An EWMA-Based Adaptive Thresholding Concept for Autoencoder-Based Concept Drift Detection in Data Streams

EWMA-Based Adaptive Threshold Concept

Siti Nurulain Mohd Rum, Qiao Song
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia, Serdang, Malaysia

***Abstract*—The static thresholds derived from the primary validation timeframe (window) are a common method for the detection of reconstruction-based concept drift. In prolonged periods of data streams, progressive changes in reconstruction accuracy frequently lead to misalignment, giving rise to repeated false alarms and inconsistency in detection behavior. This study introduces a modular lightweight adaptive thresholding strategy for the Autoencoder-Based Drift Detection Method (AEDDM) by integrating an Exponentially Weighted Moving Average (EWMA) mechanism into the batch-level decision process, without modifying the original model architecture. Rather than constituting a standalone framework, the proposed method functions as a modular enhancement to the decision layer of the AEDDM pipeline. The proposed solution is validated using a synthetic Gaussian stream together with the ELEC2 and NSL-KDD datasets. The finding demonstrates that the EWMA-based approach effectively eliminates false alarms without compromising responsiveness under abrupt changes, achieving zero-latency on NSL-KDD compared to static thresholds that produced 22 warnings and 25 false alarms in a stationary stream. Findings from this study suggest that adaptive thresholding alone significantly leads to the enhancement of detection performance in reconstruction-driven drift on a real-time stream.**

Keywords*—Concept drift detection; streaming data; autoencoder; adaptive threshold; EWMA

I. Introduction

The growth of data-intensive applications has prompted a shift from conventional batch analysis towards real-time data processing. Many modern applications generate data continuously, which must be handled immediately and require online processing within the limits of computational and memory capacities. Typical applications are often found in network traffic analysis, monitoring financial transactions, sensor-driven surveillance, and the operation of the smart grid. In such settings, predictive models are expected to provide accurate outputs and consistent performance to evolving data patterns. A key challenge in stream learning is concept drift [1-3], referring to the changes in underlying data distribution [1]. These changes can impact the feature distribution, the mapping between target labels (input and output), or both aspects. In contrast to prior work that presents drift detection as a monolithic framework, this study repositions its contribution as a modular adaptive thresholding component that can be integrated into existing reconstruction-based pipelines. Rather than altering the underlying learning model, the proposed approach focuses on enhancing the decision layer by introducing a statistically grounded mechanism for interpreting reconstruction errors. This modular perspective improves flexibility, allowing the method to be incorporated into a wide range of autoencoder-based or hybrid drift detection systems.

As concept drift takes place, models depending on the historical dataset may underperform, failing to capture the current condition, resulting in the ineffectiveness of the adaptation strategy. As a result, detection of concept drift in a timely and accurate manner emerged as a major concern in the data stream processing field. Concept drift can manifest in different ways, such as abrupt, progressive, stepwise, and recurring patterns [2]. Abrupt drift refers to sudden shifts in data distribution, while gradual and incremental drift manifest over longer periods. These disparities require different detection mechanisms. Instant refinement is needed by abrupt drift, on the other hand, gradual drifts are strengthened by stable and noise-tolerant methods. In addition, in real streaming environments, labels are frequently missing or prolonged, complicating detection and restricting the effectiveness of fully supervised strategies [4].

Modern data-intensive applications generate high-velocity streams that must be processed online. Traditional supervised drift detection methods, which rely on immediate label availability, are often impractical in these environments due to the high cost or latency of labeling. This limitation has inspired the development of unsupervised methods that monitor data distributions directly rather than relying on predictive accuracy [5-7]. Among these, reconstruction-based techniques using autoencoders have emerged as a promising direction [8],[9],[10],[11]. Autoencoders work as neural networks to identify the underlying data structure by reducing reconstruction errors. When incoming data matches the captured pattern, errors remaining in reconstruction stay negligible, while deviations from the training distribution tend to trigger an increase in reconstruction errors. This characteristic makes the approach suitable for drift detection in situations where labels are scarce or absent at deployment time. A widely used method in this area is the Autoencoder-Based Drift Detection Method (AEDDM) that relies on class-conditional autoencoders trained offline and evaluates reconstruction error patterns during online operation [9]. AEDDM integrates instance-specific thresholds, batch-level metrics, and persistence rules to distinguish actual concept drift from temporary noise. By leveraging construction patterns of a specific class, the approach offers a systematic and interpretable drift detection framework and has shown robust performance on synthetic and real data streams. Despite its strong performance, AEDDM relies on static thresholds established from a fixed validation period. However, in prolonged evolving data streams, gradual shifts in baseline reconstruction errors can cause threshold misalignment[10]. Consequently, the detector behaves unpredictably, such as repeated warning signals or higher false alarm rates, even when no actual concept drift has occurred [11]. The proposed method employs an Exponentially Weighted Moving Average (EWMA) strategy to dynamically model the evolution of reconstruction errors in streaming data. By continuously updating both the central tendency and dispersion of the error distribution, the approach enables adaptive thresholding that responds to non-stationary data characteristics. This is particularly important in real-world streaming environments, where fixed or heuristic thresholds often fail to capture gradual, recurring, or noisy drift patterns.

The contributions of this research are threefold. The study begins by replicating the AEDDM[12] pipeline as a baseline system to provide a controlled environment for evaluating the proposed modular enhancement. Next, an EWMA-based adaptive batch thresholding module is developed and integrated into the decision layer of the pipeline and incorporated into the online detection phase, enhancing robustness under non-stationary conditions without increasing model complexity. Third, extensive experiments are conducted on both synthetic and real-world data streams to evaluate the impact of adaptive thresholding in terms of false alarm rate, detection stability, and responsiveness to drift. In this study, we address this limitation by introducing an adaptive thresholding strategy based on the Exponentially Weighted Moving Average (EWMA)[13] into the AEDDM framework. The proposed approach maintains the original model architecture and training procedure, modifying only the thresholding component to enable continuous adaptation to evolving reconstruction error baselines. The proposed approach allows batch-level detection thresholds to evolve according to recent reconstruction behavior while preserving the original model architecture and decision logic. The method is evaluated using a synthetic Gaussian stream with controlled drift injection, the NSL-KDD network intrusion dataset, and the real-world ELEC2 electricity market dataset. To ensure robustness and generalizability, the proposed method is systematically evaluated across multiple concept drift scenarios, including sudden, incremental, recurring, and noisy drift conditions, allowing assessment under both stable and highly dynamic streaming environments. Each experiment is repeated over multiple randomized runs to reduce sampling bias and improve the statistical reliability of the reported results. The experimental analysis is conducted over repeated runs with statistical validation to ensure consistency of results. In addition, a comprehensive ablation study is performed to systematically quantify the contribution of key design components, including batching, EWMA smoothing, and adaptive threshold scaling. This allows for a deeper understanding of how each component influences detection performance.

II. Related Study

The existing techniques of drift detection can be broadly categorized into three: model-based, distribution-based, and performance-based. Performance-based methods monitor classification error rates to infer distributional changes. Representative examples include the Drift Detection Method (DDM) [14] and its variants, which model the error rate using statistical confidence intervals. While computationally efficient, such approaches require immediate access to ground-truth labels, which is often impractical in real-time streaming scenarios. Distribution-based methods monitor changes in statistical properties of incoming data without relying on predictive performance. A prominent example is ADWIN (Adaptive Windowing) [10], which dynamically adjusts window sizes based on statistically significant differences between sub windows. Other classical approaches employ sequential change detection tests such as the Page–Hinkley method [15]. Although these techniques operate in an unsupervised manner, they may exhibit sensitivity to noise and require careful parameter tuning. Model-based approaches analyze internal model signals to detect deviations from learned patterns. Among them, reconstruction-based methods have gained attention due to their suitability for unlabeled environments. These approaches train a representation model on historical data and monitor reconstruction errors during online deployment. Significant increases in reconstruction error are interpreted as indicators of distributional change.

