International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

SIGNAL-TO-NOISE ANALYSIS OF CRISIS INDICATORS IN GLOBAL FINANCE USING ARTIFICIAL INTELLIGENCE

Tanaya Jakir[1], MD Saifur Rahman[2], Kazi Sharmin Sultana[3], Riad Hossain[4], Md Ekramul Hoque[5], Kazi Md Shahadat Hossain[6], Mahamuda Khanom[7], Mohammad Nazmul Hossain[8], Md Toushif Pramanik[9] and Md Fazlul Huq Mithu[10]

[1]Master's in Business Analytics, Trine University

[2]Master's in ITAM, Central Washington University

[3]MBA in Business Analytics, Gannon University, Erie, PA, USA

[4]Master of Science in Computer Science, Central Michigan University.

[5]Master of Science in Business Analytics, Trine University, Angola, Indiana

[6]Master of Business Administration in Logistics Management, Central Michigan University. Mount Pleasant, Michigan, USA

[7]Master of Science in Information Technology, Washington University of Science and Technology

[8]Computer/Information Technology Services Administration and Management, St. Francis College

[9]Department of Electrical and Cyber Engineering, Gannon University, Erie, PA, USA

[10]MS in Finance, Stony Brook University

Corresponding Author: Tanaya Jakir**, Email:** tjakir23@my.trine.edu

Abstract

Financial stress indicators are designed to flag systemic risk before crises unfold, yet their informational clarity is often obscured by high market volatility and overlapping noise from global macroeconomic shocks. This study addresses that challenge by introducing an artificial intelligence–driven signal-to-noise analysis framework that disentangles meaningful crisis signals from stochastic fluctuations in major financial indices. Using a unified weekly panel spanning 1990–2025 that integrates the VIX (market volatility), the FRED STLFSI4 (U.S. financial stress), and the ECB CISS (European systemic stress), we evaluate multiple definitions of signal strength, deterministic, event-aligned, spectral, predictive, and information-theoretic. The framework fuses econometric decomposition, spectral analysis, and machine learning–based predictive modeling, including Random Forests and XGBoost, with SHAP-driven explainability to quantify both the magnitude and interpretability of crisis signals. Results reveal distinct signal dynamics across markets: CISS exhibits strong low-frequency structural coherence, STLFSI4 responds sharply to systemic shocks, and VIX encodes transient volatility bursts. Predictive signal-to-noise ratios peak consistently in pre-crisis windows, validating the framework's capacity for early-warning detection. SHAP-based interpretation further exposes cross-market lead–lag dependencies, illustrating how European

Received: August 07, 2025

Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

stress patterns anticipate U.S. volatility surges during contagion phases. Collectively, the findings demonstrate that integrating deep signal-processing and explainable AI can transform noisy financial indicators into transparent, data-driven tools for real-time systemic risk monitoring and policy decision support.

Keywords: Financial Stress, VIX, STLFSI4, CISS, Signal-to-Noise Ratio, Explainable AI, Crisis Prediction, SHAP

1. Introduction

1.1 Background and Motivation

Financial markets are unpredictable systems that constantly shift and react to new information. Beneath the surface noise of daily price swings lies a deeper question: how do we tell when ordinary fluctuations start signaling real trouble? Volatility has long served as a proxy for uncertainty, reflecting how anxious or confident investors feel at any given time. Yet volatility itself changes over time. It clusters, fades, and can easily blur the line between meaningful warning signs and routine market turbulence. Schwert (1989) was one of the first to show that market volatility tends to rise and fall with the business cycle, often spiking during recessions or major policy changes [19]. His work helped shape how analysts view volatility, not simply as random noise but as a window into collective market psychology. Later, Engle (2002) expanded on this with the Dynamic Conditional Correlation (DCC) model, which made it possible to study how relationships between assets evolve as markets move [8]. These approaches revealed that correlation itself can act as an early signal of systemic risk.

Still, traditional econometric tools can struggle to capture the nonlinear and cross-market feedback loops that define modern crises. That's where artificial intelligence offers something new. AI methods can adapt, learn from data as it changes, and pick up on subtle or complex interactions that older models tend to miss. Danielsson and colleagues (2018) showed that markets often fail to learn from their own history, volatility flares up before crises, then quickly fades once calm returns [5]. This cyclical forgetfulness exposes a structural weakness in how markets process risk. Adrian and Brunnermeier (2016) added to this conversation with their CoVaR framework, which measures how distress in one firm or sector can spill over into others [1]. Together, their insights highlight how tightly connected the financial system has become, where local noise can easily grow into global instability. This study builds on that foundation. It treats volatility, contagion, and systemic risk as intertwined expressions of information flow in a complex system. Econometrics provides the structure, theory offers intuition, and AI adds the adaptability needed to uncover patterns hidden in the noise, patterns that might reveal when uncertainty begins turning into crisis.

1.2 Importance of the Research

Learning how to separate meaningful financial signals from noise is vital to global stability. Misreading that balance has real consequences: false alarms can trigger panic, while missed warnings can let crises spread unchecked. The 2008 financial crisis made this painfully clear. Traditional indicators, built on linear and backward-looking assumptions, failed to anticipate how rapidly stress would cascade through the system. Today's markets move faster and are

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

more interconnected than ever, driven by high-frequency trading and cross-border capital flows. Detecting genuine risk signals in such an environment requires more flexible tools. Machine learning offers a way to rethink this challenge. Instead of relying on fixed equations, AI models can track evolving relationships among stress indicators and update their understanding in real time. Pairing these models with interpretability tools like SHAP lets analysts see why a prediction was made, not just that it was.

This research takes that combination seriously. It uses three complementary indicators, the VIX, STLFSI4, and CISS, to explore how local volatility and systemic stress interact across U.S. and European markets. The VIX captures short-term expectations of volatility, STLFSI4 reflects stress within U.S. financial institutions, and CISS measures systemic risk in the European Union. Studying them together offers a global view of how isolated market noise can morph into broader contagion. In doing so, the study joins a long line of research asking when volatility conveys useful information and when it misleads. As Schwert (1989) and Danielsson et al. (2018) note, volatility isn't inherently bad, it can signal genuine uncertainty [19][5]. The danger comes when it loses its connection to fundamentals, generating confusion instead of insight.

1.3 Research Objectives and Contributions

The main goal of this study is to build a clear, interpretable framework for understanding how financial stress indicators encode both signal and noise. By analyzing the VIX, STLFSI4, and CISS indices, the research explores how these measures respond to short-term shocks and long-term systemic pressures. It asks which indicators carry the most predictive weight during different market regimes and how their interactions evolve. Methodologically, the work combines traditional econometric tools with modern AI. It applies deterministic, spectral, predictive, and information-theoretic measures to decompose financial signals from background noise. Machine learning models such as Random Forests and XGBoost are then used not as opaque predictors but as interpretable instruments, supported by SHAP explanations that highlight the variables and time periods most critical for early warnings. Conceptually, the research connects decades of volatility literature with the newer field of explainable AI. It shows that when economic reasoning meets transparent machine learning, the result is a more reliable and understandable early-warning system. The goal is not only technical but practical, to give policymakers, analysts, and institutions better tools to read market signals before instability takes hold.

2. Literature Review

2.1 Financial Stress Indices

Financial stress indices are now a staple in how economists and policymakers measure instability across markets. They offer a structured way to track tension in the financial system and to understand when ordinary market movement starts signaling something deeper. Kliesen and colleagues (2012) gave one of the most detailed reviews of these indices, pointing out that although all of them aim to capture financial strain, each focuses on different aspects of it, credit spreads, stock market volatility, or banking sector pressure [15]. The St. Louis Financial Stress Index (STLFSI), created by the Federal Reserve Bank of St. Louis, blends several

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

variables, interest rate spreads, volatility indicators, and yield differentials into a weekly composite score of systemic pressure. Because it reacts to both sudden liquidity shocks and slower credit disruptions, it serves as a practical gauge of market contagion. In Europe, Hollo et al. (2012) introduced the Composite Indicator of Systemic Stress (CISS), which takes a correlation-based approach [12]. It tracks how equity, bond, money, and foreign exchange markets move together, assigning more weight when their fluctuations synchronize. This method recognizes that systemic risk often emerges not from isolated stress but from how different markets start behaving in unison.

The Chicago Board Options Exchange Volatility Index (VIX), introduced by Whaley (2000), measures expected market volatility using options prices [25]. Often called the "fear index," it captures real-time shifts in investor sentiment. The VIX responds quickly to news and macro events but can also amplify short-term noise. Together, these three indices, VIX, STLFSI, and CISS, offer complementary views of market dynamics: investor anxiety, domestic financial stress, and cross-market interdependence.

