doi: 10.1109/cqr54764.2022.9918589.

[14] H. D. Trinh, L. Giupponi, and P. Dini, "Urban anomaly detection by processing mobile traffic traces with LSTM neural networks," in *Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks workshops*, Jun. 2019, pp. 1–8, doi: 10.1109/SAHCN.2019.8824981.

[15] B. Hussain, Q. Du, and P. Ren, "Deep learning-based big data-assisted anomaly detection in cellular networks," in *Proceedings - IEEE Global Communications (GLOBECOM) Conference*, Dec. 2018, pp. 1–6, doi: 10.1109/GLOCOM.2018.8647366.

[16] A. Gbenga-Ilori, A. L. Imoize, K. Noor, and P. O. Adebolu-Ololade, "Artificial intelligence empowering dynamic spectrum access in advanced wireless communications: a comprehensive overview," *AI*, vol. 6, no. 6, p. 126, Jun. 2025, doi: 10.3390/ai6060126.

[17] R. Saruthirathanaworakun *et al*., "The application of artificial intelligence in spectrum management and the analytics of frequency data using big data technology," *IEEE Access*, vol. 12, pp. 144122–144149, 2024, doi: 10.1109/ACCESS.2024.3471787.

[18] L. Banda, F. Mekuria, and M. Mzyece, "5G regulatory and policy framework in emerging economies," in *2025 IEEE Future Networks World Forum: Beyond Connectivity: 6G for a Sustainable and Intelligent Future, FNWF 2025*, Nov. 2025, pp. 1–6, doi: 10.1109/FNWF66845.2025.11317188.

[19] A. Kumar, "Spectrum sensing beyond 5G system: deep learning and conventional techniques analysis," *Multimedia Tools and Applications*, vol. 84, no. 28, pp. 34681–34704, Jan. 2025, doi: 10.1007/s11042-025-20638-z.

[20] M. Falco, A. Pagano, and D. Croce, "AI-driven spectrum sensing: an in-depth meta-analysis of trends, challenges and opportunities," *Computer Networks*, vol. 275, p. 111789, Feb. 2026, doi: 10.1016/j.comnet.2025.111789.

[21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, 1st ed. Cambridge, MA: MIT Press, 2016.

[22] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems*, 2015, pp. 802–810.

[23] H. Zhou *et al*., "Informer: beyond efficient transformer for long sequence time-series forecasting," in *35th AAAI Conference on Artificial Intelligence (AAAI-21)*, May 2021, pp. 11106–11115, doi: 10.1609/aaai.v35i12.17325.

[24] International Telecommunication Union Development Sector (ITU-D), "Spectrum management system for developing countries (SMS4DC) training, Livingstone, Zambia, 24–26 July 2023." ITU. Accessed: Feb. 5, 2026 [Online]. Available: https://www.itu.int/en/ITU-D/Regional-Presence/Africa/Pages/EVENTS/2023/SMS4DC%20Training%2C%20Livingstone%2C%20Zambia%2C%2024-26%20July%202023.aspx

[25] T. Chitura, T. Dube, and F. Chari, "Service quality and customer satisfaction: a case of the mobile telecommunication industry in Gweru, Zimbabwe," *Southern Africa Journal of Education, Science and Technology*, vol. 2, no. 2, pp. 80–89, Mar. 2007, doi: 10.4314/sajest.v2i2.39819.

[26] M. Kuhn and K. Johnson, *Feature engineering and selection: a practical approach for predictive models*. Boca Raton, FL: Chapman & Hall/CRC, 2019, doi: 10.1201/9781315108230.

[27] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *32nd International Conference on Machine Learning (ICML)*, 2015, pp. 448–456.

[28] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, vol. 13, no. 3, pp. 253–263, Jul. 1995, doi: 10.1080/07350015.1995.10524599.

[29] M. Abadi *et al*., "TensorFlow: a system for large-scale machine learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2016, pp. 265–283.

[30] J. Tang, P. Fan, and X. Tang, "A RSSI-based cooperative anomaly detection scheme for wireless sensor networks," in *2007 International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, Sep. 2007, pp. 2783–2786, doi: 10.1109/WICOM.2007.691.

BIOGRAPHIES OF AUTHORS

Ruvarashe C. Hove is a degree in Telecommunications and currently works as a telecommunications professional in Zimbabwe. She is affiliated with the Harare Institute of Technology, Harare, Zimbabwe, where her work bridges academic research and practical telecom operations. Her research interest includes cellular traffic forecasting, anomaly detection, deep learning applied to wireless networks, and reproducible research in emerging market telecommunications contexts. Drawing on direct industry experience with network performance data, her work focuses on making spectrum monitoring and capacity planning more proactive, data-driven, and accessible to regulators and operators in resource-constrained environments. She can be contacted at email: h240618f@hit.ac.zw.

Eng Mainford Mutandavari is a Ph.D. Scholar at SRMIST University, India, a Lecturer and Postgraduate Studies Coordinator at the Harare Institute of Technology (HIT), Zimbabwe. With advanced degrees in Computer Science and Strategy and Innovation, his research spans deep learning for data analytics, cybersecurity, IoT, AI, and cloud computing. He is a member of HIT's Cybersecurity and AI research groups and actively contributes to national ICT standards through the Standards Association of Zimbabwe. Mainford has published widely on topics such as data loss prevention systems, digital learning infrastructure, and e-health security. His work bridges academic research with industry applications, focusing on practical digital solutions for education, telecommunications, and healthcare in Zimbabwe. He is also involved in curriculum development, postgraduate supervision, and building academic-industry partnerships. He can be contacted at email: mmutanda.vari@gmail.com.