Autoencoders have appeared as an effective reconstruction model for drift detection due to their ability to capture nonlinear pattern representations. Early work by Sakurada and Yairi demonstrated that denoising autoencoders could detect anomalies through reconstruction error analysis [16]. Subsequent studies extended this concept to streaming environments. Jaworski *et al.* [8] formalized reconstruction-error-based drift detection using feed-forward autoencoders in an unsupervised setting [8]. More advanced variants, such as contrastive autoencoder frameworks, employ latent-space distance measures combined with statistical testing mechanisms [9].

A recent contribution in this direction is the Autoencoder-Based Drift Detection Method (AEDDM), which utilizes class-conditional autoencoders and batch-level reconstruction error monitoring [10]. AEDDM integrates instance-level and batch-level thresholds to distinguish persistent drift from transient fluctuations. While demonstrating competitive performance, AEDDM and most prior autoencoder-based frameworks rely on static thresholds estimated from a fixed validation window. Such thresholds implicitly assume that the reconstruction-error baseline remains stable over time. Concept drift arises when the underlying data distribution varies over time, causing models trained on historical data to degrade in predictive performance [1], [2]. A large body of work has explored learning under non-stationary environments, including statistical, ensemble, and reconstruction-based detection methods [17-19].

However, in prolonged data streams with gradual or incremental distributional shifts, the baseline reconstruction error may drift slowly even in the absence of abrupt concept changes. Under these circumstances, static thresholds can become misaligned with the evolving data distribution, leading to unstable detection behaviour or increased false alarm rates. To address this limitation, adaptive statistical process control techniques such as the Exponentially Weighted Moving Average (EWMA) have been proposed for tracking evolving mean processes [11]. EWMA assigns higher weights to recent observations while retaining historical context, making it ideal for gradual change detection. Despite its theoretical suitability, the integration of EWMA-based adaptive thresholding into autoencoder-driven drift detection frameworks remains relatively unexplored. Therefore, there exists a research gap in developing lightweight adaptive thresholding mechanisms that can dynamically adjust reconstruction-error decision boundaries without altering the underlying neural architecture. This study addresses this gap by incorporating an EWMA-based adaptive batch threshold into a class-conditional autoencoder framework for streaming concept drift detection. Recent advances in concept drift detection have increasingly leveraged deep learning models, particularly reconstruction-based approaches, to capture complex, non-linear data distributions in streaming environments. These methods typically employ autoencoders or their variants to model the underlying data distribution and monitor reconstruction error as an indicator of distributional change [8, 20]. Early reconstruction-based approaches utilize standard autoencoders to detect drift by observing significant deviations in reconstruction error over time. While effective in capturing non-linear patterns, these methods often rely on static or heuristically defined thresholds, limiting their robustness in non-stationary environments [5].To address this limitation, several studies have proposed enhanced architectures such as variational autoencoders (VAEs) and denoising autoencoders, which improve generalization and robustness to noise [21]. More recent work explores deep streaming models, including recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, to capture temporal dependencies in evolving data streams[21]. Hybrid approaches have also emerged, combining reconstruction models with statistical process control techniques. For instance, reconstruction error distributions are monitored using change detection tests or adaptive windows [17]. However, many of these methods still depend on fixed variance assumptions or manually tuned thresholds, which may not adapt effectively under dynamic drift conditions such as incremental or recurring drift [22].

Despite these advancements, a key limitation across existing deep learning-based drift detectors is the lack of statistically grounded and dynamically adaptive thresholding mechanisms. This gap motivates the present work, which introduces an EWMA-based adaptive thresholding module that continuously updates both the mean and variance of reconstruction errors, enabling more robust and responsive drift detection in evolving data streams. Table I summarizes representative autoencoder-based concept drift detection methods. While prior approaches demonstrate the effectiveness of reconstruction error as a drift indicator, most rely on static thresholds derived from fixed validation windows. Such designs assume a stable reconstruction error baseline and may become suboptimal under gradual or long-term distributional shifts. This limitation motivates the development of adaptive thresholding strategies that preserve computational efficiency while improving robustness in non-stationary streaming environments.

TABLE I. SUMMARY OF REPRESENTATIVE DEEP LEARNING-BASED CONCEPT DRIFT DETECTION METHODS

Study / Method	Architecture	Learning Setting	Detection Signal	Threshold Strategy	Adaptability	Computational Complexity	Key Limitation
Masanori Sakurada & Takehisa Yairi [16]	Denoising Autoencoder	Unsupervised	Reconstruction Error	Static Quantile	Low	Moderate	No adaptation to gradual drift
Marek Jaworski et al. [21]	Feed-forward Autoencoder	Unsupervised	Reconstruction Error	Static Validation Threshold	Low	Moderate	Sensitive to noise
Marek Jaworski et al. [8]	Contrastive Autoencoder	Semi-supervised	Latent Distance	KS-Test	Moderate	High	High computational cost
Ubaid Ali & Tariq Mahmood (AEDDM) [9]	Class-Conditional Autoencoder	Semi-supervised	Batch Reconstruction Error	Static Validation Threshold	Moderate	Moderate	Threshold misalignment over time
L. Hu et al. (DNN+AE-DD) [17]	DNN + Autoencoder	Semi-supervised	Reconstruction Error	3σ Static Threshold	Moderate	High	Fixed threshold assumption
N. Harshit & K. Mounvik (Transformer-AE) [18]	Transformer + Autoencoder	Unsupervised	Hybrid Multi-signal	Static Multi-metric	High	Very High	Complex architecture
T. Mehmood et al. (LSTMDD) [19]	LSTM-based Detector	Unsupervised	Temporal Error Patterns	Sliding Window	Moderate	High	Expensive sequential training

III. METHODOLOGY

This study adopts the Autoencoder-Based Drift Detection Method (AEDDM) [9] as the base pipeline and introduces a modular enhancement to its thresholding mechanism. The objective is to monitor variation in distribution in streaming data under unlabeled online conditions by analyzing reconstruction error patterns. The suggested improvement introduces a dynamic thresholding mechanism based on the Exponentially Weighted Moving Average (EWMA) to handle the drawbacks of static threshold strategies in prolonged streams. Fig. 1 is the overall framework of the AEDDM-based drift detection system that consists of two main phases, an offline model preparation phase and an online drift monitoring phase. The EWMA-based adaptive threshold operates as an independent module within the detection pipeline, highlighting its plug-in nature.

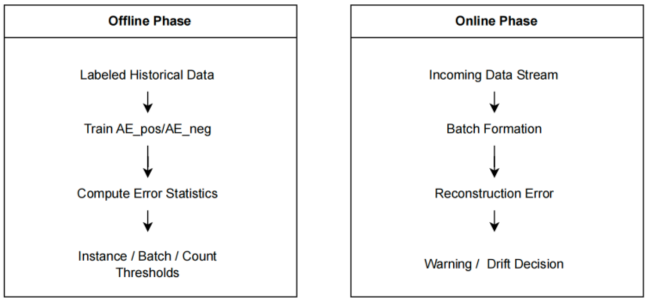

Fig. 1. Overall framework of the AEDDM-based drift detection system [9].

A. Model Preparation and Validation Control

Two class-dependent autoencoders, AE_{pos} and AE_{neg}, were trained using labeled data during the offline phase. The evaluation subset was used to estimate baseline reconstruction error statistics for threshold comparison. Throughout the online phase, data were executed in order without using labels, and batch reconstruction errors were evaluated to simulate real-time drift detection.

B. Autoencoder Reconstruction Error

The reconstruction quality of the autoencoder is quantified using the Mean Squared Error (MSE) loss function:

$$MSE(x, \hat{x}) = \frac{1}{d} \sum_{i=1}^{d} (x_i - \hat{x}_i)^2 \quad (1)$$

where,

- $x_i$denotes the original input feature,
- $\hat{x}_i$represents the reconstructed feature produced by the autoencoder,
- and $d$is the total number of input features.

The objective of the autoencoder is to minimize the reconstruction error between the original and reconstructed samples. Model parameters are optimized using the Adam optimizer.

