

## Trigger-amplifier-resilience dynamics of short-horizon systemic risk: An interpretable machine-learning early warning system for Indonesia

 Reni Lestari<sup>\*</sup>

<sup>1</sup>Faculty of Economics and Business Universitas Trisakti, Indonesia; 221022401003@std.trisakti.ac.id (R.L.).

**Abstract:** This research develops an interpretable early warning system for short-horizon systemic risk using the Trigger-Amplifier-Resilience (TAR) framework. Rather than treating crises as static outcomes, the framework considers them as evolving processes, where shocks in financial markets act as initial triggers, are subsequently transmitted through financial linkages, and are ultimately shaped by underlying macro-financial conditions. Using monthly Indonesian data from 2015 to 2024, the analysis applies a Random Forest model and uses SHAP decomposition to unpack the drivers behind predicted risk. Crisis episodes are identified based on market-based stress thresholds, and model performance is evaluated through an expanding-window out-of-sample approach. The results suggest that the machine-learning model improves crisis detection relative to a linear benchmark, pointing to the relevance of nonlinear dynamics. More importantly, the decomposition reveals a consistent sequencing pattern: market signals tend to move first, amplification mechanisms follow with a short delay, and resilience factors become more relevant as stress unfolds. Taken together, these findings indicate that combining prediction with interpretation can provide more useful real-time signals, allowing policymakers to respond more precisely to different phases of systemic stress.

**Keywords:** *Early warning system, Indonesia, Interpretable machine learning, SHAP, Systemic risk, Trigger-amplifier-resilience framework.*

### 1. Introduction

Three distinct factors contribute to the widespread increase in systemic financial risks within emerging markets. These three interdependent factors are (i) high dependence on volatile capital inflows that are classified as “pull factors”; (ii) extreme sensitivity of emerging markets to foreign central banks’ monetary policy changes; and (iii) the direct exposure to global risk appetite and investor sentiment [1]. Consequently, changes in global liquidity conditions can induce large corrections in local asset prices and exchange rates, often accompanied by shifts in sovereign risk premia. Moreover, the same channels also enable the rapid transmission of financial instability across financially open systems. Underlying this transmission is the Global Financial Cycle, which is characterized by sudden reversals in capital inflows and highly disruptive movements in the financial markets, with emerging economies being particularly vulnerable to such financial shocks [2, 3].

Financial surveillance is often constrained by the lag between market-based stress signals and the deterioration of macroeconomic or banking indicators within Indonesia’s financial regulatory framework. During the commodity price collapse (2015-2016), emerging market turbulence (2018), and the COVID-19 pandemic (April 2020), equity, foreign exchange, and government securities markets signaled financial stress ahead of the slower-moving economic fundamentals, limiting policymakers’ ability to identify and respond to systemic risk in a timely manner [4, 5]. This sequence illustrates Indonesia’s recurrent vulnerability to rapidly developing financial crises.

Traditional approaches to constructing Early Warning Systems (EWS) have historically focused on vulnerabilities that develop over medium to long-term periods using linear econometric models and

gradually evolving indicators such as credit expansion, leverage ratios, and external imbalances [6-8]. While these classical EWS models are effective at identifying structural vulnerabilities (lack of resilience), they are poor at capturing volatility risk resulting from sudden decreases in asset prices, rather than from the gradual accumulation of balance sheet weaknesses [9, 10]. However, EWS models have begun using machine learning to develop more accurate forecasts of systemic risk [11]; this growing body of literature emphasizes forecasting accuracy over model interpretability. No existing model incorporates the full crisis sequence, trigger, amplification, and resolution, as an integrated temporal process. The lack of a temporal structure in existing EWS models, especially in developing countries, limits their ability to capture the sequential evolution of financial crises.

By developing a framework to identify and categorize financial vulnerabilities, this study fills a significant gap regarding oversight of short-term systemic risk. The Trigger-Amplifier-Resilience (TAR) approach will be used alongside established risk channels, categorized into three temporal groups: fast-moving market indicators act as triggers, financial interconnectedness and balance-sheet weaknesses are amplifiers, and macroeconomic and banking sector fundamentals are resilience factors that ultimately prevent system collapse. An interpretable machine-learning early warning system using a Random Forest model connected with SHAP decompositions will be developed to implement this framework of three descriptive groupings over monthly datasets from Indonesia during 2015-2024.

