International Journal of Accounting and Economics Studies, 13 (2) (2026) 123-132

International Journal of Accounting and Economics Studies

Website: www.sciencepubco.com/index.php/IJAES
https://doi.org/10.14419/7x458m39
Research paper

Self-Adaptive Machine Learning Models for Financial Risk Forecasting: Handling Non-Stationarity in Banking and Cryptocurrency Time Series

Proshanta Kumar Bhowmik [1]*, Dil Tabassum Subha [2], Abdur Rahim [3], Arshad Ali Mohammed [4], Maksuda Begum [5], Redoyan Chowdhury [6], Tanveer Ehsan Chowdhury [7], Mohammed Wajeed [8], Dr. Khan Mohammad Elyas [9], Mahamuda Akter Shati [10]

[1] *MBA in Business Analytics, Trine University, Angola, IN, USA*
[2] *Master of Science in Business Analytics, Grand Canyon University, USA*
[3] *Master of Science in Computer Science, University of New Haven*
[4] *College of Graduate and Professional Studies, Trine University*
[5] *Master of Business Administration, Trine University*
[6] *Management Information System (MIS), International American University*
[7] *Management Information System (MIS), International American University (IAU)*
[8] *Masters of Science in Information Technology Management, St.Francis College*
[9] *International Development Consultant and Researcher, Innovation and Incubation Center for Enterprises (IICE)*
[10] *Master's in Business Analytics, Grand Canyon University*
**Corresponding author E-mail: pbhowmik23@my.trine.edu*

Received: January 17, 2026, Accepted: February 7, 2026, Published: February 15, 2026

Abstract

Financial markets rarely sit still. They drift, jump, and occasionally snap into entirely new regimes, with volatility behaving differently depending on the moment. These shifts steadily erode the usefulness of static risk forecasting models. Approaches built on fixed parameters or scheduled retraining tend to lag behind reality, especially when markets are under stress and conditions change fast. This paper investigates whether self-adaptive machine learning models can generate reliable and practical volatility forecasts in live settings, without relying on repeated full retraining cycles. Using daily closing prices from the S&P 500 index and Bitcoin as examples of traditional and cryptocurrency assets, the task is set up as a strictly out-of-sample, one-step-ahead volatility forecasting problem. We compare static models, periodically retrained models, and fully self-adaptive models across calm periods and clearly defined stress regimes, including the COVID-19 crash and major cryptocurrency boom and bust cycles. The results show that self-adaptive models deliver stronger performance across both asset classes. They achieve lower forecast errors, more stable errors during stress, quicker recovery after shocks, and sharply lower computational demands. Recursive EWMA-based models stand out for their solid performance using constant memory and lightweight updates, while online gradient-based learners adapt flexibly without drifting into instability. Taken together, the evidence points to continuous adaptation as a structurally superior approach for financial risk forecasting under non-stationarity. The study demonstrates that self-adaptive models can be deployed in real-time risk systems and sets the stage for future work that connects adaptive forecasting with evolving explainability and regulatory reporting requirements.

Keywords: *Adaptive Learning, Banking Risk; Cryptocurrency; Financial Risk Forecasting; Non-Stationarity.*

1. Introduction

1.1. Background and motivation

Financial risk forecasting sits at the center of how banks, asset managers, and regulators operate day to day. Decisions about capital allocation, portfolio design, stress testing, and regulatory compliance all depend on having risk estimates that are both accurate and timely. The difficulty is that financial time series almost never behave in a stable way. Market behavior shifts in response to macroeconomic shocks, policy actions, technological change, and feedback from investor behavior. These forces produce structural breaks, clustered volatility, and shifting regimes. Such patterns are not edge cases. They define how real markets behave, especially during periods of stress. Ray (2025) shows that major financial disruptions often spread across multiple markets at the same time, reinforcing the idea that risk evolves over time and across assets, instead of remaining fixed or isolated [21]. This creates tension with many traditional risk models, which rest on parameters estimated from past data and treated as stable.

Volatility modeling has long been a cornerstone of financial risk measurement, particularly for widely used metrics such as Value at Risk. One of the most influential practical approaches is the RiskMetrics framework developed in the mid 1990s, which brought exponential weighting into mainstream risk practice through the use of exponentially weighted moving averages [24]. EWMA gained traction because it is easy to implement and because it places more emphasis on recent observations, allowing volatility estimates to react as market conditions evolve. By moving away from the idea of constant variance, EWMA represented an early attempt to make adaptation operational in risk models. Its widespread use across financial institutions reflects a clear recognition that adaptive behavior matters in settings shaped by non-stationarity. Later work took a closer look at the strengths and limits of the RiskMetrics approach. Danielsson and de Vries (1997) offered one of the first systematic evaluations of EWMA-based volatility estimation, pointing out where it performs well and where it struggles, especially during extreme market movements and in the presence of heavy-tailed returns [6]. Their findings showed that EWMA adapts more quickly than fixed variance estimates, yet it still tends to understate risk during severe tail events. Zumbach (2007) extended this line of research by introducing variants such as skewed EWMA models, designed to capture asymmetries and tail behavior more accurately, acknowledging that negative shocks often affect volatility differently from positive ones [32]. Taken together, this body of work positions EWMA as a reference point for adaptive volatility modeling, not as a complete solution, but as a baseline against which more advanced adaptive methods can be assessed.

The importance of adaptation becomes even clearer with the rise of cryptocurrencies. These markets run continuously, operate without centralized regulation, and are heavily shaped by speculative forces. The result is extreme volatility and frequent regime changes. Such conditions intensify non-stationarity and expose the fragility of models that remain fixed or update infrequently. At the same time, recent research in financial machine learning has stressed that adaptation under concept drift extends beyond predictive accuracy alone. Shivogo (2025) argues that changing financial populations and market environments call for systems that adjust their behavior over time to remain reliable and fair under concept drift [28]. Although this argument is developed in the context of credit scoring, the core insight carries over to market risk forecasting, where both data distributions and the meaning of risk evolve continuously. This broader literature motivates a fresh examination of adaptive volatility models across traditional markets and cryptocurrency markets alike.

1.2. Problem statement

Despite extensive research on volatility and financial risk, there remains a clear disconnect between methodological advances and what is used in practice. Many operational risk systems still depend on static models trained once on historical data or on models that are retrained at fixed intervals. These approaches are simple to deploy, yet they fit poorly with the realities of non-stationary financial time series. Static models assume that the statistical properties of returns and volatility remain stable over time, an assumption that breaks down during crises, regime changes, and periods of elevated uncertainty. In such moments, these models often deliver acceptable results during calm conditions and then falter when accurate risk estimates matter most. Periodic retraining is often presented as a practical fix, though it brings its own set of difficulties. Retraining requires decisions about window sizes, update frequencies, and validation procedures, all of which depend on choices that lack a clear theoretical anchor. Retraining also carries high computational and operational costs, especially for complex machine learning models. When markets change quickly, retraining schedules can fall behind current dynamics, leaving models misaligned with reality for extended periods. Danielsson and de Vries (1997) noted early on that delayed adaptation leads to systematic underestimation of risk during turbulent episodes, even for models designed to be responsive [6]. This concern grows stronger in modern environments where data volumes are large, and decisions often need to be made in real time.