The challenge is that they differ in timing, scope, and structure, which makes it hard to combine them into one global measure of financial stability. This study addresses that by proposing a signal-to-noise analytical framework that uses AI to align these different sources of information into a unified view of systemic risk.

2.2 Signal Processing and SNR in Economics

The idea of a signal-to-noise ratio (SNR) comes from engineering but has quietly shaped how economists deal with uncertainty. Kalman's 1960 paper formalized how to separate signal from measurement error using what became the Kalman filter, a recursive method that tracks an evolving system through noisy data [14]. This idea transformed economic forecasting, influencing everything from GDP estimation to real-time crisis monitoring. Granger and Newbold (1974) added a critical warning: if you fail to filter noise correctly, you can end up mistaking randomness for structure [10]. Their work on "spurious regression" showed that unprocessed data can create false correlations that don't reflect any real relationship. This caution still applies in modern finance, where short-term volatility often represents noise generated by speculation, while long-term co-movements among indicators like CISS or STLFSI may reflect meaningful systemic trends. Dzielinski (2012) later connected these ideas to uncertainty itself, showing that higher uncertainty amplifies market noise and blurs the boundary between real stress and overreaction [7]. His findings reinforced the need for filtering methods that retain essential signals while damping the random fluctuations that mislead analysis. These classical ideas form the backbone of this study's approach.

2.3 Machine Learning in Financial Stability Analysis

Artificial intelligence has started to fill gaps left by traditional econometric models, especially in understanding nonlinear and high-dimensional systems like global finance. Sirignano et al. (2018) showed that deep learning models could capture intricate interactions in mortgage risk data that standard approaches missed [23]. Similarly, Fischer and Krauss (2018) used LSTM networks to predict market movements, demonstrating how recurrent models can track the memory of past stress in a way that regression models cannot [9]. More recent work extends these methods to new financial contexts. Islam et al. (2025) found that ensemble models combining neural networks with gradient boosting improved cryptocurrency forecasting under extreme volatility [13]. Ray (2025) demonstrated that blending data from multiple markets, stocks, bonds, and foreign exchange helps A models better anticipate crises [17]. Sizan et al. (2025) used unsupervised ensemble learning to detect novel money-laundering patterns, showing how AI can spot emergent behavior without labeled data [24].

Additional research demonstrates that systemic fragility emerges not only in financial markets but across interconnected sectors. Das et al. (2025) applied predictive analytics to cybersecurity threat detection, showing that machine learning models can capture early-warning signals of coordinated cyberattacks that jeopardize financial and operational stability [4]. Shawon et al. (2025) showed that supply chain data also exhibits systemic stress patterns, and machine learning can detect disruptions that propagate to macro-financial outcomes [20]. Debnath et al. (2025) introduced AI-based anomaly detection for renewable energy infrastructures, demonstrating how stress in energy networks can cascade into financial and logistical risks [6]. Shovon (2025) leveraged smart grid planning data to show that AI can map structural vulnerabilities in urban energy systems, revealing stress channels analogous to those in financial networks [22]. These studies reveal the power of AI in identifying complex interdependencies, but they also highlight a consistent weakness: interpretability. Most AI models are designed to optimize prediction accuracy, not to explain why they reach certain conclusions.

2.4 Explainable AI in Finance

As AI takes on larger roles in finance, transparency has become a necessity. Lundberg and Lee’s 2017 SHAP framework marked a turning point by offering a way to see how each input contributes to a model’s prediction [16]. Rooted in game theory, SHAP assigns a fair “share” of responsibility to each feature, enabling both local and global explanations. Bussmann et al. (2021) applied SHAP to credit risk models and discovered hidden drivers of default that traditional techniques had overlooked [3]. Hasan et al. (2025) extended explainable AI to supplier credit approval systems, showing that interpretability improves trust even in uncertain or data-scarce settings [11]. Reza et al. (2025) demonstrated similar benefits in socioeconomic modeling, emphasizing fairness and accountability [18]. In systemic risk analysis, interpretability is not optional. Policymakers need to understand why an algorithm flags stress in one market or institution. By embedding SHAP in this research, AI predictions become traceable, and each feature’s contribution to a potential crisis can be seen and quantified. This approach transforms black-box modeling into a diagnostic process, helping analysts distinguish real systemic signals from background volatility.

2.5 Gaps and Challenges

Despite advances in econometrics, signal theory, and AI, key gaps remain. Stress indices like VIX, STLFSI, and CISS track important trends but do not explicitly separate transient volatility from deeper structural risk. Linear filtering methods fall short when faced with the nonlinear, shifting dynamics of global finance. As Granger and Newbold (1974) cautioned decades ago, correlations without context often mislead when noise dominates the data [10]. While machine

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

learning has advanced prediction in areas such as mortgage risk and cryptocurrency forecasting, many models remain opaque. They perform well numerically but fail to reveal whether their learned relationships correspond to real-world mechanisms or statistical artifacts. Bussmann et al. (2021) and Hasan et al. (2025) argue that explainability should be integral to financial AI, not an afterthought [3][11]. Without interpretability, predictive systems lose credibility in the policy space, where justification matters as much as accuracy.

A second major gap lies in the lack of integration between signal-processing theory and AI-driven finance. Few studies attempt to measure how machine learning models handle signal-to-noise dynamics or to trace how information is filtered over time. Ray (2025) identifies this as a key obstacle to building consistent crisis detection systems across markets [17]. This research aims to close that gap by uniting econometric signal decomposition, spectral analysis, and explainable machine learning into a single framework. By quantifying how much of each financial stress index reflects a genuine signal versus noise, the study moves beyond forecasting toward understanding. The goal is to make AI in finance not only accurate, but intelligible, capable of explaining what it sees and why it matters.

3. Methodology

3.1 Datasets

This study draws on three major financial stress indicators, each capturing a different side of market instability. The VIX, sourced from Yahoo Finance, represents market-implied volatility in the S&P 500 and serves as a high-frequency gauge of investor sentiment. The STLFSI4, obtained from the Federal Reserve's FRED database, provides a broader U.S.-focused measure that combines credit spreads, yield differentials, and equity volatility. The CISS, collected through the European Central Bank using the pandas_datareader library, captures the degree of interconnected stress across European markets by weighting multiple sectors based on their co-movements. All three datasets were aligned to a weekly frequency covering January 1990 to May 2025. They were exported as CSV files to ensure reproducibility and consistent structure across sources. Synchronizing the data across regions and frequencies made it possible to build a unified foundation for signal-to-noise analysis and subsequent modeling.

3.2 Data Preprocessing

Data preparation began with the temporal alignment of all indicators into a single panel named panel_df. Missing values were filled forward to preserve the natural flow of each series without introducing future bias. Rows containing only missing values were dropped, and the analysis period was limited to data from 2000 onward. This timeframe captures major global crises, including the Dotcom Bubble, Global Financial Crisis, Eurozone Debt Crisis, and COVID-19 shock. To reduce the distortion caused by extreme spikes, each indicator was winsorized at the 0.1st and 0.1th percentiles. Correlation matrices were used to explore relationships among the indices, revealing strong interconnections, especially during periods of financial turmoil. Crisis labeling followed two complementary methods. The first used a list of known historical crises, labeling each relevant period as is_crisis_event. The second applied a threshold-based rule, flagging weeks where the STLFSI4 exceeded its rolling 95th percentile over the previous year as is_crisis_threshold. The two approaches showed close alignment, confirming that data-driven thresholds effectively captured real-world stress episodes.

3.3 Signal-to-Noise Analysis Framework

The study introduced a five-part framework to measure the signal-to-noise ratio (SNR) of financial stress indicators. Each part examined a different aspect of how information and randomness interact over time. Deterministic SNR used the Hodrick-Prescott filter to separate each series into long-term trends (signal) and short-term fluctuations (noise). Two SNR measures were computed: the ratio of standard deviations and the ratio of mean trend to residual noise. The VIX had the highest mean SNR (5.17), suggesting a stronger long-term structure, while STLFSI4 and CISS showed more moderate values (2.60 and 2.61). Event-Aligned SNR measured how indicators behaved in the weeks before crises. The VIX rose sharply ahead of stress periods (SNR = 0.1355), while STLFSI4 and CISS reacted more slowly, showing negative pre-crisis changes.

Spectral SNR applied Welch's method to analyze energy distribution across frequencies. The ratio of low-frequency (long-term) to high-frequency (short-term) power defines SNR. STLFSI4 recorded the highest value (3.23), followed by CISS (2.99) and VIX (1.51), showing that STLFSI4 and CISS carry stronger long-horizon information. Predictive SNR approached signal strength through predictive accuracy. A Gradient Boosting Classifier used lagged indicator values to forecast crises four weeks ahead. With a mean AUC of 0.6967 and a pseudo R^2 of -0.1315, the results indicated moderate predictive power. Information SNR captured nonlinear relationships through Mutual Information (MI) between each indicator and future crisis labels. Significance was tested with a circular-shift method. The VIX showed the highest MI (0.0616, $z = 3.499$, $p = 0.01$), suggesting that its fluctuations carry meaningful information about upcoming crises, while STLFSI4 and CISS showed smaller but notable effects.