The study makes three main contributions. First, it develops an operational TAR framework for short-horizon systemic risk, which is a mechanism-based approach for describing how systemic crises develop through the sequential interaction of triggers, amplifiers, and resilience, and empirically validates the temporal ordering of these components. Second, it shows that interpretable machine learning can generate accurate predictions of systemic risk while also providing insight into the underlying transmission mechanisms. Third, it introduces a policy-oriented scanning tool that allows government regulators to develop macroprudential policy responses based on their knowledge of both the dominant transmission channel of systemic risk and the stage of risk development.

## 2. Literature Review

Systemic risk and EWS have led to the independent development of three major research areas. Each area emphasizes different aspects of crisis formation, including various timeframes and pathways. This section integrates these three areas to justify the combined TAR framework developed in this paper.

### 2.1. Medium-to Long-Term Vulnerability Indicators

Traditional early warning systems (EWS) have primarily focused on long-term, slow-moving macroeconomic and financial "imbalances" that accumulate over time before leading to crises. The literature [6, 7, 12] has shown that variables such as credit growth, leverage, current account deficits, and banking sector fragility are relatively strong predictors of medium- to long-term financial crises. However, since these indicators evolve slowly and tend to reflect structural weaknesses rather than imminent financial stress, they offer limited help in identifying real-time imminent financial stress. This inability of traditional indicators to provide early warning is especially problematic for highly open economies that face crises over months rather than years [9].

### 2.2. Short-Horizon Market Triggers and Forward-Looking Signals

A different literature has emerged detailing the importance of high-frequency and market indicators for measuring immediate stress in the market. Asset prices, exchange rates, country risk premiums, and volatility indicators are all considered forward-looking and adjust quickly based on changes in global liquidity, investors' risk preferences, and investor sentiment [4, 5]. Many emerging-market countries experience sudden and extreme changes to these indicators following external shocks. These crises serve as early warning signs preceding significant declines in actual macroeconomic valuations [3, 13].

### 2.3. Amplification through Interconnectedness and Spillovers

Financial stress typically does not remain isolated to a single market or institution; it can spread from one market or institution to another via networks and "spillover" indices, as shocks flow through interbank linkages, cross-asset correlations, and common exposures to magnify the initial disturbance [14, 15]. Herding behavior, sentiment-based trading, and attention-driven trading are additional behavioral mechanisms that can create synchronization among market participants and larger movements [16]. Amplification of these behaviors occurs with relatively short latency. Hence, some triggers evolve into systemic events while others dissipate before generating broader instability.

### 2.4. Resilience through Fundamentals and Buffers

The third component of assessing the resiliency of a country's financial system is the measurement of that system's ability to withstand shocks. There are three key components of resiliency assessment: macroeconomic fundamentals, buffers in the banking sector, and the prevailing policy framework or environment. The combination of these components will define how severe and how long-lasting the stress will be when a shock occurs. Macroeconomic conditions and the buffers that banks maintain to protect themselves from shocks typically do not cause financial crises; however, they will govern the severity of the financial crisis and whether it escalates into a systemic crisis [9, 17]. The interaction between macroeconomic conditions and the health of the banking sector is particularly important. For example, when GDP growth declines, this may reveal vulnerabilities in an economy's underlying balance sheets; however, banks with high capital levels will serve as automatic stabilizers and cushion the effects of the economic downturn [18, 19].