The problem is amplified by the growing interconnectedness of financial markets. Ray (2025) documents how crises often unfold through linked movements across equities, bonds, and foreign exchange markets, indicating that shifts in risk can occur rapidly and spread widely across asset classes [21]. In settings like these, models updated only occasionally struggle to keep pace with the speed and scale of change. Cryptocurrency markets push this challenge even further, with volatility regimes that can shift within days or hours, making infrequent retraining unsuitable for real-time risk management. At a more abstract level, this challenge can be understood through the lens of concept drift, where the relationship between inputs and risk outcomes evolves. Shivogo (2025) emphasizes that systems facing concept drift need continuous adaptation to remain reliable, rather than relying on fixed assumptions or delayed updates [28]. Although much of the adaptive learning literature focuses on classification tasks such as credit scoring, the same underlying issue applies directly to volatility forecasting. The central problem addressed in this paper is the weakness of static and periodically retrained models in the face of persistent non-stationarity, and the need for adaptive methods that update risk estimates continuously and efficiently as new data arrives.

1.3. Research objective and contibutions

The main objective of this research is to evaluate, in a systematic way, whether self-adaptive machine learning models that update continuously without full retraining can deliver stronger financial risk forecasts under non-stationarity than static models or models retrained at intervals. The study builds on the long-established role of EWMA as a reference point for adaptive volatility modeling in market risk practice, as formalized in the RiskMetrics framework [24]. The aim is to place modern adaptive learning methods within a clear empirical comparison grounded in practical use. The emphasis is not on introducing new modeling ideas for their own sake. The emphasis is on understanding the real trade-offs between adaptation speed, predictive accuracy, stability, and computational efficiency. A key goal is to evaluate adaptive models beyond average performance alone, with close attention to how they behave during periods of market stress. Earlier studies of EWMA-based models show that sensitivity to recent data improves volatility estimation, with limitations that become visible in extreme regimes [6], [32]. By examining EWMA alongside other adaptive approaches such as online stochastic gradient descent, this work seeks to clarify the settings in which simple adaptive mechanisms are sufficient and the settings in which more flexible learning approaches add value. The use of both equity market data and cryptocurrency data allows the analysis to cover a wide range of non-stationarity, from relatively established markets to highly speculative and fast-moving ones.

This objective is also shaped by a broader shift in financial machine learning toward adaptive systems that operate under concept drift. Shivogo (2025) presents evidence that adaptation is becoming a core requirement across financial applications, highlighting the relevance of continuous learning approaches across domains [28]. Ray (2025) points to the growing importance of machine learning methods for anticipating and managing systemic risk across markets, suggesting that stronger volatility forecasts support a deeper understanding of financial instability [21]. Within this wider setting, the specific objective of this paper is to determine whether self-adaptive volatility models can function as reliable, low-overhead components of real-time financial risk systems, and to identify the market conditions under which their advantages over static or retrained models become most visible.

This paper contributes to the study of financial risk forecasting under non-stationarity in several ways. It introduces a unified experimental framework that compares static models, periodically retrained models, and fully self-adaptive models using the same data sources, prediction targets, and evaluation procedures across equity and cryptocurrency markets. By anchoring the analysis in clearly defined stress regimes drawn from observed market events, the study shifts attention beyond average accuracy toward recovery behavior, stability, and failure patterns that matter for operational risk management. The paper also offers a detailed empirical evaluation of computational efficiency, bringing operational considerations into the discussion of continuous adaptation versus retraining-based strategies. Taken together, these contributions provide evidence-grounded guidance on the use of self-adaptive machine learning models in real-time financial risk forecasting and clarify the settings in which lightweight adaptive approaches deliver meaningful advantages over more complex or resource-intensive alternatives.

2. Literature Review

2.1. Financial risk and volatility modeling

Traditional econometric frameworks, such as GARCH, formalized volatility as an evolving process. While these models remain the gold standard in regulatory and accounting work, their reliance on fixed parameters and stationarity assumptions often leads to calibration failure during structural breaks or extreme market regimes. Realized volatility measures improve precision by capturing actual price variability, while long memory dynamics suggest that simple exponential smoothing, such as EWMA, can approximate, but not fully capture, shifting persistence structures. In cryptocurrency markets, machine learning methods allow flexible modeling of extreme volatility, whereas socioeconomic ML highlights the generality of adaptive prediction under non-stationary conditions [3], [2], [19], [14], [22]. Taken together, these studies demonstrate that while traditional and realized-volatility approaches provide a foundation for financial risk modeling, persistent non-stationarity in both mature and emerging markets motivates exploration of models capable of continuous adaptation.

2.2. Machine learning for financial time series

Deep learning architectures, particularly Long Short-Term Memory (LSTM) networks, have demonstrated superior capacity for capturing non-linear return dependencies compared to classical benchmarks. However, the typical offline training paradigm used in these studies results in static models that lack the parameter-level flexibility to respond to shifting data-generating processes [9], [26], [13], [19]. Applications to cryptocurrency time series underscore this limitation, as high volatility and rapid regime shifts can degrade predictive performance when model parameters remain fixed. Similarly, supervised learning approaches applied to socioeconomic risk assessment, such as income prediction and disparity modeling, reveal that static ML models share vulnerabilities with traditional econometric techniques under evolving patterns. These findings collectively emphasize that adaptability, rather than model capacity alone, is essential for robust forecasting in dynamic financial environments.

2.3. Non-stationarity and concept drift

Non stationarity in financial time series is often discussed through the idea of concept drift, where the statistical relationship between inputs and targets changes over time. Gama et al. (2014) offer a detailed taxonomy of concept drift, separating it into sudden, gradual, incremental, and recurring forms, and reviewing adaptation strategies developed across machine learning research [11]. This perspective maps naturally onto financial markets, where drift can arise from economic cycles, regulatory shifts, or changes in market sentiment. In practice, financial time series rarely exhibit a single clean form of drift. Multiple types often overlap, which makes both detection and response more difficult. A number of studies propose explicit techniques for managing concept drift in financial forecasting. Cavalcante et al. (2016) introduced a framework that combines online sequential extreme learning machines with explicit drift detection, reporting improved forecasts when drift is identified, and models are retrained in response [4]. While promising in controlled experiments, this approach depends heavily on reliable drift detection, which is challenging in financial data that tends to be noisy and ambiguous. Oliveira et al. (2017) explored a particle swarm optimization-based forecasting system designed to handle drifting environments, showing gains under changing conditions while still relying on reoptimization steps when shifts occur [18]. These methods show that recognizing drift can improve performance, yet they often rely on retraining or reconfiguration stages that introduce delays and added computational burden.

There is also work that focuses on domain-specific drift detection for finance. Neri (2021) proposed drift detectors tailored to financial time series, arguing that generic detectors may miss patterns unique to market data [17]. Such tools can improve sensitivity to structural change, though they also increase system complexity and place greater emphasis on deciding when adaptation should be triggered. A different viewpoint prioritizes continuous adaptation over explicit detection. Shivogo (2025), working in the area of credit scoring, demonstrates that both predictive performance and fairness can deteriorate under concept drift when models are not updated continuously, and argues for systems that evolve alongside incoming data [28]. While this work centers on classification and fairness, the underlying message transfers directly to volatility forecasting, where delayed adaptation can be costly during periods of market stress.