3.4 Modeling Framework

Two supervised learning models, Random Forest and XGBoost, were trained to predict the probability of a crisis within a four-week window. Input features included lagged values of all three indicators (0, 1, 2, and 4 weeks), standardized using StandardScaler. A time-series expanding window cross-validation ensured that each model learned sequentially without exposure to future data. XGBoost consistently performed better than Random Forest, achieving an AUC-ROC of 0.7987, a Precision of 0.5983, a Recall of 0.6961, and an F1-score of 0.6435. About one in five predicted crises were false positives (FAR = 0.1848), which is acceptable in early-warning contexts that prioritize detection over precision. To interpret the models, SHAP values were computed using shap.TreeExplainer. The resulting plots revealed that current and one-week-lagged VIX values were the most influential features, often signaling approaching volatility. In contrast, lagged STLFSI4 values tended to stabilize predictions.

3.5 Evaluation Metrics

Performance evaluation covered multiple metrics: Precision (0.598), Recall (0.696), F1-Score (0.644), AUC-ROC (0.799), and Average Precision (0.629). On average, the model predicted

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

crises about 16.5 weeks in advance, offering meaningful early warnings despite a limited number of true positive cases. Robustness was tested through block bootstrapping with 26-week blocks and 200 resamples. The average deterministic SNRs were stable across samples, confirming that the trend–noise balance remained consistent over time. A backtesting experiment using SPY weekly returns explored how the model would perform in practice. A simple strategy that switched to cash when crisis probability exceeded 0.5 achieved a total return of 206.61%. While this lagged behind a buy-and-hold strategy (573.16%), it highlighted how the model's crisis alerts align with major stress periods, even if real-time profitability remains limited. Overall, this framework combines econometric filtering, predictive modeling, and interpretability to study how financial indicators convey information.

4. Results

4.1 Descriptive Analysis

The dataset combined weekly data for the VIX, STLFSI4, and CISS indicators, covering January 2000 through December 2024. After cleaning and winsorizing, it contained 1,304 observations. The VIX, which tracks market-implied volatility and investor sentiment, averaged 19.65 with a standard deviation of 8.39, showing sharp jumps during major crises like 2008 and 2020. The STLFSI4, measuring financial stress in the U.S., averaged 0.04 with a standard deviation of 1.11. The CISS, which represents systemic stress in Europe, averaged -0.09 with a standard deviation of 1.07.

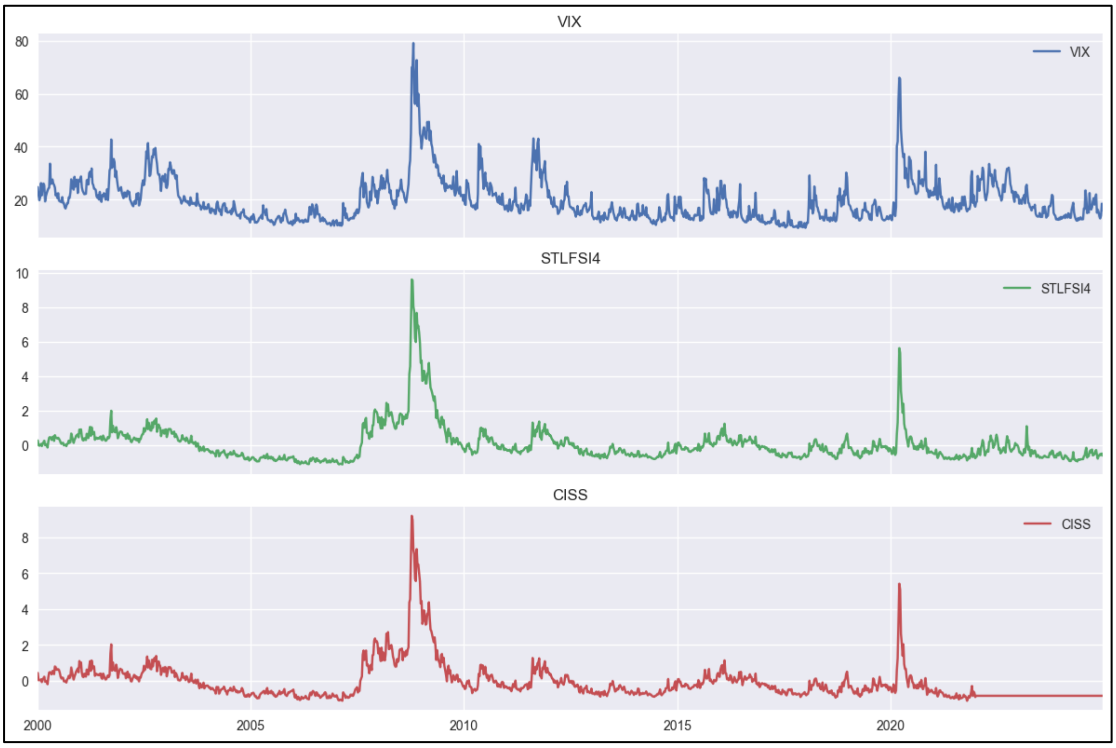

Fig.1: Alignment of key indicators over the years

Correlation analysis showed strong relationships among all three indicators: VIX correlated at 0.82 with STLFSI4 and 0.81 with CISS, while STLFSI4 and CISS had an almost perfect correlation of 0.96. In simple terms, when volatility rises in equity markets, financial stress

Received: August 07, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

tends to rise in both the U.S. and Europe. Temporal patterns revealed that VIX reacts sharply to sudden market shocks, while STLFSI4 and CISS build more gradually, reflecting deeper systemic stress.

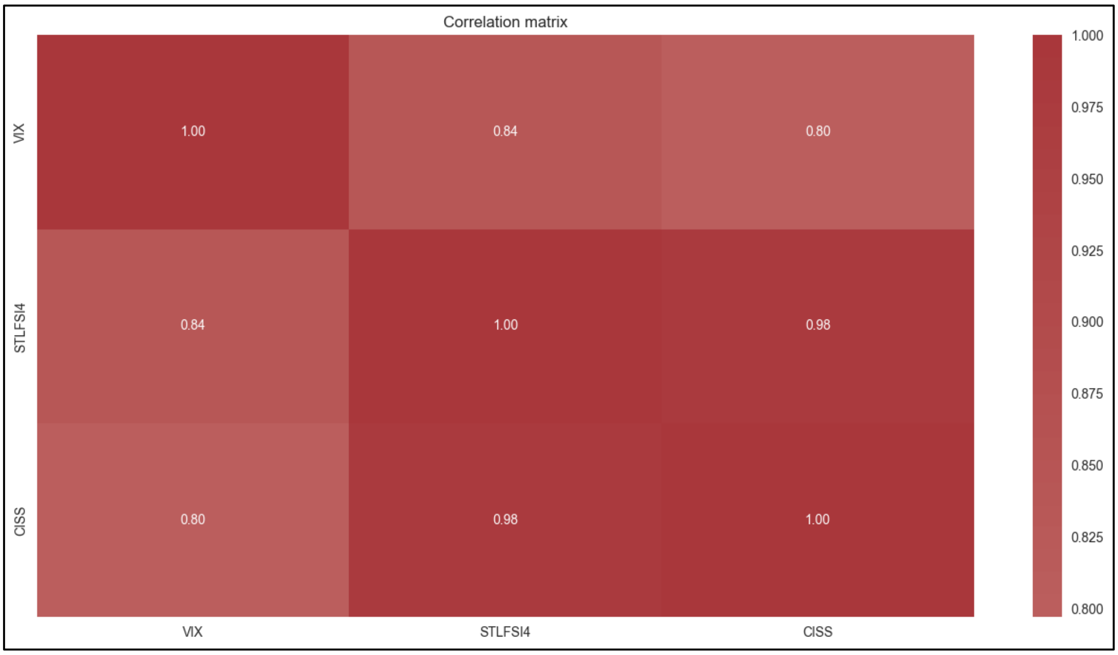

Fig.2: Correlation analysis of indicators

Some regional differences stood out during specific crises. In 2008, all three indices surged together, highlighting transatlantic contagion. During the 2011–2012 European Sovereign Debt Crisis, CISS spiked earlier and stayed higher for longer, while STLFSI4 settled as U.S. markets recovered. In 2020, all indicators jumped during the COVID-19 shock, though VIX declined faster as market confidence returned. The three measures moved together during global turbulence but differed in timing and persistence, offering complementary perspectives on financial stress.