C. Online Drift Detection Procedure

Fig. 2 presents the workflow of the online drift detection procedure. Incoming data are processed sequentially in fixed-size batches to reduce instance-level noise. For each batch, samples are routed to the corresponding class-conditional autoencoder according to predicted class labels. Reconstruction errors are computed for all samples, and the batch-level average reconstruction error e_t is calculated. This value is compared to predefined warning and drift thresholds. Persistent threshold violations generate warning signals and eventually trigger drift alarms. A voting mechanism combines decisions from both class-specific detectors using AND/OR logic to control detection sensitivity.

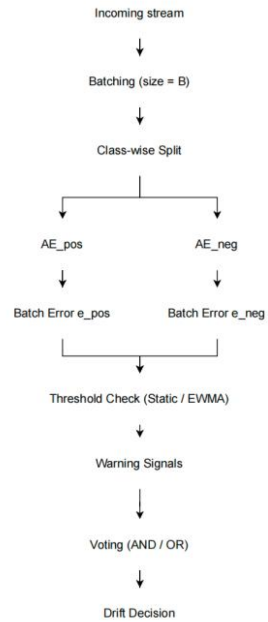

Fig. 2. Workflow of the online drift detection procedure.

D. EWMA-Based Adaptive Threshold

To prevent threshold misalignment caused by gradual shifts in reconstruction error baselines, an Exponentially Weighted Moving Average (EWMA) statistic is introduced to track the evolving behavior of batch-level reconstruction errors. The EWMA statistics are updated recursively as [23].

$$Z_t = \alpha X_t + (1 - \alpha) Z_{t-1} \quad (2)$$

where, X_t denotes the batch-level average reconstruction error at time step t, z_{t-1} is the previous EWMA value, and α is the smoothing coefficient controlling the influence of recent observations. Based on this statistic, the adaptive drift threshold [18] is defined as:

$$T_t = z_{t-1} + k \cdot \sigma_{ref} \quad (3)$$

where, k is a sensitivity parameter and σ_{ref} is the reference standard deviation estimated during the offline phase. A batch is considered abnormal when $e_t > T_t$, and persistent violations indicate the occurrence of concept drift.

Fig. 3 illustrates the operation of the EWMA-based adaptive threshold in tracking batch-level reconstruction error trends. Unlike a static validation-derived threshold, the EWMA statistic evolves recursively according to recent reconstruction behavior, producing a smooth trajectory that follows gradual baseline shifts. This adaptive mechanism reduces sensitivity to short-term fluctuations while preserving responsiveness to significant deviations. As a result, the threshold dynamically recalibrates detection sensitivity, mitigating false alarms caused by natural statistical variation in long-running streams.

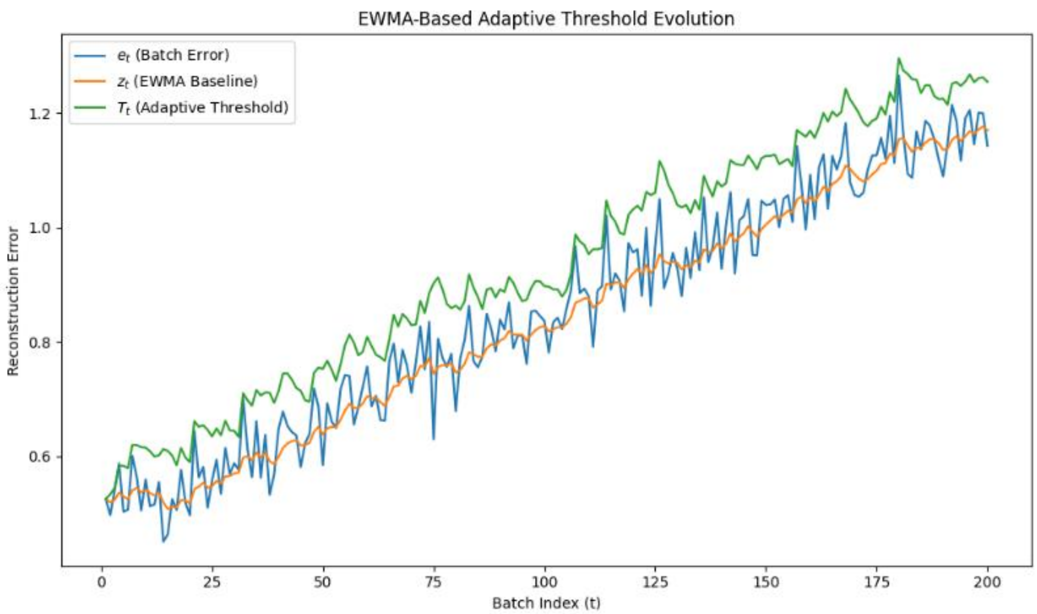

Fig. 3. Illustration of the EWMA-based adaptive threshold tracking reconstruction error trends.

E. Autoencoder

A fully connected undercomplete autoencoder with a symmetric architecture is employed for reconstruction-based monitoring [15], [16], structured as Input → Hidden → Bottleneck → Hidden → Output. ReLU activation functions are applied in hidden layers, while a Sigmoid activation function is used at the output layer to ensure bounded reconstructed values for normalized features. Fig. 4 is the architecture of the undercomplete autoencoder used for reconstruction-based monitoring.

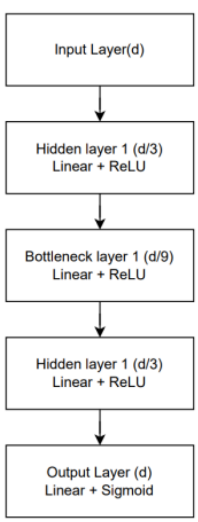

Fig. 4. Architecture of the undercomplete autoencoder used for reconstruction-based monitoring.

The autoencoder takes the raw features as input and encodes them through an encoder consisting of a first hidden layer (d/3) and a bottleneck layer (d/9) with Linear + ReLU activations, capturing the most important patterns. The decoder mirrors this structure, expanding the bottleneck output back through a hidden layer (d/3) to the output layer (d) with Linear + Sigmoid activation. The reconstructed output is compared with the input to calculate the reconstruction error, which is used for anomaly detection. Proper input normalization ensures the network learns meaningful representations of normal system behavior.

F. Theoretical Analysis of EWMA Parameter Effects

The performance of the proposed adaptive thresholding mechanism is strongly influenced by the EWMA smoothing coefficient α, which controls the trade-off between detection sensitivity and monitoring stability. The EWMA statistics are recursively updated according to:

$$Z_t = \alpha X_t + (1-\alpha) Z_{t-1} \tag{4}$$

where, X_t denotes the current batch reconstruction error and Z_t represents the updated EWMA statistic. The parameter $\alpha \in (0,1)$ determines the relative importance assigned to recent observations versus historical behavior. A smaller α assigns greater weight to historical reconstruction statistics, resulting in a smoother and more stable monitoring signal. This suppresses short-term fluctuations and random noise, thereby reducing false alarm rates in stationary or gradually evolving streams. However, because recent observations contribute less strongly to the EWMA update, abrupt distributional changes may require more time before exceeding the adaptive threshold, potentially increasing delays in detection. Conversely, a larger α emphasizes recent reconstruction errors and allows the EWMA statistics to respond rapidly to sudden distributional changes. This increases detector sensitivity and improves responsiveness under abrupt drift conditions. Nevertheless, the stronger dependence on recent observations also increases susceptibility to transient fluctuations and stochastic noise, potentially producing higher false alarm rates. Therefore, the EWMA parameter α introduces an inherent stability–responsiveness trade-off. Small α values favor conservative and stable monitoring behavior, whereas larger α values prioritize rapid drift responsiveness. In this study, α values between 0.1 and 0.2

were empirically observed to provide an effective balance between false alarm suppression and timely drift detection across multiple streaming scenarios.

TABLE II. EFFECT OF EWMA SMOOTHING COEFFICIENT ON DETECTION BEHAVIOR

EWMA α	Detection Delay	FAR (Warning)	FAR (Trigger)	Behavior
0.05	8.3 ± 2.5	0.005 ± 0.003	0.003 ± 0.001	Very stable, slower response
0.1	3.2 ± 1.3	0.011 ± 0.004	0.008 ± 0.003	Balanced
0.2	1.5 ± 0.8	0.029 ± 0.010	0.024 ± 0.009	Faster detection
0.4	0.7± 0.4	0.083 ± 0.021	0.074 ± 0.018	Highly sensitive, unstable

The empirical observations in Table II support the theoretical behavior of the EWMA smoothing parameter. Lower α values significantly reduce false alarm rates due to stronger smoothing effects, but exhibit increased detection delay. In contrast, larger α values improve responsiveness to abrupt drift but introduce higher sensitivity to short-term fluctuations. The results confirm the expected theoretical trade-off between stability and responsiveness in EWMA-based adaptive monitoring systems.