2.4. Adaptive and online learning methods

Adaptive and online learning methods offer a natural way to keep models up to date as new data comes in. Shalev Shwartz (2012) provides a foundational discussion of online learning and online convex optimization, laying out theoretical guarantees for methods such as online stochastic gradient descent under both adversarial and stochastic settings [25]. This body of work shows that a model can maintain low regret relative to the best fixed predictor in hindsight, even when the data distribution keeps changing. That insight fits financial risk forecasting well, where change is expected rather than unusual. In practical use, adaptive learning often appears through recursive estimation, exponential smoothing, or incremental updates to model parameters. EWMA-based approaches remain among the simplest and most widely used adaptive tools in finance, continuously updating volatility estimates as each new observation arrives. Their appeal comes from the fact that they adapt in real time without retraining, which makes them well-suited for operational settings. More sophisticated online learning algorithms build on the same idea by updating many parameters at once, supporting richer model structures while keeping computational costs under control.

Adaptive learning also shows up in financial decision-making beyond market risk. Hasan et al. (2025) examine explainable machine learning systems for supplier credit approval in data-sparse environments, pointing to the importance of adaptability when data distributions evolve [13]. Their work centers on explainability and credit decisions rather than volatility forecasting, yet it still underlines the practical value of adaptive learning in financial systems that operate under uncertainty and ongoing change. Across these different applications, a consistent theme appears. Models must update incrementally, run with low computational overhead, and remain dependable as conditions shift. Even so, much of the adaptive learning literature stays focused on theory or on domains outside finance. Empirical studies that apply online learning ideas to financial volatility forecasting are still relatively scarce, particularly those that compare adaptive models directly with static and retrained approaches under realistic stress conditions. This gap suggests that adaptive and online learning methods are well understood at a conceptual level, while their real-world benefits for financial risk forecasting still need careful and systematic empirical validation.

2.5. Research gap

The existing literature shows meaningful progress in volatility modeling, machine learning for financial time series, and approaches to concept drift. At the same time, several gaps remain unresolved. Traditional econometric models such as GARCH offer a clear and structured starting point, and they remain influential in both academic and regulatory work. Their performance weakens when markets exhibit strong non-stationarity. Machine learning models, including LSTM-based architectures, provide flexibility and often achieve higher average accuracy. In practice, they are usually trained in static or batch settings, which limits how quickly they respond as market conditions evolve [10], [31]. Research on concept drift has introduced detection-driven retraining strategies, with an emphasis on identifying discrete change points and reoptimizing models. These designs add delays and operational complexity that matter in live financial systems [4], [18], [15]. A further limitation is the way machine learning is commonly applied in finance. Many studies focus on cross-sectional or static prediction problems instead of time-varying risk dynamics. Reza et al. (2025) highlight this pattern in socioeconomic risk modeling, where predictions are assessed without explicitly accounting for temporal non-stationarity [22]. Even in areas where drift is openly acknowledged, such as credit scoring, the emphasis has often centered on fairness or interpretability, with less attention given to continuous updating of predictive risk estimates over time. Adaptive learning methods grounded in online optimization theory provide a compelling direction, supported by strong theoretical foundations. Their use in market volatility forecasting remains limited in empirical depth and scope [25].
The most notable gap lies in the absence of studies that directly compare fully online adaptive models with static and periodically retrained models under clearly defined financial stress regimes using real market data. Many existing evaluations emphasize average accuracy and overlook recovery behavior, error surges, and stability during crisis periods. This omission is especially important for cryptocurrency markets, where volatility levels are extreme, and regime changes occur rapidly, placing strong pressure on conventional modeling assumptions [14]. Unlike prior studies that rely on batch retraining or external drift detection, this study evaluates inherently self-adaptive volatility models that update continuously without retraining, specifically assessing their stability and recovery behavior under explicitly defined stress regimes. The present study addresses this gap through a systematic and stress-aware comparison of self-adaptive models and traditional baselines across both banking and cryptocurrency settings, with a focus on practical performance, stability, and computational efficiency in non-stationary environments.

3. Methodology

3.1. Data sources and assets

This study makes use of two financial assets chosen to reflect different market structures and risk environments. The S&P 500 index serves as a representative benchmark for traditional banking, equity, and institutional markets. It paints the picture of a mature and highly liquid setting shaped by regulatory oversight, large capital flows, and volatility patterns that tend to evolve over longer horizons. Bitcoin (BTC-USD) is included as a representative cryptocurrency asset as it operates in a decentralized environment with limited formal regulation, strong speculative behavior, high volatility levels, and frequent regime changes. Examining these assets side by side allows the analysis to cover both established financial systems and emerging digital markets, creating a demanding setting for evaluating adaptive risk forecasting models under diverse non-stationary conditions. Daily closing price data for both assets are collected programmatically using the finance library, supporting reproducibility and reducing the risk of manual handling errors. For the S&P 500, the dataset begins on 1 January 2000 and spans multiple macroeconomic cycles, including the dot-com bubble, the global financial crisis of 2008, and the COVID-19 shock. Bitcoin data starts on 1 January 2013, marking the period when daily pricing becomes sufficiently reliable and market liquidity improves to support meaningful volatility analysis. Price series are automatically adjusted for corporate actions where relevant, ensuring consistency in return calculations. These datasets form the empirical base for all modeling, diagnostics, and stress testing carried out in the experimental pipeline.

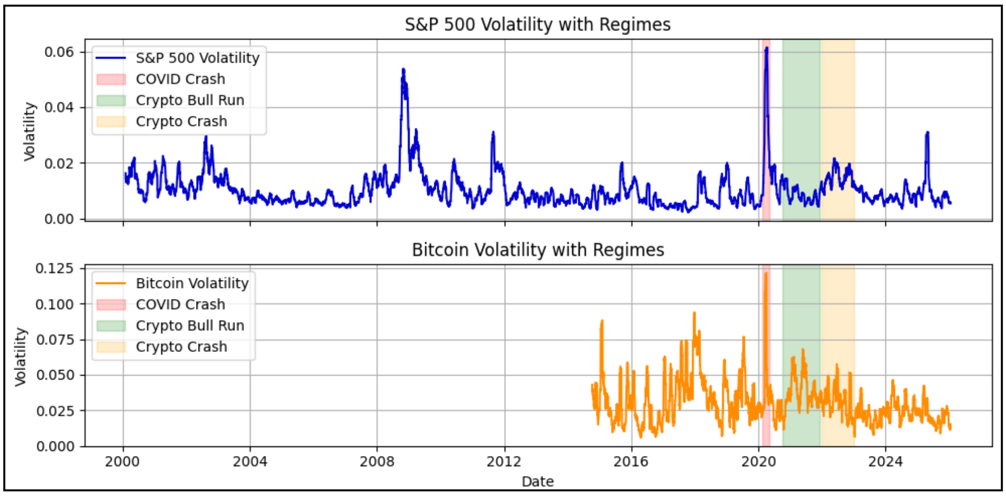

Fig. 1: Bitcoin and S & P 500 Volatility with Regimes.