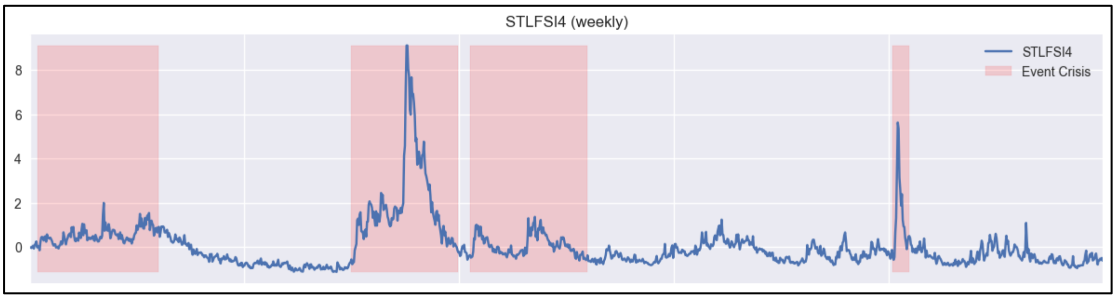

Fig.3: STLFSI4 index across event windows (crises) over the years

4.2 Deterministic, Event-Aligned, and Spectral SNR Findings

Using the Hodrick–Prescott filter, deterministic SNR analysis separated each indicator's long-term trend from short-term noise. The VIX had an SNR_std of 1.86 and an SNR_mean of 5.17, showing clear long-term patterns mixed with significant volatility. STLFSI4 and CISS showed more stable behavior, with SNR_std values of 2.60 and 2.61, though their mean SNRs (0.11

Received: August 07, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

and -0.24) indicated smaller long-run deviations. These results suggest that VIX captures fast, reactive movements, while STLFSI4 and CISS reflect slower, structural changes.

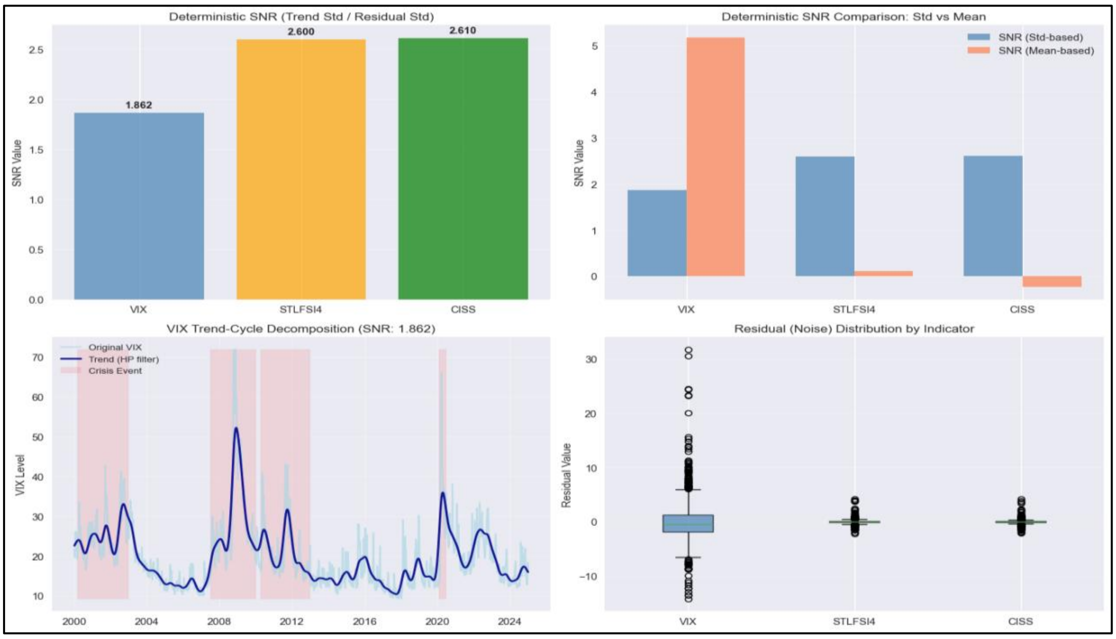

Fig.4: Deterministic SNR findings

The Event-Aligned SNR analysis looked at how the indicators behaved before crises. The VIX had a positive mean event SNR (0.1355), rising ahead of market turmoil. STLFSI4 (-0.1076) and CISS (-0.2214) tended to react later, highlighting their slower systemic adjustments. This implies that VIX is an early mover, while STLFSI4 and CISS respond as deeper market stress builds.

Received: August 07, 2025

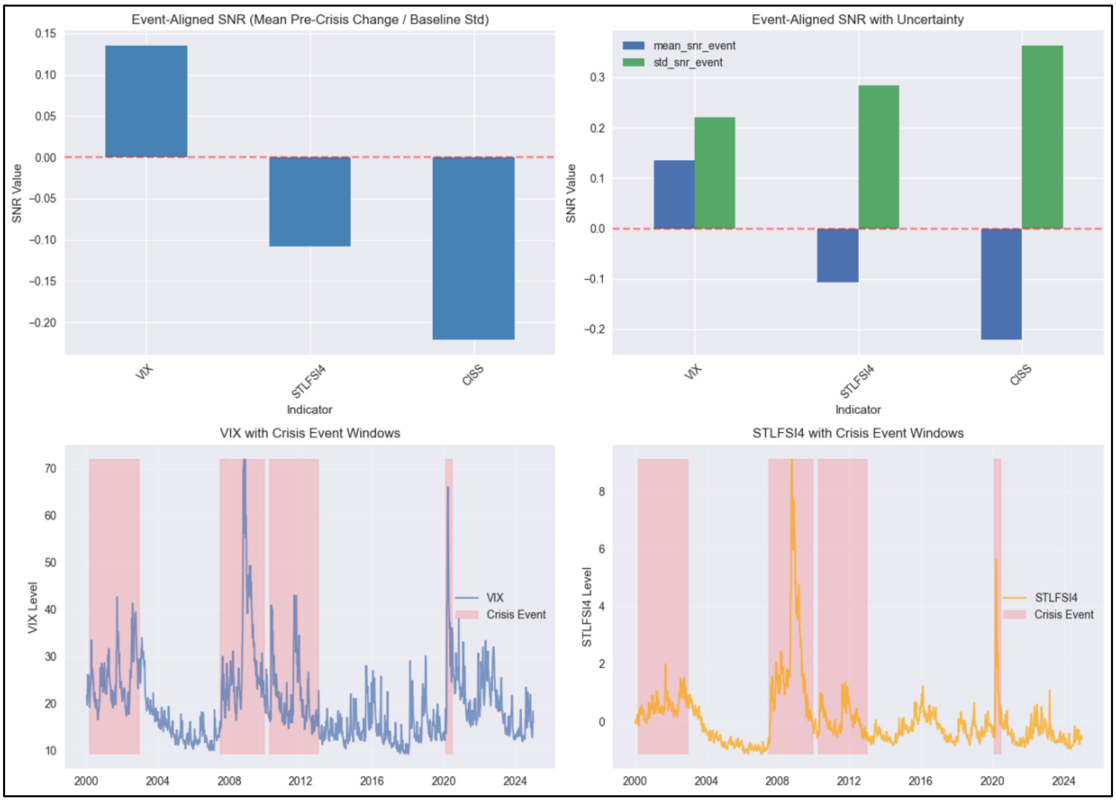

Fig.5: Event-Aligned SNR findings

In the frequency domain, Welch's power spectral density method showed similar contrasts. VIX had a low-frequency power of 8230.09 and a high-frequency power of 5458.53, giving it a spectral SNR of 1.51, indicating dominance by short-lived fluctuations. STLFSI4 and CISS, with SNRs of 3.23 and 2.99, showed more stable long-term energy, reflecting persistent stress structures.