### 2.5. Machine Learning in Early Warning Systems

Recent advances in machine learning have enabled early warning systems (EWSs) to achieve better predictions by capturing nonlinear relationships, interactive effects, and high-dimensional dependence, which are difficult to model with traditional econometric methods Beutel et al. [17] and Chen and Guestrin [20]. Sarlin and Peltonen [21] demonstrate that self-organizing maps can visualize high-dimensional financial stability data, providing intuitive monitoring tools for policymakers. However, most ML-based EWSs prioritize predictive accuracy while neglecting interpretability. Many models treat predictors as unordered inputs and offer reduced-form temporal lag structures, limiting their usefulness for diagnosing transmission mechanisms. Recent advances in explainable AI, such as SHAP attribution methods, help mitigate some limitations by estimating each variable's contribution to the final prediction [22]. Nonetheless, these efforts tend to focus on contributing variables rather than interpreting transmission mechanisms associated with the final prediction, and they generally lack a structural framework linking contributions to individual phases of crisis formation.

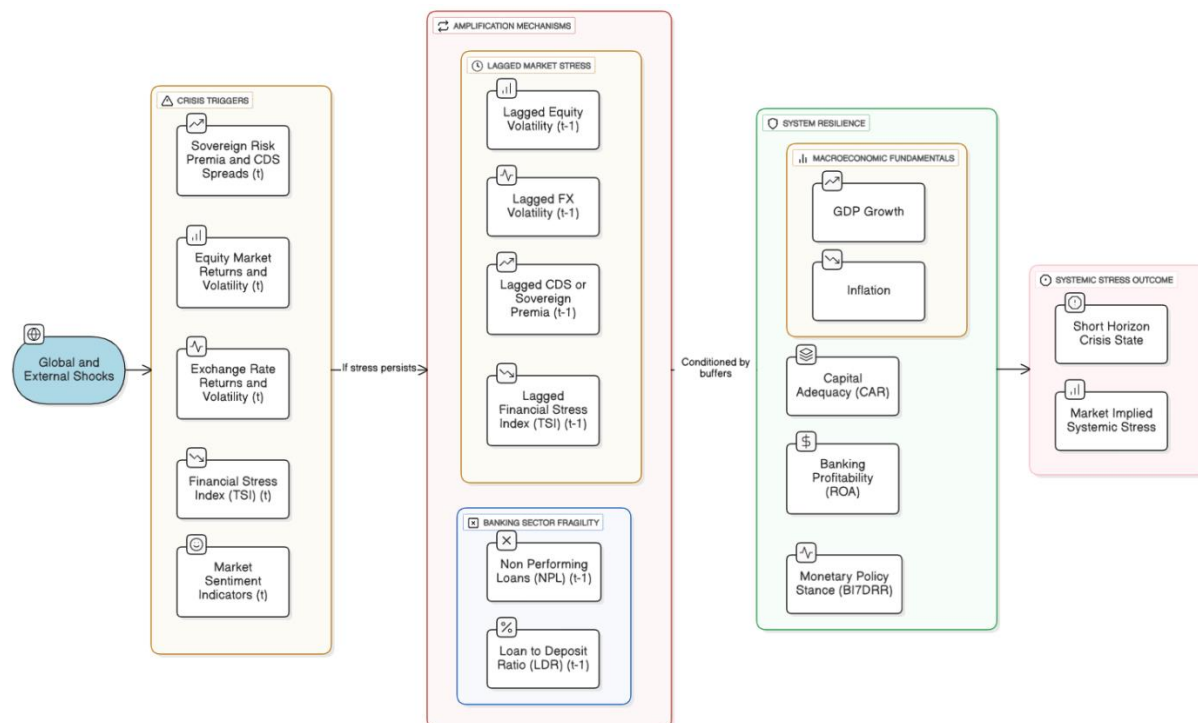
### 2.6. Synthesis and Contribution

Current research has focused individually on the roles of trigger, amplification, and resilience factors in predicting early warning systems; however, no existing frameworks have provided an integrated model for how these three channels are connected in time-based systems designed to assist decision support processes on short time frames. This research provides that missing link through the development of the Trigger Amplifier Resilience (TAR) framework. The TAR framework offers a defined list of risk factors that allows users to organize their risk factors based on a step-wise process as part of building an interpretable machine-learning early warning system. An additional benefit is that the result is a surveillance tool that provides both advanced automated detection of stress and real-time diagnoses of the root causes of stress to appropriate policy responses.