3.2. Data preprocessing and transformation

First, we go through the price series and remove any missing values, keeping only the closing price. The idea is to focus on the daily picture without getting distracted by tiny intraday swings. After that, we calculate log returns by taking the difference of the natural log of consecutive prices (see Appendix A, Fig.A1). This makes it easier to compare movements across assets and keeps the numbers on a sensible scale, while still reflecting actual market behavior. We measure realized volatility with a 21-day rolling standard deviation of these log returns. Twenty-one days strikes a balance: it reacts to sudden market shocks but isn't overly sensitive to random noise, so it gives a meaningful signal for predicting the next day. Everything is lined up in time properly: we only ever use information up to day t to forecast day t+1. That way, nothing sneaks in from the future. The programmatic setup handles supervised learning sequences, rolling forecasts, and online updates to make sure our tests are genuinely out-of-sample.

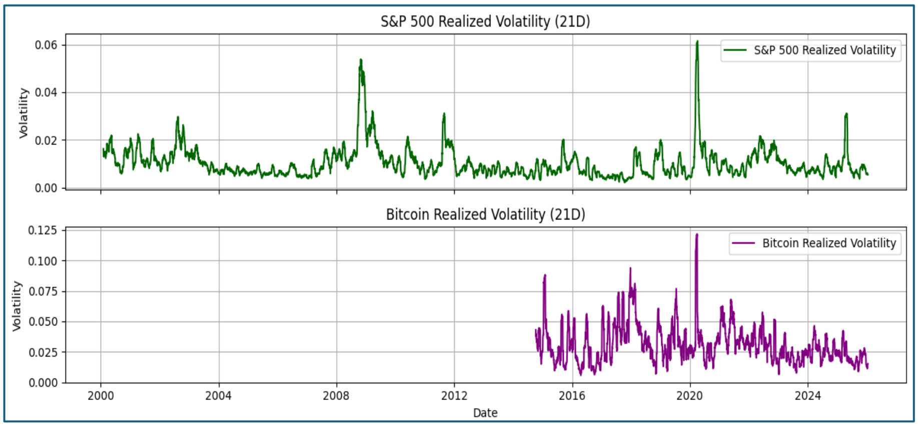

Fig. 2: Bitcoin and S & P 500 Realized Volatility.

3.3. Non-stationarity diagnostics

Looking at the series for the rolling diagnostics (see Appendix A, Fig.A2) over time, we can see that neither one behaves in a perfectly steady way. Bitcoin, in particular, jumps around displaying more variation than other assets. Running Augmented Dickey-Fuller tests shows that returns sometimes seem to reject a unit root, but realized volatility almost always fails the test. In other words, volatility sticks around and evolves in ways that simple stationary assumptions can't capture. Even when returns look a little stable, other aspects like higher moments shift depending on the market conditions. That's why models that keep learning and adjusting as new data comes in are a better fit than ones that assume everything stays fixed

3.4. Problem formulation

The forecasting task is defined as a one-step-ahead volatility prediction problem. At each point in time, the goal is to predict the next day's realized volatility using only information available up to the current day. Realized volatility serves as the prediction target, and the input set includes lagged returns, lagged volatility, and simple nonlinear transformations derived directly from observed values. This setup reflects the realities of operational risk management, where volatility estimates are updated daily and used immediately to inform decisions. Model evaluation follows a strictly rolling out-of-sample procedure. The data are split in chronological order, with no random shuffling. Static models are trained once and evaluated forward through time. Periodically retrained models update parameters at fixed intervals. Self-adaptive models revise parameters incrementally at every step. Performance is measured using error-based metrics computed only on unseen data, alongside stress regime analyses that focus on behavior during extreme market events. This formulation ensures that all reported results correspond to realistic deployment settings, with attention placed on robustness, stability, and recovery in non-stationary financial environments rather than idealized in-sample performance.

3.5. Modeling and evaluation

The experimental setup looks at three modeling approaches, each chosen to reflect a different level of adaptability when the data refuse to sit still over time: static models, periodically retrained models, fully self-adaptive models. These groupings are meant to focus attention on adaptation itself, without muddying the picture through differences in model size, expressiveness, or feature sets.Static models provide the reference point, closely matching what is still common practice in financial risk forecasting. The first static model is a GARCH-based volatility specification, estimated once using an initial historical window, then left unchanged for the remainder of the evaluation. It models conditional heteroskedasticity through a parametric form, while quietly assuming that parameters estimated in the past remain appropriate far into the future. The second static model is a Long Short-Term Memory network, trained a single time on historical observations, then applied without further updates. While the LSTM architecture can learn rich temporal structure, freezing its parameters after training leaves it exposed when market behavior shifts or structural breaks emerge.

Periodically retrained models are included to represent a middle ground that is often suggested in practice. Here, the same model structures used for the static baselines are retained, but their parameters are re-estimated on rolling windows of fixed length. Retraining happens at predefined points in time, with older observations dropped as the window moves forward. This allows the model to absorb newer information while remaining within a batch learning mindset. Retraining frequency plus computational cost are tracked explicitly, making it possible to weigh predictive improvements against operational burden. This mirrors real deployment settings, where retraining introduces delays, resource demands, plus the risk of instability during handover periods. Self-adaptive models sit at the center of this study, built to update continuously without full retraining cycles. The first adaptive method is an online stochastic gradient descent regressor that updates its parameters at each time step using the latest observation. It reflects true online learning, with a fixed update cost per step plus immediate response to new data. The second adaptive model is an exponentially weighted moving average volatility estimator, which updates volatility recursively using a fixed decay rate with constant memory. Despite its straightforward structure, EWMA remains a standard reference point for adaptive risk estimation, making it useful for comparing modern online learning approaches with long-established adaptive techniques. Taken together, these models make it possible to examine how varying degrees of adaptivity shape forecasting accuracy, stability over time, plus practical feasibility in live settings.

3.6. Evaluation metrics

Model performance is assessed using a mix of accuracy, stability, plus efficiency measures intended to capture typical behavior as well as responses to shifting conditions. Predictive accuracy is measured using root mean squared error plus mean absolute error between predicted volatility values versus realized outcomes. These measures offer complementary views of error size plus sensitivity to large deviations. All accuracy metrics are computed strictly out of sample across rolling prediction windows. Beyond summary accuracy, absolute error paths are examined through time to show how performance evolves across regimes, especially during periods of market stress. This time-based view highlights sudden error spikes, extended performance deterioration, plus recovery patterns that aggregate statistics often hide. Computational efficiency is evaluated by recording update frequency, retraining events, plus total computation time for each class of model. For periodically retrained models, retraining cost is measured directly, while for self-adaptive models, per-step update cost is logged. This makes it possible to weigh predictive gains against operational demands, a necessary step when considering real-time deployment.