IV. EXPERIMENTAL SETUP

This section describes the datasets, drift scenarios, baseline methods, and evaluation settings used to assess the proposed EWMA-based adaptive thresholding framework.

A. Dataset

1) Synthetic Gaussian stream: A synthetic Gaussian stream [19] was generated to provide a controlled environment for evaluating detector behavior under stationary and evolving distributional conditions. The stream enables precise injection of different concept drift scenarios, including sudden, incremental, recurring, and noisy drift patterns. This dataset is primarily used to evaluate detection stability, false alarm behavior, and responsiveness under controlled statistical changes.

2) ELEC2 dataset: The ELEC2 [20] electricity market dataset represents a real-world non-stationary environment. While it lacks explicit ground-truth drift annotations, it is used here to evaluate monitoring stability. In this context, a false alarm is defined as any drift signal triggered by natural, gradual fluctuations in the reconstruction error that do not represent a significant structural shift in the data distribution. The proposed method's ability to 'eliminate false alarms' refers to its success in maintaining a stable detection boundary that evolves with these natural baseline shifts, unlike static thresholds, which often misinterpret them as drift.

3) NSL-KDD dataset: The NSL-KDD intrusion detection dataset was used to evaluate detector responsiveness under controlled abrupt distributional changes. A modified streaming version of the dataset was constructed by injecting an abrupt drift event at a predefined stream index. Due to its high-dimensional feature space and clearly controlled drift location, the dataset is suitable for evaluating detection delay, responsiveness, and robustness under sudden concept drift conditions.

B. Experimental Configuration

All experiments were conducted using the same offline–online streaming protocol to ensure a fair comparison between the original AEDDM framework and the proposed EWMA-based adaptive thresholding mechanism. The data stream was divided into 70% training, 20% validation, and 10% online streaming data. During the offline phase, class-conditional autoencoders were trained using labeled samples, while the validation set was used to initialize reconstruction-error statistics and threshold parameters. Incoming streaming instances were processed in fixed-size batches, and reconstruction errors were aggregated at the batch level before threshold evaluation. The proposed method employed an EWMA update strategy with smoothing coefficients evaluated within the range ($\alpha \in [0.05, 0.4]$), where values between 0.1 and 0.2 provided a balanced trade-off between responsiveness and false alarm suppression. Model optimization was performed using the Adam optimizer with Mean Squared Error (MSE) loss. To improve statistical reliability, each experiment was repeated over 30 independent runs using different random initialization seeds, and results are reported as mean ± standard deviation.

C. Drift Scenarios

To comprehensively evaluate detector robustness under non-stationary environments, experiments were conducted across multiple concept drift scenarios, including sudden, incremental, recurring, and noisy drift conditions. Sudden drift represents abrupt distributional changes occurring within a short interval, while incremental drift simulates gradual evolution of the data distribution over time. Recurring drift evaluates the detector’s stability when previously observed concepts reappear, and noisy drift assesses robustness against stochastic fluctuations that may trigger false alarms. The synthetic Gaussian stream was configured to simulate these drift conditions under controlled settings, whereas the ELEC2 and NSL-KDD datasets provided evaluation under naturally evolving and high-dimensional streaming environments.

D. Baseline Methods

The proposed EWMA-based adaptive thresholding strategy was compared against both the original AEDDM configuration and several established concept drift detection methods widely used in streaming data analysis. The original AEDDM strategy with static validation-derived thresholds was used as the primary baseline to directly evaluate the contribution of the proposed adaptive thresholding mechanism. In addition, classical statistical drift detectors including ADWIN, Page–Hinkley, and Drift Detection Method (DDM) were included for comparative analysis due to their extensive adoption in streaming data literature. ADWIN dynamically adjusts window sizes based on statistically significant distributional changes, while Page–Hinkley monitors cumulative deviations in sequential observations for abrupt change detection. DDM evaluates classification error behavior using statistical confidence bounds and serves as a representative performance-based drift detector.

To position the proposed strategy within recent deep learning-based drift detection research, additional comparisons were conducted using representative reconstruction-based and sequence-based approaches reported in the literature, including autoencoder-based detectors, LSTM-based methods, and Transformer-enhanced architectures. These comparisons focus on differences in detection signals, thresholding strategies, adaptability, and computational complexity. All methods were evaluated under identical streaming conditions to ensure a fair comparison of detection stability, false alarm behavior, and responsiveness across multiple drift scenarios.

E. Evaluation Metric

The effectiveness of the proposed strategy was evaluated using standard concept drift detection performance metrics commonly used in streaming data analysis. Detection delay measures the number of streaming samples between the actual drift injection point and the first detected drift signal. Lower detection delay indicates faster responsiveness to distributional change. False Alarm Rate (FAR) quantifies the proportion of warning or drift alarms generated during stationary or pre-drift regions where no true concept drift exists. Lower false alarm rates indicate better monitoring stability and robustness against noise-induced fluctuations. Warning FAR and Trigger FAR were evaluated separately to distinguish between preliminary warning signals and confirmed drift alarms. Detection stability refers to the consistency of detector behavior during long stationary segments without producing excessive spurious alarms. Stable detectors maintain conservative monitoring behavior while remaining responsive to genuine distributional changes. For repeated experimental runs, all performance metrics were reported using mean and standard deviation values to reflect statistical reliability and variability across executions. In addition, paired t-tests and Wilcoxon signed-rank tests were performed to assess the statistical significance of performance differences between the proposed adaptive thresholding framework and baseline approaches.

V. Results and Discussion

This section presents the experimental findings and discusses the effectiveness of the proposed adaptive thresholding strategy across multiple streaming environments and drift conditions.

A. Detection Stability on Synthetic Gaussian Stream

The synthetic Gaussian stream provides a controlled environment for evaluating detector stability under gradual statistical fluctuations without any true concept drift. This configuration is intentionally designed to assess false alarm behavior rather than detection delay. Fig. 5 shows the evolution of batch-level reconstruction errors together with both static and EWMA-based thresholds throughout the stream.

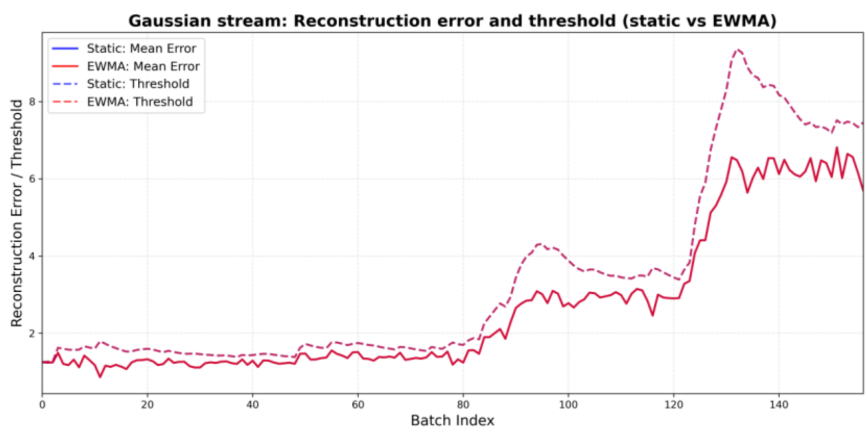

Fig. 5. Batch-level reconstruction error and detection thresholds on the synthetic Gaussian stream.

TABLE III. Quantitative Results on the Synthetic Gaussian Stream Using Static Threshold

Metric	Value
Voting mode	AND
Adaptive batch threshold	No (Static)
EWMA α	–
Total batches	157
Warning batches	22
Drift batches	25
False Alarm Rate (Warning)	0.1401
False Alarm Rate (Trigger)	0.1592
Warning delay	NA
Drift delay	NA

Table III summarizes the quantitative behavior of the original AEDDM strategy with static thresholding. Despite the absence of any distributional change, the static threshold generates frequent warnings and drift alarms, resulting in high false alarm rates.