### 3. Methods

#### 3.1. Conceptual Framework: Trigger-Amplifier-Resilience (TAR)

The TAR framework provides a systematic model of short-horizon systemic risk development. Systemic crises are treated not merely as static cross-sectional outcomes but as dynamic processes. The TAR framework identifies three separate but related mechanisms in crisis events. The first mechanism consists of triggers (events that create systemic stress through abrupt changes in markets and market expectations). In addition, there are amplification mechanisms (mechanisms that propagate and accelerate systemic stress through short-term persistence of the initial shock, subsequent spillover effects, and banking system fragility). Finally, there are resilience factors (factors that shape the financial system's ability to absorb shocks and mitigate stress intensity).



**Figure 1.**  
Trigger-Amplifier-Resilience (TAR) Framework.

Figure 1 illustrates the temporal structure of systemic risk development within the TAR framework.

Crisis triggers are signaled by concurrent market indicators ( $t$ ), which exhibit rapid repricing across many markets. Because stress is transmitted to other markets via lagged conditions and financial vulnerabilities of the sector, the amplifying mechanism operates primarily on a short time frame ( $t - 1$ ). The use of the lagged period in the monthly frequency for one period reflects the short horizon for which systemic risk and early warning systems are typically developed, as financial propagation occurs within a one-to-three-month period [23, 24]. Conversely, macroeconomic fundamentals, banking and capital buffers, and policy stances related to the resilience factors of sectors are developed over longer time frames, thereby affecting how shocks will materialize in terms of systemic results.

Formally, the probability of systemic stress at time  $t$  can be represented schematically as:

$$\Pr(\text{Crisis}_t = 1) = f(T_t, A_{t-1}, R_t),$$

where  $T_t$  denotes market-based crisis triggers,  $A_{t-1}$  denotes amplification mechanisms capturing short-run propagation, and  $R_t$  denotes resilience factors reflecting macroeconomic and banking-sector conditions. The TAR framework has a clear empirical implication: if TAR truly reflects how systemic risk develops in Indonesia over time, then SHAP decomposition should indicate that trigger blocks are the most important direct cause just prior to a crisis, that amplifier blocks remain highly elevated after trigger spikes (in the months following), and that resilience blocks develop slowly over time and have changing signs and strengths over the business cycle.

### 3.2. Empirical Implementation: From TAR Framework to Random Forest Model

#### 3.2.1. Feature Vector Construction: Operationalizing the TAR Blocks

To implement the TAR framework as a predictive model, a feature vector is created that corresponds exactly to the three theoretical blocks. For each month  $t$ , the full predictor set  $\mathbf{X}_t$  is:

$$\mathbf{X}_t = [ \underbrace{T_{1,t}, T_{2,t}, \dots, T_{8,t}}_{\text{Trigger Block (contemporaneous)}} \mid \underbrace{A_{1,t-1}, A_{2,t-1}, \dots, A_{7,t-1}}_{\text{Amplifier Block (lagged one month)}} \mid \underbrace{R_{1,t}, R_{2,t}, \dots, R_{5,t}}_{\text{Resilience Block (contemporaneous)}} ]$$

**Table 1.**  
TAR Block Structure and Variable Mapping.

TAR Block	Timing	Description	Variables	Count
Trigger	$t$	Contemporaneous market disturbances that initiate stress.	Equity return, equity volatility, FX return, FX volatility, CDS spread, sovereign risk premium, spillover index, market sentiment	8
Amplifier	$t - 1$	Lagged propagation mechanisms that magnify stress	Lagged equity volatility, lagged FX volatility, lagged CDS spread, lagged sovereign risk premium, lagged spillover index, NPL ratio, LDR	7
Resilience	$t$	Slow-moving conditioning factors that absorb shocks	GDP growth, CPI inflation, policy rate, CAR, ROA	5
Total				20

The TAR framework's temporal structure is incorporated directly into the model architecture. Triggers are introduced simultaneously to capture immediate market repricing; amplifiers are introduced one month later to account for the propagation of shocks through various financial linkages; and resilience factors are introduced concurrently as gradual conditioning variables representing the market's ability to absorb impacts.