3.7. Stress testing plus regime analysis

To reveal where models break down while judging robustness, the evaluation includes explicit stress testing grounded in historically observed market regimes. Stress periods are defined through fixed calendar windows tied to widely documented events. These include the COVID-19 market crash, marked by sudden volatility escalation in equity markets; a cryptocurrency bull market phase characterized by rapid price increases plus heightened speculative activity; plus a subsequent cryptocurrency crash phase involving sharp drawdowns with pronounced volatility spikes. Making these regimes explicit keeps the stress tests reproducible while avoiding subjective interpretation. Within each stress regime, model behavior is examined using focused diagnostics that emphasize stability rather than average accuracy. Reported measures include the maximum error spike observed during the stress window, which captures worst-case failure; the variance of prediction errors within the regime, reflecting stability under sustained turbulence; plus the recovery time required for errors to fall back within a predefined tolerance relative to pre-shock levels. Additional diagnostics look at oscillatory responses plus overreaction after shocks, revealing whether adaptive models overshoot or display unstable adjustment patterns. By design, this analysis does not hide or smooth unfavorable outcomes. Instead, it brings instability into view where it actually appears, allowing a transparent comparison of how static, retrained, plus self-adaptive models behave under extreme yet realistic financial conditions.

4. Results

4.1. Overall predictive performance

Across both the S&P 500 index and the BTC-USD cryptocurrency series, clear differences appear among static, periodically retrained, and fully self-adaptive modeling approaches when evaluated using a strictly rolling out-of-sample protocol. Static models, including the once-trained GARCH specification and the fixed LSTM, show the weakest predictive accuracy across both assets. Their RMSE and MAE values remain high and tend to worsen as time passes. This pattern is especially visible in the cryptocurrency data, where volatility behavior shifts quickly and does not align well with the stable assumptions built into fixed-parameter models. Periodically retrained models perform better than static baselines by realigning predictions through rolling-window refitting. This leads to noticeable reductions in both RMSE and MAE. These improvements are inconsistent and depend heavily on retraining frequency and window length. Delays following regime changes remain visible, with prediction errors often lagging behind new market conditions. Self-adaptive models deliver the strongest overall predictive performance in both markets. The online SGD regressor achieves consistently lower average RMSE and MAE by updating parameters with each new observation. The EWMA model remains competitive by continuously revising volatility estimates through its recursive update structure. The ordering of model performance stays consistent across assets, suggesting that the gains from continuous adaptation reflect a structural advantage in dealing with non-stationary financial time series rather than an effect tied to a specific market.

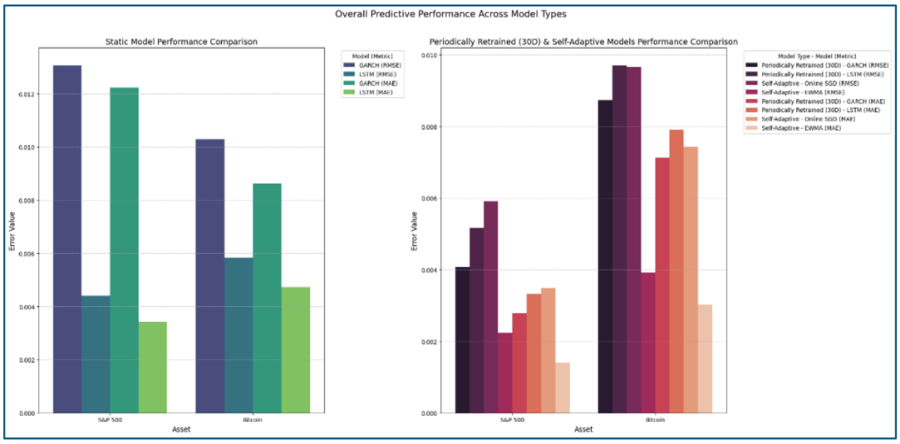

Fig. 3: Static, Periodically Retrained, and Fully Self-Adaptive Modeling Outcomes.

4.2. Stress regime, stability, and recovery results

Differences in model behavior become most apparent during explicitly defined stress regimes, including the COVID-19 equity market crash, the cryptocurrency bull market phase, and the subsequent crypto market collapse. At the start of the COVID-19 shock, static models display sharp and persistent error spikes. These spikes reflect an inability to revise volatility expectations as conditions change. Periodically retrained models reduce the severity of these spikes once retraining occurs. Error correction arrives with a delay, and pronounced overshooting appears immediately after the shock. Comparable dynamics emerge during extreme cryptocurrency regimes. Static and retrained LSTM models struggle to follow rapid volatility expansions and contractions during these periods. Self-adaptive models show a different pattern. Maximum error spikes remain substantially lower across all stress events. The EWMA model reacts quickly to large return shocks by inflating volatility estimates, while the online SGD regressor continuously updates its coefficients, keeping forecasts closer to realized outcomes. Error trajectories during regime transitions appear smoother for adaptive models, with fewer oscillations around retraining points. These results point to adaptation speed as the primary driver of stress-period forecasting performance, with model complexity playing a secondary role.

Looking past peak error behavior, adaptive models also show stronger stabilization and recovery once shocks pass. After each identified stress event, self-adaptive approaches return to pre-shock error levels much sooner than static and retrained models. Static models often do not recover within the observed window, with error variance staying elevated long after volatility conditions settle. Periodically retrained models do regain stability over time, though the path is uneven. Abrupt parameter resets introduce short-lived instability that interrupts recovery. Both the EWMA and online SGD models show a different pattern. Error variance contracts quickly after shocks, and recovery times remain consistently shorter across assets and regimes. Post-shock error variance is also lower, pointing to faster recovery and steadier long-run behavior. This effect stands out in the cryptocurrency series, where adaptive models avoid the extended oscillations seen in retrained LSTM forecasts. These results indicate that continuous adaptation acts as a natural stabilizer during regime shifts, limiting both the size and duration of forecasting errors.

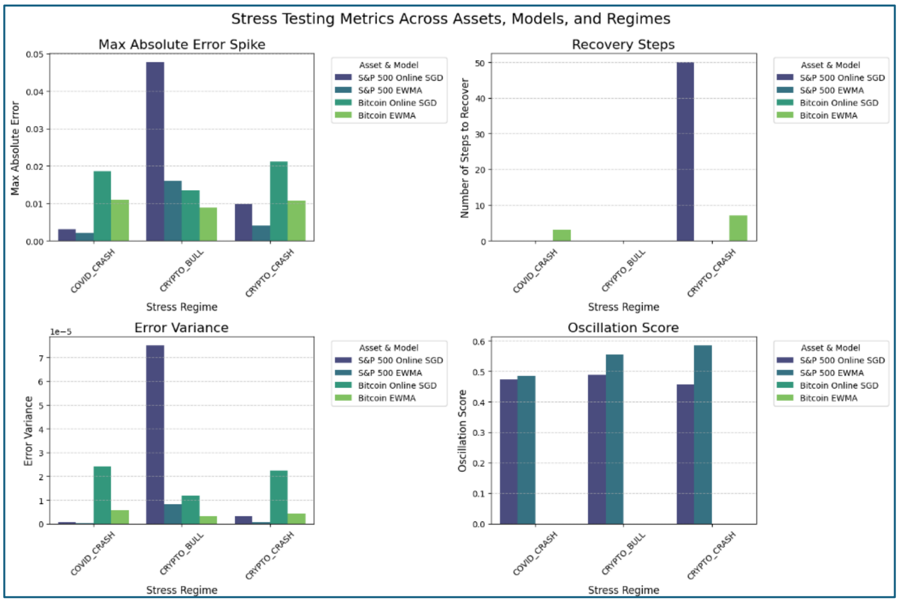

Fig. 4: Stress Regime, Stability, and Model Recovery Outcomes.