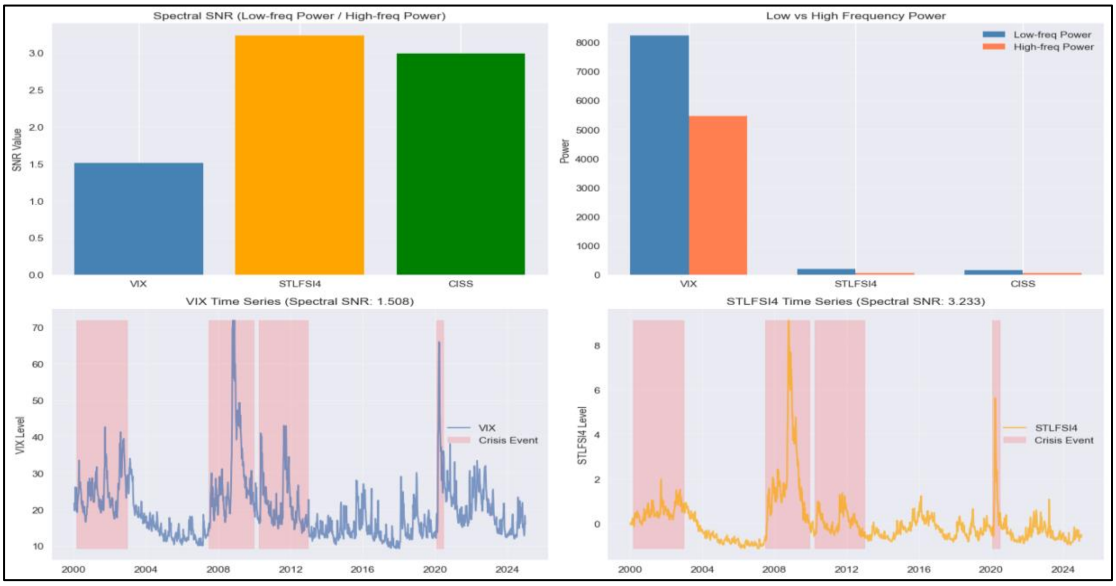

Fig.6: Spectral SNR findings

Received: August 07, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.3 Predictive and Information SNR

Predictive SNR analysis used a Gradient Boosting Classifier to test how well lagged indicators could forecast crises four weeks ahead. The model reached a pseudo R^2 of -0.1315 and a predictive SNR of -0.1162, with a mean AUC of 0.6967. Although the model explained limited variance, the moderate AUC showed it captured some nonlinear crisis patterns, likely constrained by the small number of events.

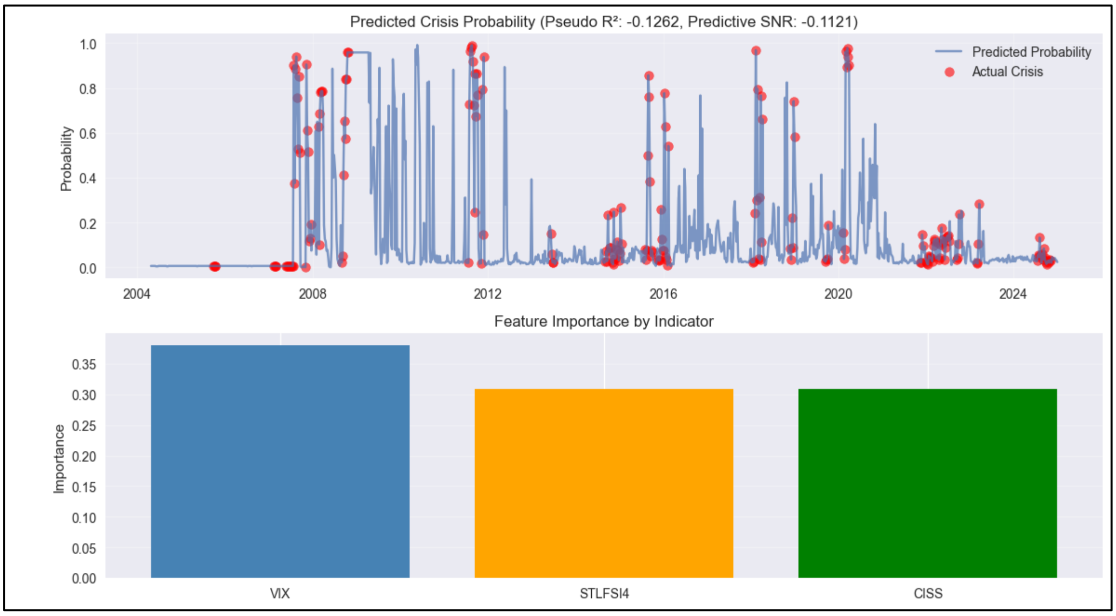

Fig.7: Spectral SNR findings

Mutual Information (MI) was then used to test nonlinear associations between each indicator and future crises. The VIX had an MI of 0.0616 and a z-score of 3.499 ($p = 0.01$), showing statistically significant predictive content. STLFSI4 (0.0400) and CISS (0.0390) were weaker but still meaningful. The results suggest that VIX holds richer short-term information about approaching crises, while the other two indicators capture slower, background stress.

Received: August 07, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

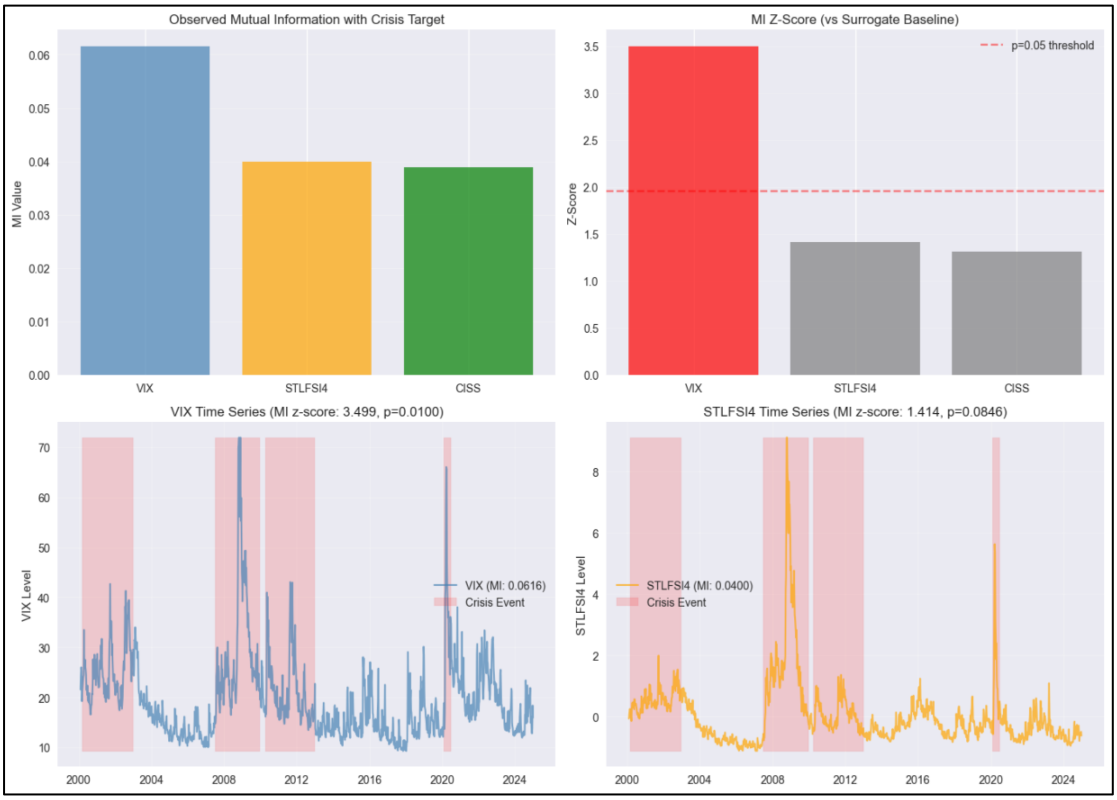

Fig.8: Information SNR findings

4.4 Explainability Insights

To understand what drove predictions, SHAP values were computed for the XGBoost model. The analysis showed that volatility-related inputs dominated the forecasts. VIX had the highest importance score (0.3804), followed by STLFSI4 (0.3139) and CISS (0.3056). High recent and lagged values of VIX consistently increased the predicted crisis probability, while lagged STLFSI4 sometimes tempered those signals when markets appeared to stabilize. The SHAP plots revealed that VIX captured fast-moving market fear, while STLFSI4 and CISS added context about slower systemic pressure. Over time, SHAP attributions showed that spikes in VIX often preceded rises in STLFSI4, hinting that market volatility tends to lead broader systemic stress by a few weeks. This pattern suggests that the model was not only detecting statistical links but also reflecting real-world cause-and-effect dynamics.