#### 3.2.2. Empirical Estimation Equation

The Random Forest model estimates crisis probability as follows:

$$\hat{P}(\text{Crisis}_t = 1 \mid \mathbf{X}_t) = f(T_{1,t}, \dots, T_{8,t}, A_{1,t-1}, \dots, A_{7,t-1}, R_{1,t}, \dots, R_{5,t})$$

where  $f(\cdot)$  is an ensemble of  $B = 500$  classification trees.

This specification captures the temporal structure of TAR; triggers represent immediate repricing, amplifiers represent propagation, and resilience factors condition the severity of systemic stress.

#### 3.2.3. Random Forest Algorithm

Random Forest is chosen because it can (i) capture nonlinear TAR interactions, (ii) achieve better recall than linear benchmarks, which is important for policy, and (iii) make SHAP-based interpretation easier. The generalized random forest framework [25] provides theoretical guarantees for consistency and asymptotic normality, making it especially appropriate for policy-relevant predictions where robustness is critical. Alternative ensemble methods, including XGBoost [20], were also considered, but Random Forest was selected because of its stronger compatibility with SHAP-based interpretability.

Recent improvements to gradient boosting for extreme quantile regression [26] look promising for predicting tail risk, but they are still too computationally expensive for high-frequency monitoring.

Each tree  $b$  is generated from a bootstrap sample. At each node,  $m = \sqrt{20} \approx 4$  characteristics are randomly chosen. The ideal partition optimizes the decrease in Gini impurity:

$$\hat{i}(\tau) = 1 - \sum_{c=0}^1 \hat{p}_{c,\tau}^2, \Delta i(j, s, \tau) = i(\tau) - \frac{n_L}{n} i(\tau_L) - \frac{n_R}{n} i(\tau_R)$$

Trees are grown to a maximum depth of 10, requiring a minimum of 5 samples per leaf. Final predictions aggregate across trees:

$$\hat{P}(\text{Crisis}_t = 1 \mid \mathbf{X}_t) = \frac{1}{B} \sum_{b=1}^B \hat{f}_b(\mathbf{X}_t)$$

### 3.2.4. Model Rationale and Hyperparameters

Although tree-based models are scale-invariant, predictors are standardized (z-scores) to facilitate SHAP comparability and block-level aggregation across variables measured in different units.

$$\tilde{x}_{j,t} = \frac{x_{j,t} - \bar{x}_j}{\sigma_j}$$

Model hyperparameters are determined by five-fold cross-validation. The final specification has 500 trees, a maximum depth of 10, a minimum of 5 samples per leaf, and a feature subsampling ratio of  $m/p = \sqrt{p}/p \approx 0.24$ . The model was built using the scikit-learn Python library [27].

### 3.3. Data and Variable Construction

The empirical analysis employs monthly data from January 2015 through December 2024. Indonesia's prolonged integration into global capital markets and resulting exposure to external shocks make it an ideal case for this analysis.

#### 3.3.1. Crisis Variable Construction

A time series of monthly binary indicators representing short-term systemic stress, defined as episodes of extreme financial market stress rather than ex-post banking or macroeconomic crises.

**Table 2.**  
Crisis Variable Construction.

Stress Channel	Indicator	Threshold	Rationale
Sovereign Risk	5-year sovereign CDS spread	> 250 bps	Corresponds approximately to the upper tail of the in-sample distribution (within the 90-97.5 <sup>th</sup> percentile range); indicates acute sovereign stress consistent with market-based early warning benchmarks [4].
Equity Market	Monthly return of the IHSG index	< -10%	Corresponds approximately to the lower tail of the in-sample distribution (within 5-10 <sup>th</sup> percentile range); signals broad-based valuation collapse [5].
Foreign Exchange	IDR/USD monthly depreciation	> 10%	Corresponds approximately to the upper tail of the in-sample distribution (within 90-97.5 <sup>th</sup> percentile range); consistent with capital-flight benchmarks [1].