4.3. Computational efficiency

From a computational standpoint, self-adaptive models outperform the alternatives by a wide margin. Periodically retrained models carry substantial overhead from repeated batch optimization, with costs rising in step with retraining frequency and window length. This burden becomes especially visible for deep learning models, where retraining introduces delays that clash with real-time risk monitoring needs. Static models demand little computation during inference, though their predictive usefulness fades over time, shifting the burden toward larger forecasting errors. Self-adaptive models deliver stronger predictive accuracy while using far less computation than periodic retraining approaches. The EWMA model runs with constant memory and minimal per-update cost, while the online SGD regressor updates parameters through lightweight incremental steps. Adaptive models also remove the need for fixed retraining schedules, which lowers operational complexity and reduces system brittleness. These efficiency gains support the case for self-adaptive methods in real-world financial risk systems, where low latency, scalability, and resilience to regime change remain core requirements.

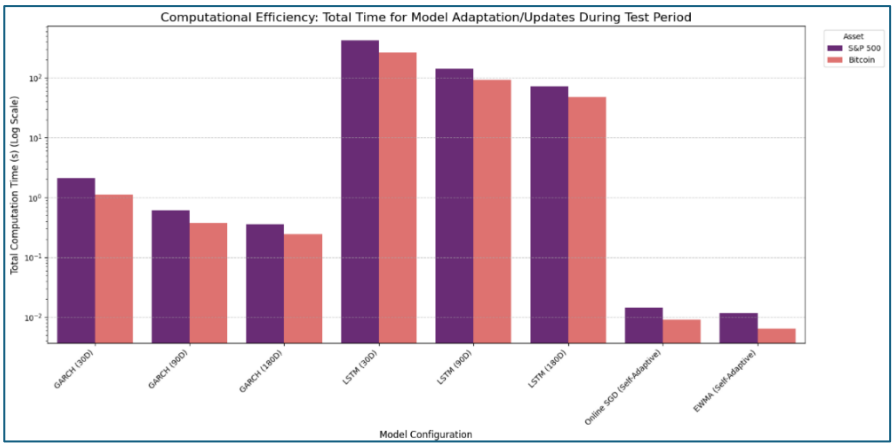

Fig. 5: Computational Efficiency Across Models.

5. Discussion

5.1. Effectiveness of self-adaptive learning

The strong performance of self-adaptive models seen in this study follows naturally from how they stay aligned with data that keep changing, rather than depending on spaced-out retraining cycles that trail behind shifts in market behavior. In financial markets, non-stationarity is not an edge case. Volatility patterns drift, jump, and sometimes change direction without warning, which means parameter estimates age quickly. Online learning treats estimation as something ongoing, not a task completed once and revisited later. Mocanu et al. (2015) make this point clearly, noting that sequential learning frameworks fit financial time series well because they limit regret under shifting distributions and avoid the instability that comes with repeated batch retraining [14].

Evidence from other high-stakes domains reinforces this conclusion. Chouksey et al. (2025) develop an early warning system for financial risk that continuously updates with evolving macroeconomic and sentiment indicators, demonstrating improved timeliness in anticipating stress conditions relative to static frameworks [5]. Similarly, adaptive behavioral machine learning effectively tracks evolving patterns in cross-chain bridge illicit movements, exemplifying robustness to distributional shifts in complex datasets [27]. Research on micro-inflation cluster forecasting with ML shows that ongoing parameter updates improve both forecast accuracy and responsiveness to economic change [20]. Signal-to-noise analyses in global crisis indicators further highlight how adaptive approaches handle emerging patterns more effectively than static alternatives [15]. These findings, together with work showing that continuous adaptation reduces forecast errors and operational overhead, illustrate that self-adaptive learning offers tangible benefits for volatile financial contexts beyond the specific markets studied here [30], [1].

5.2. Banking vs cryptocurrency dynamics

Differences in model behavior across banking and cryptocurrency assets reflect underlying distinctions in volatility formation and persistence. Traditional equity markets, represented by the S&P 500, exhibit measured regime transitions, stronger mean reversion, and stabilizing institutional features such as regulation, liquidity provision, and circuit breakers. Cryptocurrency markets, in contrast, experience abrupt regime shifts, longer volatility clusters, and weaker stabilizing mechanisms, increasing the cost of delayed adaptation. Dyhrberg (2016) documents early evidence that Bitcoin volatility combines features of speculative commodities and alternative hedging instruments [9], explaining why retraining-based strategies often fail in crypto markets. Rolling-window retraining assumes recent history predicts the near future, a condition frequently violated in BTC-USD, leading to larger error spikes and slower stabilization. Self-adaptive methods succeed across both asset classes because updates occur continuously rather than being bounded by fixed windows. Their effectiveness derives not from model complexity but from responsiveness, which allows forecasts to track extended bull and crash phases while controlling noise, reducing error variance, and shortening recovery times. The central takeaway is that model performance depends less on asset type and more on the intensity and persistence of non-stationarity. As portfolios increasingly integrate digital and traditional assets, adaptive risk frameworks are better positioned to maintain stability across diverse volatility environments.

5.3. Practical deployment implications

The results have clear implications for operationalizing real-time financial risk systems, particularly where latency, robustness, and regulatory compliance matter. Self-adaptive models fit naturally into continuous risk monitoring because they remove the need for fixed retraining schedules and reduce the fragility associated with episodic batch optimization. For financial institutions, continuous adaptive models can be integrated into Value-at-Risk (VaR) and Expected Shortfall (ES) reporting, and into supervisory stress-testing pipelines, helping ensure real-time alignment with market conditions and regulatory expectations for timely risk reporting and capital adequacy assessments. Their constant-update mechanisms support strong governance, auditability, and transparent decision-making, which are essential for internal model validation and external audit processes. Analogues from other industries reinforce these practical advantages. Hasan et al. (2025a) and Shawon et al. (2025) show that adaptive machine learning pipelines enhance supply chain resilience by continuously adjusting to shifting conditions [12], [26]. In smart grid systems, adaptive models maintain performance under changing loads and infrastructure dynamics, demonstrating feasibility for systems with constrained resources while preserving operational stability [29]. In cybersecurity and energy infrastructure, adaptive detection frameworks achieve high performance with modest computational overhead, a property that makes large-scale deployment feasible without excessive energy costs [7], [8].

Chouksey et al. (2025), Rahman (2025), and Jakir (2025) further provide evidence that early warning systems using adaptive ML attain timely risk detection, making them relevant to regulatory contexts where proactive identification of stress events can inform supervisory action [5], [20], [15]. Additionally, adaptive unsupervised and anomaly-driven methods demonstrate appeal for scenarios such as anti-money-laundering monitoring, suggesting that evolving models can operate with minimal human intervention and little reliance on labeled data [30]. These cross-domain examples show that adaptive machine learning meets real operational needs by offering scalable support for real-time, regulation-aware risk management and financial reporting. Self-adaptive learning thus sits at the core of modern financial risk infrastructure, well beyond the specific asset classes analyzed here.