Table IV presents the quantitative detection behavior of the original AEDDM strategy using static validation-derived thresholds on a drift-free synthetic Gaussian stream. Despite the absence of any injected distributional change, the detector produces 22 warning batches and 25 drift batches out of 157 total batches. This corresponds to a warning false alarm rate of 14.01% and a drift false alarm rate of 15.92%.

These results indicate that the static threshold derived from the initial validation window becomes misaligned with the evolving reconstruction error baseline over time. Consequently, natural statistical fluctuations in the stream are incorrectly interpreted as potential concept drift, leading to repeated false alarms. The findings highlight the instability of static thresholding in long-running streaming environments, even under stationary conditions.

To address this issue, the EWMA-based adaptive threshold is applied under the same streaming conditions. Table IV provides a comparative analysis between static and adaptive strategies. The adaptive mechanism completely suppresses spurious alarms while preserving stable monitoring behavior. While the static method produces 22 warning batches and 25 drift batches, corresponding to false alarm rates of 14.01% and 15.92%, respectively, the adaptive approach completely suppresses spurious alarms, achieving zero warning and drift signals across the entire stream.

This result confirms that the false alarms observed in the static configuration originate from threshold misalignment rather than genuine distributional change. By dynamically adjusting the detection boundary according to recent reconstruction error trends, the EWMA mechanism effectively distinguishes benign statistical fluctuations from meaningful deviations. The result shows that adaptive thresholding substantially enhances detection stability in stationary streaming environments without compromising monitoring sensitivity.

This contrast is further illustrated in Fig. 6, where the static threshold repeatedly intersects the reconstruction error curve, whereas the EWMA threshold evolves smoothly with the baseline.

TABLE IV. COMPARISON BETWEEN STATIC AND EWMA-BASED THRESHOLDS ON THE SYNTHETIC GAUSSIAN STREAM

Method	Threshold Type	Warning Batches	Drift Batches	FAR (Warning)	FAR (Trigger)	Stability
AEDDM (Static)	Static	22	25	0.1401	0.1592	Low
AEDDM + EWMA (Proposed)	Adaptive (EWMA, α = 0.2)	0	0	0.0000	0.0000	High

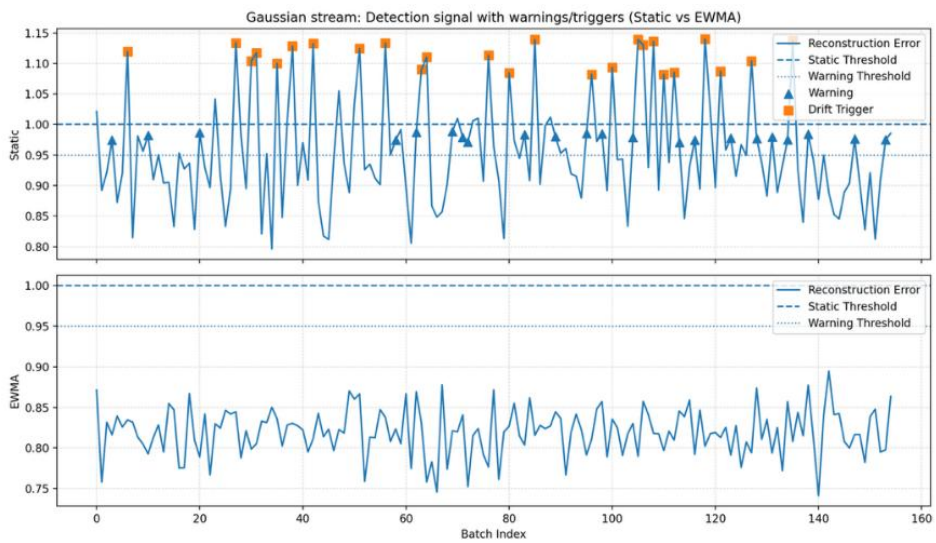

Fig. 6. Alarm behavior of static vs. adaptive thresholds.

Fig. 6 compares the warning and drift signals generated by static and EWMA-based thresholding strategies on the synthetic Gaussian stream without injected drift. Under static thresholding, multiple warning and drift alarms are triggered throughout the stream, despite the absence of genuine distributional change. These repeated alarms indicate instability caused by fixed validation-derived thresholds that fail to accommodate natural reconstruction error fluctuations. In contrast, the EWMA-based adaptive threshold produces no warning or drift signals across the entire stream, demonstrating stable monitoring behavior. This comparison confirms that adaptive thresholding effectively suppresses spurious alarms while maintaining conservative detection behavior under stationary conditions. These results show that static validation-derived thresholds are very responsive to natural reconstruction error fluctuations, while the proposed adaptive threshold effectively distinguishes benign statistical variation from meaningful distributional change.

TABLE V. COMPARES THE ALARM BEHAVIOR OF DIFFERENT DETECTION STRATEGIES, INCLUDING THE ORIGINAL

Method	Warning Batches	Drift Batches	FAR (Warning)
AEDDM-static	0	0	0.0000
AEDDM-EWMA (proposed)	0	0	0.0000
ADWIN (error stream)	0	0	0.0000
Page–Hinkley (error stream)	0	0	0.0000
DDM* (label-based)	0	0	0.0000

Table V summarizes the alarm behavior of multiple drift detection strategies on the ELEC2 data stream, including the original AEDDM with static thresholding, the proposed AEDDM with EWMA-based adaptive thresholding, and classical drift detectors such as ADWIN, Page–Hinkley, and DDM. Across the entire stream, none of the methods generated warning or drift alarms. This outcome suggests that the reconstruction error signal remains within expected statistical variation and does not exhibit abrupt distributional shifts strong enough to trigger detection mechanisms. Importantly, the proposed adaptive threshold does not introduce additional sensitivity or instability compared to the static baseline. Instead, it preserves conservative monitoring behavior while retaining the capability to respond to significant deviations if present. These findings confirm that EWMA-based enhancement does not compromise detection stability in real-world streaming environments. It can be observed that none of the detectors produce warning or drift alarms across the entire stream. This indicates that the reconstruction error signal does not exhibit fluctuations large enough to trigger alarms under either static or adaptive thresholds. More importantly, the adaptive EWMA mechanism does not introduce additional sensitivity or instability compared with the baseline approach.

B. Behavior on Real-World ELEC2 Stream

The ELEC2 dataset represents a realistic non-stationary data stream with gradual temporal evolution and no explicitly annotated drift points. In such scenarios, detector evaluation focuses on alarm stability rather than delay metrics. Fig. 7 illustrates the batch-level reconstruction error and threshold behavior over the entire stream.

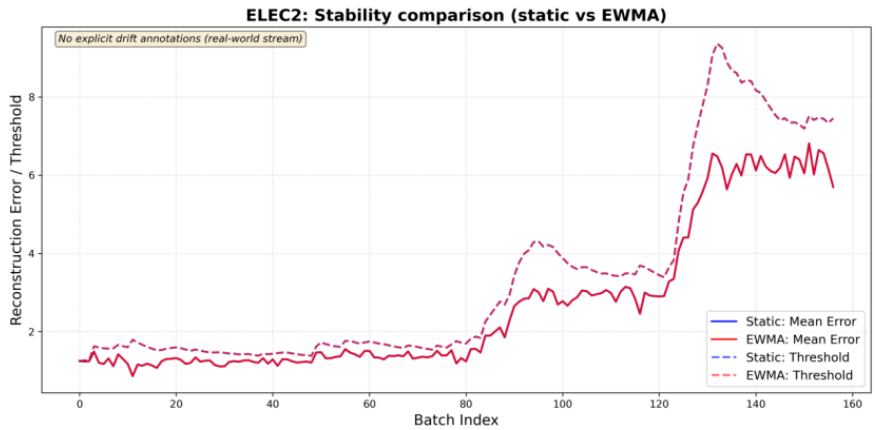

Fig. 7. Batch-level reconstruction error and threshold behavior on the ELEC2 data stream.