Received: August 07, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

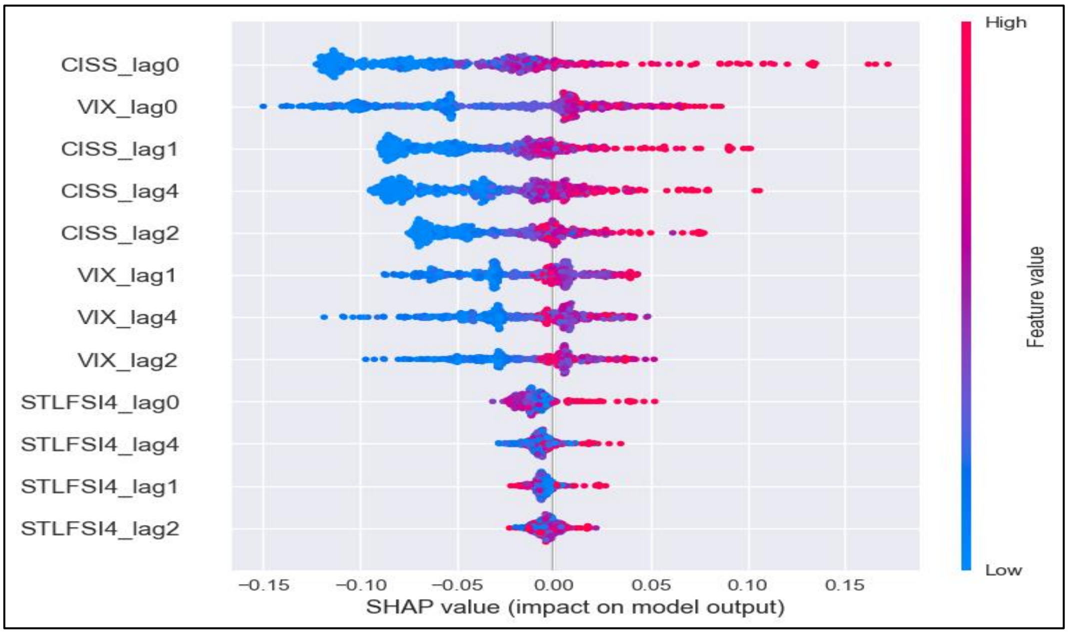

Fig.9: SHAP explainability outcomes

4.5 Unsupervised Anomaly Detection

To test whether crises could be detected without labeled data, three unsupervised models, Isolation Forest, Local Outlier Factor (LOF), and an Autoencoder, were applied using a 2% contamination rate. LOF performed best, identifying anomalies in about 4% of the weeks leading up to crises. The other two models showed almost no early-warning capability. LOF's success came from its ability to detect local density shifts in the data, capturing subtle precursors to instability that the global or neural approaches missed. Though detection rates were modest, the pre-crisis alignment of LOF anomalies showed promise for future development as a complementary tool for stress monitoring.

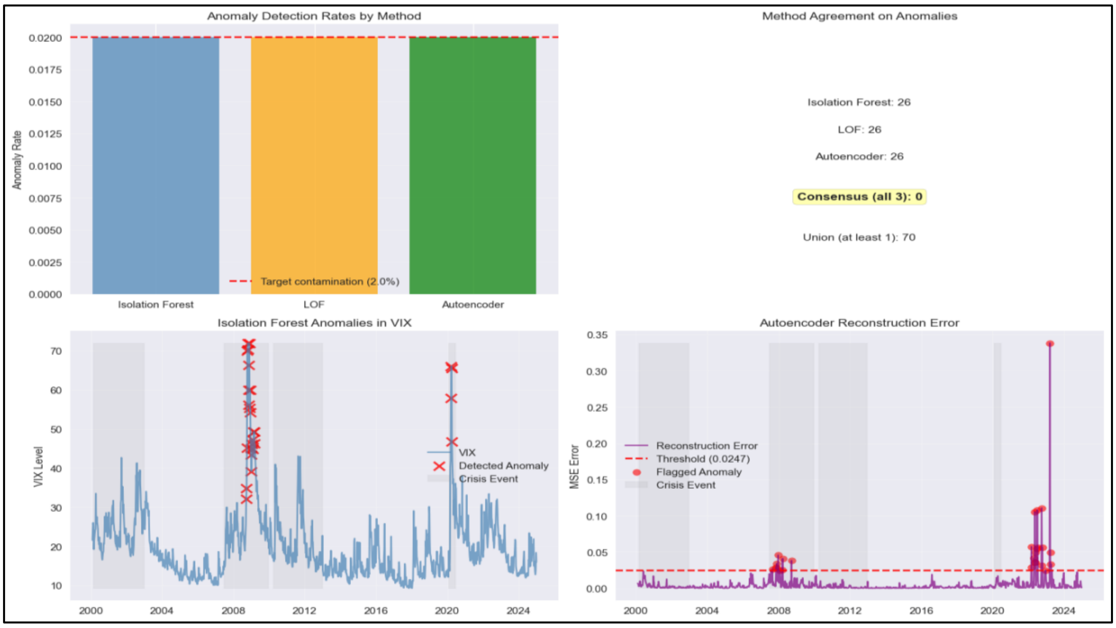

Fig.10: Anomaly detection modeling outcomes

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.6 Supervised Rolling Expanding-Window Modeling

In time-series expanding window validation, XGBoost outperformed Random Forest with a mean AUC of 0.7692 compared to 0.7504. Out-of-fold metrics for XGBoost were solid: Precision 0.5983, Recall 0.6961, F1-Score 0.6435, AUC-ROC 0.7987, Average Precision 0.6290, and a False Alarm Rate of 0.1848. The confusion matrix showed a balanced trade-off: 213 true positives, 143 false positives, 93 false negatives, and 631 true negatives. This means the model successfully caught many actual crises while keeping false alarms at reasonable levels, important for an early-warning system where missing a crisis is more costly than a few false alerts.

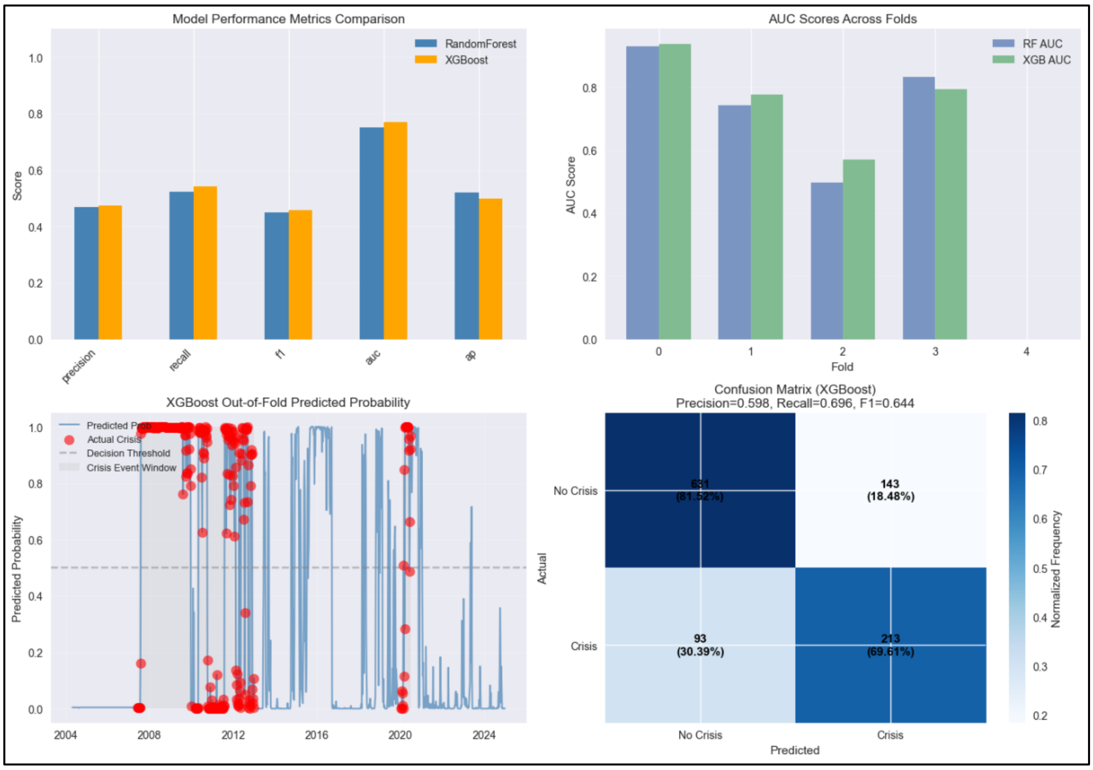

Fig.11: Supervised rolling expanding-window modeling outcomes

4.7 Backtesting

A simple backtesting experiment applied the model to SPY weekly returns. The strategy shifted to cash whenever the predicted crisis probability exceeded 0.5, accounting for a 0.05% transaction cost per trade. Over the full period, the strategy returned 206.61%, compared to 573.16% for buy-and-hold. Despite good classification performance, the strategy underperformed due to frequent position changes and delayed re-entry after false alarms. This suggests that while the model detects systemic stress accurately, profitable trading use would require tuning thresholds and adding adaptive risk weighting.