6. Limitations and Future Work

Even with the strong empirical support for self-adaptive learning in volatility forecasting under non-stationarity, the study has limits that are worth stating clearly. The modeling setup is intentionally focused on univariate volatility forecasting for individual assets. This choice makes it easier to attribute performance differences to adaptivity itself, not to cross-asset effects, yet it leaves out systemic risk channels such as volatility spillovers, contagion, and correlation breakdowns that matter for portfolio and institutional risk management. The empirical analysis also relies on a narrow asset set consisting of the S&P 500 index and Bitcoin. These assets were chosen to reflect structurally different market types, though they cannot capture the full range of behavior seen in fixed income, foreign exchange, commodities, or emerging market assets. In addition, all models use price-based inputs alone. Macroeconomic variables, order book signals, and flow-based measures are excluded, even though such information can be informative during certain stress episodes. This setup emphasizes adaptivity instead of feature breadth, which may understate the gains possible with richer data.

These same constraints point directly to meaningful directions for future work. One natural extension is the development of multivariate self-adaptive risk models that learn volatility dynamics jointly across assets. This would allow real-time estimation of changing correlations and broader measures of systemic risk, which are especially relevant for portfolio construction, stress testing, and balance sheet management. Another important step is to embed adaptive volatility forecasts into downstream risk metrics such as Value at Risk and Expected Shortfall, enabling a full assessment of adaptive learning within regulatory risk frameworks. Future studies could also examine regime-aware adaptive learning rates, where the intensity of updates shifts with market conditions. This may improve stability during extreme shocks while keeping models responsive during quieter periods.

An especially important area for future research sits at the intersection of adaptivity and explainability. The current study centers on predictive accuracy and stability, though continuously updating models raises questions about transparency, auditability, and fairness as parameters evolve. Shivogo (2025) shows that under concept drift, explanations and fairness constraints need to evolve alongside predictions to prevent misleading interpretations and unintended bias amplification [28]. Applied to financial risk forecasting, this insight points toward adaptive explanation frameworks that change in step with volatility models, keeping risk estimates interpretable and defensible across regimes. Building an explainable, regulation-aware adaptive risk system, therefore, re represents a critical next step toward the practical use of self-adaptive machine learning in financial institutions.

7. Conclusion

This study shows that self-adaptive machine learning models offer a fundamentally stronger way to forecast financial risk in environments shaped by non-stationarity. When volatility prediction is treated as an ongoing learning process instead of a sequence of fixed fits or scheduled retraining steps, adaptive models consistently achieve lower forecast errors, more stable behavior, and quicker recovery after market shocks. Across both the S&P 500 and Bitcoin, these models hold up during extreme regimes where static and retrained approaches suffer large error surges and slow adjustment. The results make a clear point: non-stationarity is not a nuisance that can be engineered away with heavy transformations or frequent refits. It is a defining feature of financial markets that needs to be handled directly in the modeling approach. One of the main contributions of this work is the direct comparison between retraining-based adaptation and true online learning. Retraining is often presented as a practical stand-in for adaptivity, yet the empirical evidence reveals clear shortcomings. Responses to regime shifts arrive late, instability appears around rolling window boundaries, and computational demands grow quickly. Recursive and online models avoid these issues by updating continuously with constant or near-constant resource use, which makes real-time deployment feasible without giving up predictive performance. The strong showing of EWMA-based models also highlights an important lesson. Effective adaptivity does not depend on complex architectures. It depends on well-chosen mechanisms for incremental updates that stay aligned with changing data-generating processes.

Looking across traditional equity markets and cryptocurrency markets strengthens these conclusions. The two asset classes differ sharply in volatility behavior and regime persistence, yet self-adaptive models perform reliably in both cases. This consistency suggests that continuous adaptation provides a common framework that works across very different financial instruments. The implication matters as financial systems grow more interconnected and risk moves quickly across markets. In that setting, adaptive volatility forecasting becomes a key building block for broader systemic risk monitoring and early warning tools. From a practical perspective, the results point strongly toward using self-adaptive models in real-time risk engines, stress testing workflows, and regulatory reporting systems. Their low computational demands, steady behavior under stress, and transparent update rules fit operational environments where reliability and responsiveness matter every day. More broadly, this work frames self-adaptive learning as a central design principle for the next generation of financial risk systems. The emphasis moves away from episodic recalibration and toward continuous alignment with market conditions as they evolve.

By demonstrating superior performance, stability, and computational efficiency under non-stationarity, this study positions self-adaptive machine learning as a practical tool for accounting and regulatory workflows. Adaptive forecasts can directly feed into VaR, Expected Shortfall, and stress-testing reports, providing decision-makers with timely, transparent, and auditable risk metrics. In doing so, the work contributes not only to financial machine learning research but also to the interdisciplinary mission of IJAES, linking data-driven model innovation to actionable economic and accounting decisions. In doing so, the study provides a solid empirical base for future research on adaptive, explainable, and regulation-aware risk modeling in a financial landscape defined by ongoing change.

References

[1] Aashish, K. C., Zamil, M. Z. H., Mridul, M. S. I., Akter, L., Sharmin, F., Ayon, E. H., … Malla, S. (2025). Towards eco-friendly cybersecurity: Machine learning-based anomaly detection with carbon and energy metrics. International Journal of Applied Mathematics, 38(9s). https://doi.org/10.12732/ijam.v38i9s.822.

[2] Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. Econometrica, 71(2), 579–625. https://doi.org/10.1111/1468-0262.00418.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1.

[4] Cavalcante, R. C., Oliveira, G. H. F. M., & Minku, L. L. (2016). An approach to handle concept drift in financial time series based on online sequential extreme learning machines and explicit drift detection. In Proceedings of the International Joint Conference on Neural Networks (IJCNN). https://doi.org/10.1109/IJCNN.2015.7280721.

[5] Chouksey, A., Dola, A., Antara, U. K., Begum, S., Ahmed, T., Sultana, T., & Zabin, N. (2025). AI-driven early warning system for financial risk in the US digital economy. International Journal of Applied Mathematics, 38(9s). https://doi.org/10.12732/ijam.v38i9s.838.

[6] Danielsson, J., & de Vries, C. G. (1997). Evaluating the RiskMetrics methodology in measuring volatility. Journal of Risk, 1(4), 5–26.

[7] Das, B. C., et al. (2025). AI-driven cybersecurity threat detection: Building resilient defense systems using predictive analytics. arXiv preprint arXiv:2508.01422. https://doi.org/10.14419/hysdg957.

[8] Debnath, S., et al. (2025). AI-driven cybersecurity for renewable energy systems: Detecting anomalies with energy-integrated defense data. International Journal of Applied Mathematics, 38(5s). https://doi.org/10.12732/ijam.v38i5s.367.

[9] Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar: A GARCH volatility analysis. Finance Research Letters, 16, 85–92. https://doi.org/10.1016/j.frl.2015.10.008.