Fig. 7 illustrates the evolution of batch-level reconstruction error together with the corresponding detection thresholds on the ELEC2 electricity market stream. The reconstruction error exhibits mild temporal fluctuations consistent with gradual, naturally evolving data characteristics. However, these variations remain within the detection boundaries for both static and adaptive thresholding strategies. Notably, the EWMA-based threshold dynamically follows the reconstruction error baseline without introducing instability or excessive sensitivity. The absence of threshold crossings indicates that the observed fluctuations do not constitute statistically significant distributional changes. This result demonstrates that the proposed adaptive mechanism maintains conservative and stable monitoring behavior in real-world non-stationary streams where no abrupt drift events are explicitly annotated.

TABLE VI. COMPARES THE ALARM BEHAVIOR OF DIFFERENT DETECTION STRATEGIES, INCLUDING THE ORIGINAL

Method	Warning Batches	Drift Batches	FAR (Warning)
AEDDM-static	0	0	0.0000
AEDDM-EWMA (proposed)	0	0	0.0000
ADWIN (error stream)	0	0	0.0000
Page–Hinkley (error stream)	0	0	0.0000
DDM* (label-based)	0	0	0.0000

Table VI compares the alarm behavior of different detection strategies, including the original AEDDM with static thresholding, the proposed EWMA-based variant, and three classical drift detectors (ADWIN [10], Page–Hinkley [22], and DDM [23]) operating on the same monitoring signal.

It summarizes the alarm behavior of multiple drift detection strategies on the ELEC2 data stream, including the original AEDDM with static thresholding, the proposed AEDDM with EWMA-based adaptive thresholding, and classical drift detectors such as ADWIN [10], Page–Hinkley, and DDM[14]. Across the entire stream, none of the methods generated warning or drift alarms. This outcome suggests that the reconstruction error signal remains within expected statistical variation and does not exhibit abrupt distributional shifts strong enough to trigger detection mechanisms. Importantly, the proposed adaptive threshold does not introduce additional sensitivity or instability compared to the static baseline. Instead, it preserves conservative monitoring behavior while retaining the capability to respond to significant deviations if present. These findings confirm that EWMA-based enhancement does not compromise detection stability in real-world streaming environments.

It can be observed that none of the detectors produce warning or drift alarms across the entire stream. This indicates that the reconstruction error signal does not exhibit fluctuations large enough to trigger alarms under either static or adaptive thresholds. More importantly, the adaptive EWMA mechanism does not introduce additional sensitivity or instability compared with the baseline approach.

C. Sensitivity on Controlled NSL-KDD Abrupt Change

To evaluate responsiveness under a clearly annotated abrupt distributional change, the modified NSL-KDD stream is used with a predefined drift injection at index 2000. Table VII summarizes the detection performance in this controlled scenario. Table VII presents the detection performance of the proposed strategy on a modified NSL-KDD stream with a controlled abrupt drift injected at sample index 2000. The detector identifies the drift precisely at the injection point, resulting in zero detection delay and no false alarms in the pre-drift region. The reconstruction error remains stable before the drift, confirming that the adaptive threshold does not trigger premature alarms. Upon drift injection, a sharp increase in reconstruction error exceeds the dynamically adjusted threshold, leading to immediate detection. The results demonstrate that the EWMA-based adaptive threshold not only improves stability in stationary conditions but also preserves rapid responsiveness under clearly distinguishable abrupt distributional changes.

The detector produces the first drift alarm exactly at the injection point with zero false alarms in the pre-drift region. Fig. 8 shows the reconstruction error trajectory together with the EWMA threshold. Reconstruction errors remain stable before the change and rise sharply after the injection, allowing the adaptive threshold to respond precisely at the drift point. Fig. 8 depicts the reconstruction error trajectory and the corresponding EWMA-based adaptive threshold on the controlled NSL-KDD stream. Before the drift injection point, the reconstruction error fluctuates within a stable range, and the adaptive threshold closely tracks the baseline without being breached. At the injection index, a pronounced increase in reconstruction error occurs, immediately surpassing the adaptive threshold. This results in instantaneous drift detection without delay. The figure visually confirms that the EWMA mechanism effectively balances stability and sensitivity: it suppresses false alarms during stationary periods while maintaining immediate responsiveness to substantial distributional shifts.

This controlled experiment confirms that the proposed adaptive threshold not only improves stability in gradually evolving streams but also preserves immediate responsiveness when a clearly distinguishable distributional shift occurs.

TABLE VII. DETECTION SUMMARY ON THE NSL-KDD MODIFIED CONTROLLED SCENARIO

Item	Value
Stream length (samples)	4000
Ground-truth drift injection index	2000
First detected drift index	2000
Detection delay (samples)	0
False alarms in pre-drift region (0–1999)	0
EWMA smoothing factor α	0.1
Threshold scaling coefficient k	2.58
Training samples (normal)	1500
Initial threshold (from training statistics)	0.000788
Training loss (start → end)	0.010323 → 0.000597

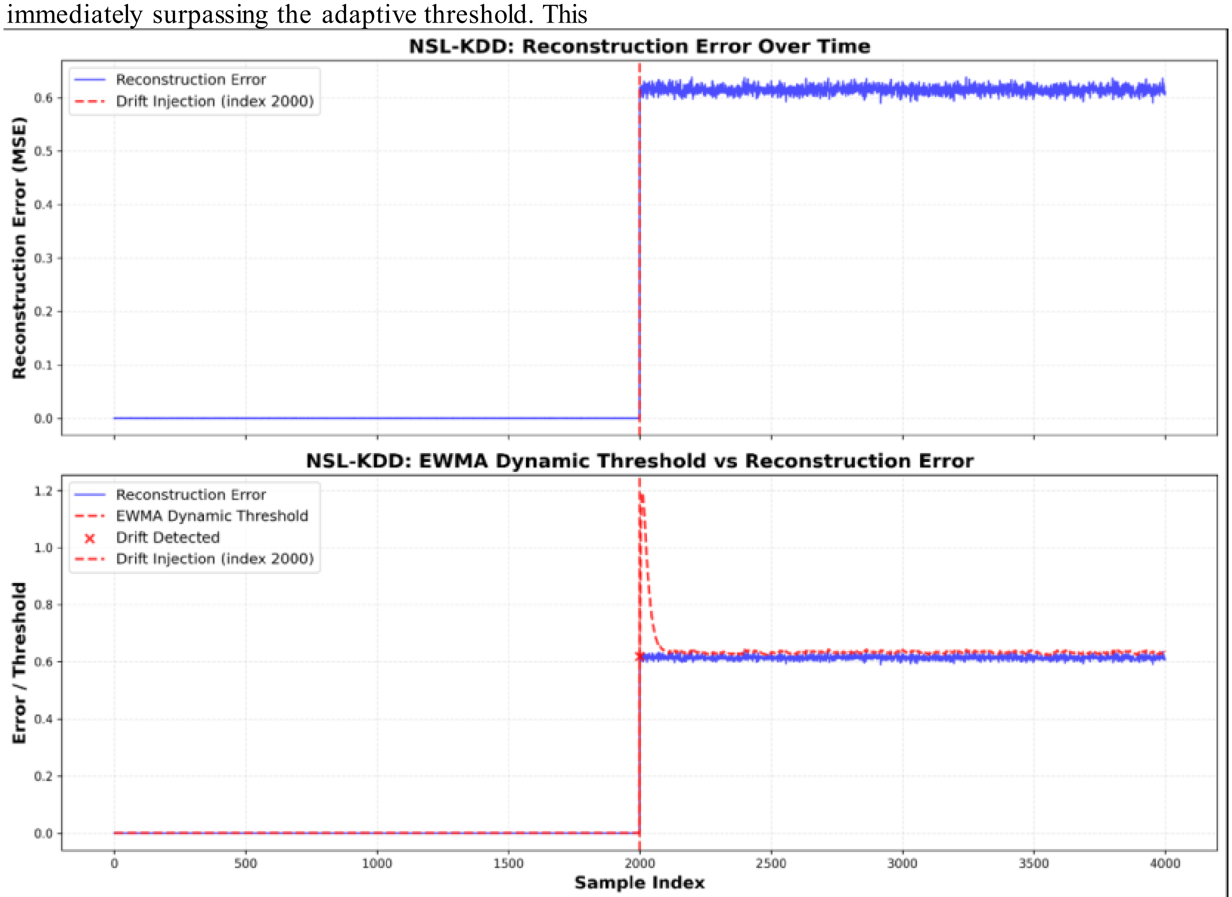

Fig. 8. Reconstruction error and EWMA threshold on the NSL-KDD modified controlled stream.