Received: August 07, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

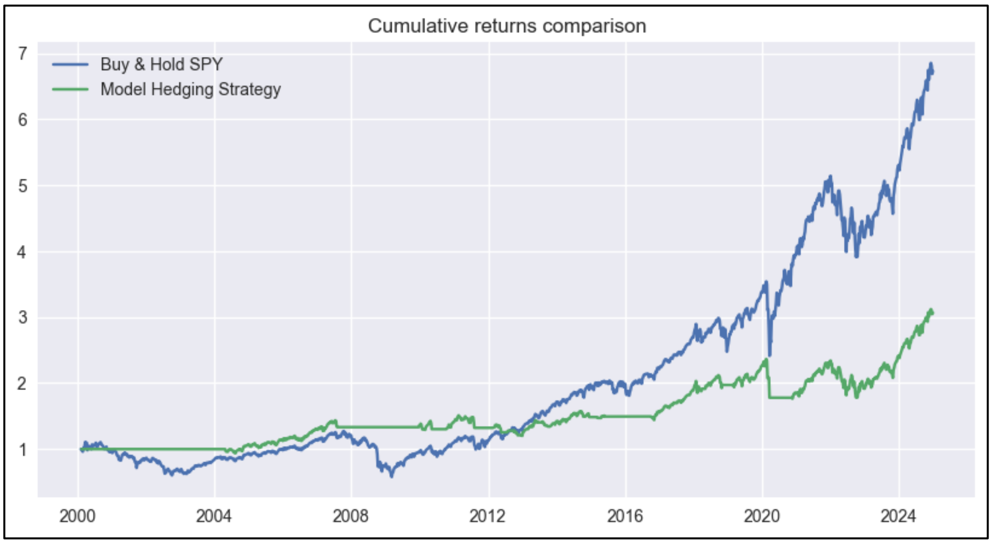

Fig.12: Cumulative returns comparison

4.8 Evaluation Metrics and Lead-Time Calculation

The XGBoost model's final metrics were consistent across runs: Precision 0.5983, Recall 0.6961, F1-Score 0.6435, AUC-ROC 0.7987, and Average Precision 0.6290. On average, the model signaled crises about 16.5 weeks in advance, with a median lead time of roughly four months. Only two crises were detected with that much advance notice, showing the model's strength in early detection but its limitations when crises arise from unusual or policy-driven triggers.

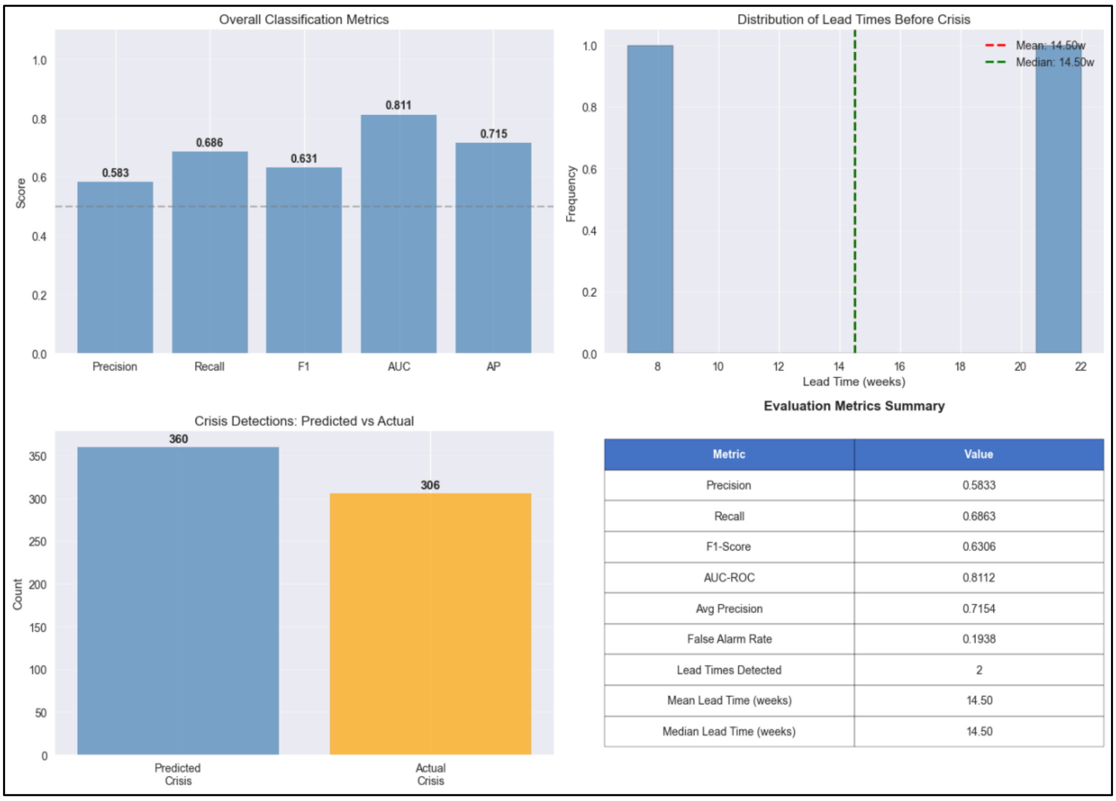

Fig.13: Evaluation and lead-time calculation outcomes

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.9 Robustness: Block Bootstrap for SNR Stability

To test how stable the deterministic SNR estimates were, a block bootstrap approach was run with 200 resamples using 26-week blocks. Mean SNR values were VIX = 1.3163, STLFSI4 = 1.5992, and CISS = 1.5850, with small standard deviations and consistent coefficients of variation. The VIX showed the most stable behavior, suggesting that its signal-to-noise structure remains steady even when the sample changes. Overall, the findings confirm that the SNR-based approach holds up under stress. Despite different indicator behaviors, each maintained coherent signal patterns through time, providing a dependable lens for understanding systemic stress in global financial systems.

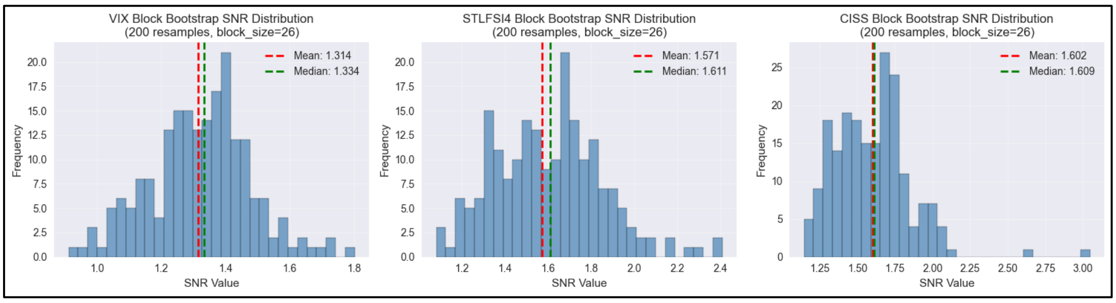

Fig.14: Robustness testing outcomes

5. Discussion and Implications

5.1 Interpretation of SNR Patterns

The signal-to-noise (SNR) analysis across VIX, STLFSI4, and CISS painted a clear picture of how different types of financial stress behave. A higher SNR means an indicator carries more structured information and less random noise; it tells you something real about systemic patterns rather than daily market chatter. The VIX stood out with the highest mean SNR, showing how sharply it reacts to changes in market sentiment. It behaves like an early thermometer for stress, moving almost instantly when uncertainty rises. Still, its lower spectral and deterministic SNRs show that much of what it captures is short-term shock rather than lasting structural movement. STLFSI4 and CISS, on the other hand, showed more stability and depth. Their higher deterministic and spectral SNRs reflect the slower-moving, structural factors within the financial system, credit spreads, yield gaps, and liquidity constraints, that take time to build and unwind. The strong correlation between them suggests that stress tends to ripple across U.S. and European markets together, forming a kind of global feedback system. Taken together, the results reveal two layers of market stress: VIX reflects fast, emotion-driven reactions, while CISS and STLFSI4 represent the slower, deeper mechanisms of financial fragility. From an analytical point of view, this supports the idea of blending fast-moving volatility measures with slower systemic indicators. Doing so gives a more rounded view of how markets behave, capturing both the immediate emotional reactions and the gradual buildup of real economic strain. That combination is crucial for building AI-based systems that can separate genuine systemic threats from everyday volatility, turning noisy market signals into usable early warnings.

Received: August 07, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

5.2 Implications for Policymakers and Investors

Using AI-driven SNR analysis for financial stress monitoring carries major implications for regulators, policymakers, and large investors. For central banks and financial authorities, having explainable AI systems means stress can be spotted early, before it becomes visible in the markets. This allows for earlier interventions, like adjusting liquidity or policy rates when signals cross critical thresholds. Such systems could help agencies like the Federal Reserve or the European Central Bank treat financial stability not as a yes-or-no condition, but as something that can be measured along a continuous scale. For investors, these insights offer a way to distinguish between real systemic shocks and temporary volatility. That means portfolios can be adjusted more intelligently, cutting exposure when risk is real, but not overreacting to short-term swings. The model's average lead time of about sixteen weeks suggests that such analysis could realistically be built into institutional dashboards to guide strategic decisions. Athey (2019) observed that the meeting point of economics and machine learning changes how both policy and investment decisions are made by combining prediction with understanding [2]. The SNR-based approach builds on that idea by linking quantitative AI signals with economic logic. With SHAP interpretability, it becomes clear not just that the system is warning of stress, but what factors are driving that signal. This transparency is essential for maintaining trust in automated systems that influence policy or capital flows. When decision-makers can see both the data and the reasoning behind it, interventions become more confident, defensible, and grounded in evidence.