[10] Fedorov, V., Panchenko, V., & Shpilkin, A. (2019). Stock price volatility prediction with long short-term memory. CS230: Deep Learning, Stanford University Project Report.

[11] Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 44:1–44:37. https://doi.org/10.1145/2523813.

[12] Hasan, M. R., Rahman, M. A., Gomes, C. A. H., Nitu, F. N., Gomes, C. A., Islam, M. R., & Shawon, R. E. R. (2025). Building robust AI and machine learning models for supplier risk management: A data-driven strategy for enhancing supply chain resilience in the USA. Advances in Consumer Research, 2(4).

[13] Hasan, M. S., et al. (2025). Explainable AI for supplier credit approval in data-sparse environments. International Journal of Applied Mathematics, 38(5s). https://doi.org/10.12732/ijam.v38i5s.380.
[14] Islam, M. Z., et al. (2025). Cryptocurrency price forecasting using machine learning: Building intelligent financial prediction models. arXiv preprint arXiv:2508.01419. https://doi.org/10.14419/s0pktr58.
[15] Jakir, T. (2025). Signal-to-noise analysis of crisis indicators in global finance using artificial intelligence. International Journal of Applied Mathematics, 38(10s), 1815–1836. https://doi.org/10.12732/ijam.v38i10s.1075.
[16] Mocanu, D. C., Fidon, L., & Sminchisescu, C. (2015). Online learning for financial time series prediction. In Lecture Notes in Computer Science (Vol. 8955, pp. 419–428).
[17] Neri, F. (2021). Domain-specific concept drift detectors for predicting financial time series. arXiv preprint arXiv:2103.14079.
[18] Oliveira, G. H. F. M., Cavalcante, R. C., Cabral, G. G., Minku, L. L., & Lacerda, A. L. F. (2017). Time series forecasting in the presence of concept drift: A PSO-based approach. In Proceedings of the IEEE International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 239–246). https://doi.org/10.1109/ICTAI.2017.00046.
[19] Proietti, T. (2014). Exponential smoothing, long memory and volatility forecasting. MPRA Paper No. 57230. https://doi.org/10.2139/ssrn.2475784.
[20] Rahman, M. S. (2025). Machine learning–enabled early warning system for detecting micro-inflation clusters in the US economy. International Journal of Applied Mathematics, 38(12s), 2743–2769. https://doi.org/10.12732/ijam.v38i12s.1585.
[21] Ray, R. K. (2025). Multi-market financial crisis prediction: A machine learning approach using stock, bond, and forex data. International Journal of Applied Mathematics, 38(8s), 706–738. https://doi.org/10.12732/ijam.v38i8s.602.
[22] Reza, S. A., et al. (2025). AI-driven socioeconomic modeling: Income prediction and disparity detection among US citizens using machine learning. Advances in Consumer Research, 2(4).
[23] Reza, S. A., Rahman, M. K., Rahman, M. D., Sharmin, S., Mithu, M. F. H., Hasnain, K. N., ... & Kabir, R. (2025). Machine learning enabled early warning system for financial distress using real-time digital signals. arXiv preprint arXiv:2510.22287.
[24] RiskMetrics Group. (1996). RiskMetrics technical document (4th ed.). J.P. Morgan/Reuters.
[25] Shalev-Shwartz, S. (2012). Online learning and online convex optimization. Foundations and Trends in Machine Learning, 4(2), 107–194. https://doi.org/10.1561/2200000018.
[26] Shawon, R. E. R., et al. (2025). Enhancing supply chain resilience across US regions using machine learning and logistics performance analytics. International Journal of Applied Mathematics, 38(4s). https://doi.org/10.12732/ijam.v38i4s.225.
[27] Shawon, R. E. R., Buiya, M. R., Pant, S., Al Jobaer, M. A., Chowdhury, M. S. R., Kawsar, M., ... & Ali, M. (2025). Detecting illicit cross-chain fund movement: Behavioral machine learning models for bridge-based laundering patterns. International Journal of Applied Mathematics, 38(12s). https://doi.org/10.12732/ijam.v38i12s.1399.
[28] Shivogo, J. (2025). Fair and explainable credit-scoring under concept drift: Adaptive explanation frameworks for evolving populations. arXiv preprint arXiv:2511.03807.
[29] Shovon, M. S. S. (2025). Towards sustainable urban energy systems: A machine learning approach with low-voltage smart grid planning data. International Journal of Applied Mathematics, 38(8s), 1115–1155. https://doi.org/10.12732/ijam.v38i8s.631.
[30] Sizan, M. M. H., et al. (2025). Machine learning-based unsupervised ensemble approach for detecting new money laundering typologies in transaction graphs. International Journal of Applied Mathematics, 38(2s). https://doi.org/10.12732/ijam.v38i2s.88.
[31] Velankar, S., & Allu, R. (2021). Stock price forecasting using a time-series long short-term memory network. Finance, Accounting and Business Analysis, 3(1), 42–58.
[32] Zumbach, G. (2007). From JP Morgan's standard-EWMA to skewed-EWMA. MPRA Paper No. 1009022.

Appendix A

Supplementary Figures

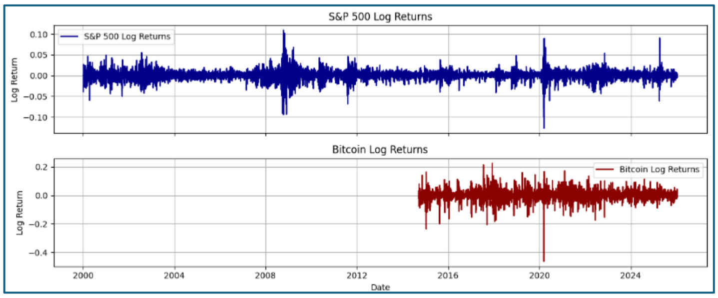

Fig. A1: Bitcoin and S&P 500 Log Returns.

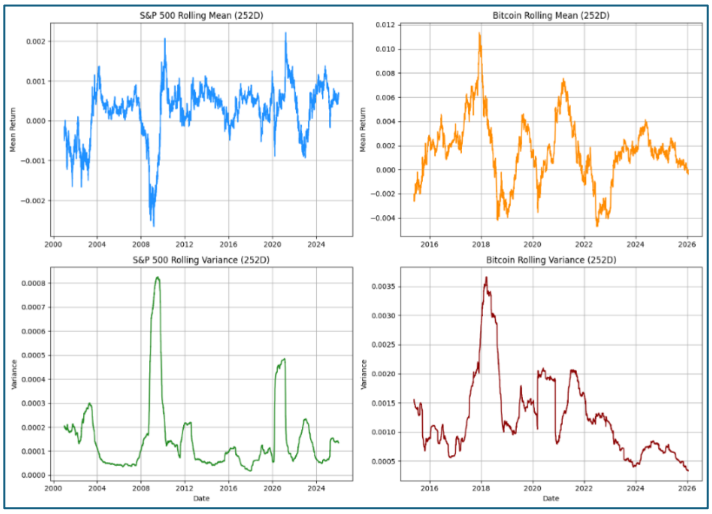

Fig. A2: Bitcoin and S&P 500 Rolling Variance and Rolling Mean Over A Trading Year (252 Days).