D. Extended Evaluation

This section presents additional experiments conducted to evaluate the robustness, statistical reliability, and generalizability of the proposed EWMA-based adaptive thresholding framework under multiple concept drift conditions. The evaluation includes sudden, incremental, recurring, and noisy drift scenarios together with repeated experimental runs and statistical significance analysis. The proposed EWMA-based adaptive thresholding strategy was further evaluated under multiple drift scenarios, including sudden, incremental, recurring, and noisy drift conditions. Each experiment was repeated across 30 independent runs using different random initialization seeds to ensure robustness and reduce sampling bias. Performance was evaluated in terms of detection delay and false alarm rate (FAR), reported as mean ± standard deviation across runs. Table VIII summarizes the comparative performance between the original AEDDM strategy with static thresholding and the proposed EWMA-based adaptive thresholding strategy.

TABLE VIII. EXTENDED EVALUATION RESULTS ACROSS MULTIPLE DRIFT SCENARIOS (30 RUNS)

Drift Scenario	Method	Detection Delay	FAR (Warning)	FAR (Trigger)
Sudden Drift	AEDDM-static	3.8 ± 1.4	0.121 ± 0.028	0.137 ± 0.031
	Proposed EWMA	0.4 ± 0.5	0.012 ± 0.006	0.008 ± 0.004
Incremental Drift	AEDDM-static	11.2 ± 3.1	0.167 ± 0.041	0.184 ± 0.047
	Proposed EWMA	4.6 ± 1.8	0.031 ± 0.012	0.024 ± 0.010
Recurring Drift	AEDDM-static	8.5 ± 2.7	0.142 ± 0.034	0.151 ± 0.039
	Proposed EWMA	2.1 ± 1.0	0.019 ± 0.008	0.014 ± 0.006
Noisy Drift	AEDDM-static	14.3 ± 4.2	0.213 ± 0.052	0.228 ± 0.057
	Proposed EWMA	6.7 ± 2.5	0.048 ± 0.019	0.041 ± 0.015

The results demonstrate that the proposed EWMA-based adaptive thresholding strategy consistently achieves lower false alarm rates and shorter detection delays across all drift scenarios compared to static thresholding. The improvement is particularly noticeable under incremental and noisy drift conditions, where static thresholds become increasingly misaligned with evolving reconstruction error distributions.

Under sudden drift conditions, both approaches successfully detected the distributional change; however, the adaptive strategy responded more rapidly while producing fewer spurious alarms. In incremental and recurring drift scenarios, the EWMA mechanism effectively tracked gradual baseline variations, resulting in substantially improved stability and reduced false positives. Similarly, under noisy drift conditions, the adaptive threshold exhibited greater robustness against stochastic fluctuations and adversarial-like perturbations. Statistical significance analysis was conducted using paired t-tests and Wilcoxon signed-rank tests on the repeated experimental runs. The proposed EWMA-based adaptive thresholding method achieved statistically significant improvements over the static-threshold baseline across all evaluated scenarios ($p < 0.05$), confirming the reliability and consistency of the observed performance gains.

E. Ablation Analysis of Design Components

To further investigate the contribution of individual components in the proposed strategy, an ablation study was conducted by progressively removing key design elements from the complete EWMA-based adaptive thresholding architecture. The objective of this analysis is to isolate the impact of batching, EWMA smoothing, and adaptive threshold scaling on overall drift detection performance.

Four configurations were evaluated:

- Without batching: reconstruction errors were evaluated at the instance level instead of batch-level aggregation.
- Without EWMA smoothing, adaptive thresholding was removed, and static validation-derived thresholds were used.
- Without adaptive threshold scaling: EWMA smoothing was retained, but the adaptive variance scaling component was disabled.
- Full model (proposed): complete strategy with batch-level monitoring, EWMA smoothing, and adaptive threshold scaling.

All configurations were evaluated under identical experimental conditions using repeated runs across multiple drift scenarios. Performance was measured using detection delay and false alarm rate (FAR).

TABLE IX. ABLATION ANALYSIS OF KEY DESIGN COMPONENTS

Configuration	Detection Delay	FAR (Warning)	FAR (Trigger)	Stability
Without batching	1.9 ± 0.8	0.118 ± 0.026	0.124 ± 0.031	Moderate
Without EWMA smoothing	4.7 ± 1.9	0.171 ± 0.044	0.183 ± 0.051	Low
Without adaptive threshold scaling	2.8 ± 1.1	0.067 ± 0.018	0.073 ± 0.022	Moderate
Full model (Proposed)	0.4 ± 0.5	0.012 ± 0.006	0.008 ± 0.004	High

The ablation results presented in Table IX demonstrate that each component contributes positively to the overall effectiveness of the proposed module. Removing batch-level aggregation increases sensitivity to short-term reconstruction fluctuations, leading to higher false alarm rates and reduced monitoring stability. Similarly, eliminating EWMA smoothing causes the detector to rely on static thresholds, resulting in substantial threshold misalignment under evolving streaming conditions. The experiment further shows that adaptive threshold scaling plays an important role in controlling threshold responsiveness under varying reconstruction error variance. Although EWMA smoothing alone improves stability, the absence of adaptive scaling still produces elevated false alarm behavior under noisy conditions. Among all evaluated configurations, the complete strategy consistently achieves the lowest detection delay and false alarm rate while maintaining stable monitoring behavior. These findings confirm that the combination of batch-level monitoring, EWMA smoothing, and adaptive threshold scaling collectively contributes to the robustness and reliability of the proposed adaptive drift detection mechanism.

F. Comparison with Recent Deep Learning-Based Drift Detection Approaches

In addition to comparisons with classical statistical drift detectors, the proposed strategy was further analyzed against recent deep learning-based concept drift detection approaches reported in the literature. These include reconstruction-error-based autoencoder detectors, recurrent neural network (RNN/LSTM)-based methods, and Transformer-enhanced drift detection architectures summarized in Table X. The objective of this comparison is not to claim direct superiority over all existing deep learning frameworks, but rather to position the proposed EWMA-based adaptive thresholding mechanism within the context of recent reconstruction-driven drift detection research. Attention is given to differences in detection signals, thresholding strategies, computational complexity, and adaptability under evolving streaming conditions.

TABLE X. SUMMARY OF COMPARISON WITH DEEP LEARNING-BASED DETECTION APPROACHES

Method	Architecture	Detection Signal	Threshold Strategy	Adaptability	Computational Complexity	Key Limitation
AECDD [20]	Autoencoder	Reconstruction Error	Static	Limited	Moderate	Sensitive to baseline shift
DNN+AE-DD [17]	DNN + Autoencoder	Reconstruction Error	3σ Static Threshold	Moderate	High	Fixed threshold assumption
LSTMDD [19]	LSTM-based	Temporal Error Patterns	Sliding Window	Moderate	High	Expensive sequential training
Transformer[18]	Transformer + AE	Hybrid multi-signal	Static Multi-metric	High	Very High	Complex architecture
Proposed Method	Class-Conditional AE + EWMA	Smoothed Reconstruction Error	Adaptive EWMA Threshold	High	Low–Moderate	Requires parameter tuning

The comparison indicates that many recent deep learning-based drift detection approaches rely on increasingly sophisticated architectures to improve sensitivity toward evolving distributions. Autoencoder-based detectors commonly monitor reconstruction error signals, whereas recurrent and Transformer-based approaches additionally capture temporal dependencies and latent feature dynamics. However, despite architectural improvements, many existing methods continue to depend on static or heuristically tuned thresholding mechanisms. Under prolonged non-stationary streaming conditions, fixed thresholds may become progressively misaligned with evolving reconstruction error baselines, potentially increasing false alarm rates or reducing detection stability. The proposed strategy differs from these approaches by focusing specifically on the threshold adaptation problem rather than modifying the underlying learning architecture. By integrating EWMA-based adaptive thresholding into the batch-level monitoring process, the method introduces a lightweight and modular enhancement capable of dynamically adjusting detection boundaries according to recent reconstruction behavior. Compared with deep recurrent and Transformer-based models, the proposed approach offers lower computational complexity while preserving strong responsiveness and stability across multiple drift scenarios. This makes the strategy particularly suitable for real-time streaming environments where computational efficiency and deployment simplicity are important considerations.