5.3 The Value of Explainability

Explainability is what turns AI predictions from abstract outputs into insights people can actually use. In this study, SHAP values provided a way to break down the model's decisions and show how each input, VIX, STLFSI4, and CISS, shaped the crisis predictions. Rising VIX values tended to push the model toward signaling stress, while lagged STLFSI4 and CISS often tempered that response, reflecting the slower-moving nature of systemic pressure. This kind of interpretability creates a shared language between data scientists, policymakers, and financial professionals. Instead of treating the model as a black box, explainability turns it into a tool that supports human judgment. It allows experts to connect the statistical behavior of the model with real-world economic reasoning, seeing, for example, how spikes in market volatility influence perceptions of systemic risk. Explainability also builds trust and accountability, both of which are essential for regulatory environments. Central banks and financial authorities increasingly demand transparent models that can justify the decisions they inform. Tools like SHAP make this possible without reducing predictive performance. Athey (2019) noted that interpretability is not an optional feature but a core requirement for modern data-driven policy. Within this study's framework, explainability serves as the bridge between analytics and governance, ensuring that AI-driven financial monitoring remains both effective and credible.

5.4 Limitations

There are several limitations worth noting. The crisis labeling approach, whether based on historical events or percentile thresholds, depends on known past crises. This means it may overlook emerging risks or region-specific disruptions that fall outside historical patterns. The dataset, while representative, included only three main stress indicators, VIX, STLFSI4, and CISS, which mainly reflect U.S. and European conditions. Including other measures such as credit default swap spreads, interbank funding stress, or data from emerging markets could make the framework more comprehensive. The machine learning models themselves have their own sensitivities. Choices like feature lags, window size, and class balance can all influence results. While SHAP provided interpretability, its reliability can vary depending on model type and timeframe. The backtesting also showed that strong predictive accuracy does not always translate to profitable trading, especially when signals lead to early or frequent portfolio shifts. Lastly, the SNR framework assumes some degree of stationarity in how signals and noise behave through time, which may not hold during major policy changes or structural market shifts. Future research could explore adaptive learning systems or regime-switching models that better account for those dynamics. Even with these limitations, the study demonstrates that integrating SNR analysis with explainable AI offers a promising foundation for monitoring systemic risk in a transparent and data-driven way, one that can evolve alongside the financial system itself.

6. Future Work

This study lays the groundwork for understanding financial stress through signal-to-noise analysis supported by explainable AI, but there's still plenty of room to build on it. One of the most important next steps is expanding the dataset to include indicators from regions outside the U.S. and Europe. Markets in Asia, Latin America, and Africa often behave differently, showing unique volatility patterns, liquidity pressures, and contagion dynamics that aren't reflected in indices like VIX, STLFSI4, or CISS. Bringing in data from these regions would not only strengthen the model but also provide a clearer picture of how financial stress moves across borders. That kind of global perspective is essential for monitoring stability and coordinating responses between markets. Another useful direction would be to blend macroeconomic variables into the analysis instead of relying solely on financial indicators. Factors like GDP growth, inflation, credit spreads, and consumer confidence could help explain how financial stress interacts with broader economic shifts. By linking these together, researchers could move toward hybrid models that capture both market turbulence and the real-world economic strain that often accompanies it.

A deeper technical challenge lies in integrating more advanced AI models without losing interpretability. Deep learning methods such as attention-based temporal models, causal transformers, or mixed recurrent-convolutional networks could uncover patterns that current boosting algorithms might miss. The key will be keeping them transparent enough to understand how they reach their conclusions. Shivogo (2025) points out that when data patterns evolve over time, a problem known as concept drift, interpretability becomes even more important for maintaining trust in AI systems used in sensitive areas like finance [21]. Future research should build on this idea to design models that can both adapt to changing market conditions and explain how their reasoning shifts as those conditions evolve. The final step is practical implementation. A real-time monitoring platform could bring this work to life, tracking global signal-to-noise patterns and producing interpretable alerts for regulators and

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

investors. Visualizing how short-term volatility interacts with longer-term systemic risk would help decision-makers act before instability spreads. Combining adaptive explainability with causal reasoning and dynamic stress modeling could turn this framework into a genuine early-warning system, one that connects research with real-world financial decision-making in a transparent and actionable way.

Conclusion

This study brought together several ways of examining financial stress indicators and looked at how they behave when viewed through signal-to-noise analysis and interpretable machine learning. By layering deterministic filters, event-based comparisons, spectral methods, predictive modeling, and information measures, the work showed that VIX, STLFSI4, and CISS each carry their own type of information about market stress. VIX reacts quickly and sharply, often moving ahead of crises, while STLFSI4 and CISS reflect slower and deeper structural pressures. These patterns appeared again and again across every version of the SNR analysis, which suggests that crises develop across multiple time horizons rather than from one simple trend.

The supervised expanding-window models showed that machine learning can pick up real structure in these indicators and deliver solid predictive performance with meaningful lead times. Still, the backtesting results made it clear that good predictions do not guarantee better trading results, which reflects the difficulty of turning early-warning signals into profitable decisions. The use of SHAP helped clarify why the model behaved the way it did, turning the predictions into something closer to a diagnostic lens instead of leaving them as opaque outputs. Taken together, the findings point to the value of combining traditional econometric thinking with interpretable AI to separate true crisis signals from everyday market noise. The framework has limits, especially in the range of indicators used and the way crises were defined, but it provides a solid starting point for future work that could include more diverse stress measures, macro-financial inputs, and adaptive methods that adjust as conditions shift over time. The study offers a clearer, more transparent path for early-warning analytics and strengthens the connection between financial stability monitoring and explainable machine intelligence.

References

[1] Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. American Economic Review, 106(7), 1705–1741.

[2] Athey, S. (2019). The impact of machine learning on economics. In The Economics of Artificial Intelligence: An Agenda (pp. 507–547). University of Chicago Press.

[3] Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. Computational Economics, 57(1), 203–216.

[4] Das, B. C., et al. (2025). AI-Driven Cybersecurity Threat Detection: Building Resilient Defense Systems Using Predictive Analytics. arXiv preprint arXiv:2508.01422.

[5] Danielsson, J., Valenzuela, M., & Zer, I. (2018). Learning from history: Volatility and financial crises. The Review of Financial Studies, 31(7), 2774–2805.

Received: August 07, 2025

[6] 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).

[7] Dzielinski, M. (2012). Measuring economic uncertainty and its impact on the stock market. Finance Research Letters, 9(3), 167–175. https://doi.org/10.1016/j.frl.2011.10.003

[8] Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate GARCH models. Journal of Business & Economic Statistics, 20(3), 339–350.

[9] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.

[10] Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. Journal of Econometrics, 2(2), 111–120.

[11] Hasan, M. S., et al. (2025). Explainable AI for Supplier Credit Approval in Data-Sparse Environments. International Journal of Applied Mathematics, 38(5s).

[12] Hollo, D., Kremer, M., & Lo Duca, M. (2012). CISS – A composite indicator of systemic stress in the financial system. European Central Bank Working Paper Series, No. 1426.

[13] Islam, M. Z., et al. (2025). Cryptocurrency Price Forecasting Using Machine Learning: Building Intelligent Financial Prediction Models. arXiv preprint arXiv:2508.01419.

[14] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Transactions of the ASME – Journal of Basic Engineering, 82(1), 35–45.

[15] Kliesen, K. L., Owyang, M. T., & Vermann, E. K. (2012). Disentangling diverse measures: A survey of financial stress indexes. Federal Reserve Bank of St. Louis Review, 94(5), 369–397.

[16] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.

[17] 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.

[18] 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).

[19] Schwert, G. W. (1989). Why does stock market volatility change over time? The Journal of Finance, 44(5), 1115–1153.

[20] 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).

[21] Shivogo, J. (2025). Fair and explainable credit-scoring under concept drift: Adaptive explanation frameworks for evolving populations. arXiv preprint arXiv:2511.03807.

[22] 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.

[23] Sirignano, J., Sadhwani, A., & Giesecke, K. (2018). Deep learning for mortgage risk. The Journal of Finance, 73(6), 2727–2771.

[24] 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).

[25] Whaley, R. E. (2000). The investor fear gauge. The Journal of Portfolio Management, 26(3), 12–17.