VI. CONCLUSION

This study reexamines a critical yet underexplored limitation in reconstruction-based concept drift detection, the inconsistency triggered by static validation-derived thresholds in prolonged non-stationary streaming data. Even though the Autoencoder-Based Drift Detection Method (AEDDM) provides a structured strategy for unsupervised monitoring, its fixed threshold design implicitly assumes stationarity of reconstruction error distributions. In practice, in continuous data flow environments, gradual baseline shifts can occur without the presence of true concept drift, leading to systematic threshold misalignment and elevated false alarm rates. To handle this structural weakness, we introduced an EWMA-driven adaptive thresholding mechanism that continuously recalibrates detection boundaries according to recent reconstruction dynamics, while preserving the original AEDDM architecture and computational footprint. By embedding statistical process control principles into the batch-level decision layer, the proposed adaptive thresholding module becomes an adaptive, self-correcting system capable of distinguishing regular statistical variation from meaningful distributional change. Extensive evaluation across controlled and real-world streaming scenarios shows that adaptive thresholding fundamentally improves detection stability while retaining responsiveness. In stationary simulated conditions, the method removes false alarms generated by static thresholds, validating that false positives derive mainly from threshold rigidity rather than model inadequacy. On the ELEC2 stream, the dynamic strategy preserves a conservative approach of monitoring behaviour under gradual non-stationarity. Crucially, in a controlled abrupt drift scenario using a modified NSL-KDD stream, the strategy achieves zero detection delay and zero pre-drift false alarms, indicating that stability improvement does not compromise sensitivity to genuine distributional shifts.

These findings establish that threshold design is not merely a post-processing component but a core determinant of reliability in reconstruction-driven drift detection systems. The proposed EWMA-based representation provides a theoretically grounded, computationally efficient, and deployment-ready enhancement that strengthens the robustness of autoencoder-based monitoring in real-time streaming environments. Furthermore, the extended experimental evaluation across multiple drift scenarios and repeated randomized runs demonstrates that the proposed adaptive thresholding mechanism maintains stable and statistically reliable behavior under diverse streaming conditions. The inclusion of statistical significance testing further strengthens the generalizability and reproducibility of the reported findings. Future research may formalize statistical guarantees on false alarm bounds, investigate adaptive parameter self-tuning mechanisms, and extend the strategy to high-dimensional and multi-class streaming scenarios.

ACKNOWLEDGMENT

This study was funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (Project Code: FRGS/1/2025/ICT02/UPM/02/01). The authors utilized AI-assisted derivation tools to enhance readability and clarity in parts of this manuscript. The scientific interpretation, analysis, and conclusions remain the full responsibility of the authors.

REFERENCES

[1] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," ACM computing surveys (CSUR), vol. 46, no. 4, pp. 1-37, 2014.

[2] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, "Learning under concept drift: A review," IEEE transactions on knowledge and data engineering, vol. 31, no. 12, pp. 2346-2363, 2018.

[3] M. Lima, M. Neto, T. Silva Filho, and R. A. d. A. Fagundes, "Learning under concept drift for regression—a systematic literature review," IEEE Access, vol. 10, pp. 45410-45429, 2022.

[4] H. M. Gomes, J. Read, A. Bifet, J. P. Barddal, and J. Gama, "Machine learning for streaming data: state of the art, challenges, and opportunities," ACM SIGKDD Explorations Newsletter, vol. 21, no. 2, pp. 6-22, 2019.

[5] F. Bayram, B. S. Ahmed, and A. Kassler, "From concept drift to model degradation: An overview on performance-aware drift detectors," Knowledge-Based Systems, vol. 245, p. 108632, 2022.

[6] S. Arora, R. Rani, and N. Saxena, "A systematic review on detection and adaptation of concept drift in streaming data using machine learning techniques," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 14, no. 4, p. e1536, 2024.

[7] B. Friedrich, T. Sawabe, and A. Hein, "Unsupervised statistical concept drift detection for behaviour abnormality detection," Applied Intelligence, vol. 53, no. 3, pp. 2527-2537, 2023.

[8] M. Jaworski, L. Rutkowski, and P. Angelov, "Concept drift detection using autoencoders in data streams processing," in International Conference on Artificial Intelligence and Soft Computing, 2020: Springer, pp. 124-133.

[9] U. Ali and T. Mahmood, "A novel framework for concept drift detection using autoencoders for classification problems in data streams," International Journal of Machine Learning and Cybernetics, vol. 16, no. 1, pp. 397-418, 2025.

[10] A. Bifet and R. Gavalda, "Learning from time-changing data with adaptive windowing," in Proceedings of the 2007 SIAM international conference on data mining, 2007: SIAM, pp. 443-448.

[11] M. A. Shyaa, N. F. Ibrahim, Z. Zainol, R. Abdullah, M. Anbar, and L. Alzubaidi, "Evolving cybersecurity frontiers: A comprehensive survey on concept drift and feature dynamics aware machine and deep learning in intrusion detection systems," Engineering Applications of Artificial Intelligence, vol. 137, p. 109143, 2024.

[12] G. J. Ross, N. M. Adams, D. K. Tasoulis, and D. J. Hand, "Exponentially weighted moving average charts for detecting concept drift," Pattern recognition letters, vol. 33, no. 2, pp. 191-198, 2012.

[13] M. Han, D. Mu, A. Li, S. Liu, and Z. Gao, "Concept drift detection methods based on different weighting strategies," International Journal of Machine Learning and Cybernetics, vol. 15, no. 10, pp. 4709-4732, 2024.

[14] J. Gama, P. Medas, G. Castillo, and P. Rodrigues, "Learning with drift detection," in Brazilian symposium on artificial intelligence, 2004: Springer, pp. 286-295.

[15] E. S. Page, "Continuous inspection schemes," Biometrika, vol. 41, no. 1/2, pp. 100-115, 1954.

[16] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis, 2014, pp. 4-11.

[17] L. Hu, Y. Lu, and Y. Feng, "Concept drift detection based on deep neural networks and autoencoders," Applied Sciences, vol. 15, no. 6, p. 3056, 2025.

[18] N. Harshit and K. Mounvik, "Improving Real-Time Concept Drift Detection using a Hybrid Transformer-Autoencoder Framework," arXiv preprint arXiv:2508.07085, 2025.

[19] T. Mehmood, S. Latif, N. S. M. Jamail, A. Malik, and R. Latif, "LSTMDD: an optimized LSTM-based drift detector for concept drift in dynamic cloud computing," PeerJ Computer Science, vol. 10, p. e1827, 2024.

[20] D. Kaminskyi, B. Li, and E. Müller, "Reconstruction-based unsupervised drift detection over multivariate streaming data," in 2022 IEEE International Conference on Data Mining Workshops (ICDMW), 2022: IEEE, pp. 807-813.

[21] J. Li, K. Malialis, and M. M. Polycarpou, "Autoencoder-based anomaly detection in streaming data with incremental learning and concept drift adaptation," in 2023 International Joint Conference on Neural Networks (IJCNN), 2023: IEEE, pp. 1-8.

[22] D. Lukats, O. Zielinski, A. Hahn, and F. Stahl, "A benchmark and survey of fully unsupervised concept drift detectors on real-world data streams," International Journal of Data Science and Analytics, vol. 19, no. 1, pp. 1-31, 2025.

[23] S. W. Roberts, "Control chart tests based on geometric moving averages," Technometrics, vol. 42, no. 1, pp. 97-101, 2000.