**Note:** A month is classified as a crisis period if at least two of the three indicators exceed their respective tail thresholds. Thresholds are calibrated to capture statistically extreme realizations in financial markets rather than being arbitrarily chosen. March-May 2020 satisfies this rule mechanically, without ex-post or narrative classification.

### 3.3.2. Independent Variables: Multiblock Predictor Architecture

Following the TAR classification, the predictor set is partitioned into three role-based categories: triggers, amplifiers, and resilience variables, each capturing distinct phases of systemic risk development. Estimation is performed on normalized data.

#### 3.3.2.1 Trigger Block: Contemporaneous Market Disturbances

Trigger variables are high-frequency, simultaneous market indicators that provoke systemic stress by reflecting abrupt repricing and changes in investor sentiment.

**Table 3.**

Trigger Block Construction.

Indicator	Raw Data	Transformation	Frequency	Source
Equity Return	Equity index level (IHSG)	Log return	Daily → Monthly	Yahoo Finance
Equity Volatility	Daily equity returns	Standard deviation	Monthly	Yahoo Finance
FX Return	USD/IDR exchange rate	Log return	Daily → Monthly	Bank Indonesia
FX Volatility	Daily FX returns	Standard deviation	Monthly	Bank Indonesia
CDS Spread	Sovereign CDS (5Y)	Level	Monthly	Refinitiv
Sovereign Risk Premium	Gov. bond yield - US Treasury	Yield spread	Monthly	FRED
Spillover Index (TSI)	Sectoral equity returns	Diebold-Yilmaz spillover	Daily → Monthly	Self-constructed
Market Sentiment Index	Search/news intensity	PCA factor	Weekly → Monthly	Google Trends

#### 3.3.2.2. Amplifier Block: Propagation Mechanisms

Amplification factors are delayed by one period and reflect the transmission of stress through financial connections, behavioral consistency, and balance-sheet weaknesses. These markers do not initiate crises but amplify and extend stress once activated.

**Table 4.**

Amplifier Block Construction.

Indicator	Raw Data	Transformation	Frequency	Source	TAR Role
Lagged Equity Volatility	Daily equity returns (IHSG)	Std. dev., lagged one period	Monthly ( $t - 1$ )	Yahoo Finance	Amplifier
Lagged FX Volatility	Daily USD/IDR returns	Std. dev., lagged one period	Monthly ( $t - 1$ )	Bank Indonesia	Amplifier
Lagged CDS Spread	Sovereign CDS (5Y)	Level, lagged one period	Monthly ( $t - 1$ )	Refinitiv	Amplifier
Lagged Sovereign Risk Premium	Gov. bond yield - US Treasury	Yield spread, lagged one period	Monthly ( $t - 1$ )	FRED	Amplifier
Lagged Spillover Index (TSI)	Sectoral equity returns	Diebold-Yilmaz spillover, lagged	Monthly ( $t - 1$ )	Self-constructed	Amplifier
Non-Performing Loan Ratio (NPL)	Non-performing loans / total loans	Level, lagged one period	Monthly ( $t - 1$ )	OJK	Amplifier (Fragility)
Loan-to-Deposit Ratio (LDR)	Total loans / total deposits	Level, lagged one period	Monthly ( $t - 1$ )	OJK	Amplifier (Fragility)

#### 3.3.2.3. Resilience Block: Structural Conditioning Factors

Resilience factors are assessed simultaneously and denote the financial system's ability to withstand shocks and alleviate stress intensity. They are not delayed since they denote the present buffering capacity pertinent to the ongoing absorption of stress already in progress.

**Table 5.**  
Resilience Block Construction.

Indicator	Raw Data	Transformation	Frequency	Source	TAR Role
GDP Growth	Real GDP level	YoY growth	Quarterly Monthly →	BPS	Resilience
CPI Inflation	Consumer price index	YoY inflation	Monthly	BPS, BI	Resilience
Policy Interest Rate (BI7DRR)	Central bank policy rate	Level	Monthly	BI	Resilience
Capital Adequacy Ratio (CAR)	Regulatory capital / risk-weighted assets	Level	Monthly	OJK	Resilience (Buffer)
Return on Assets (ROA)	Net income / total assets	Level	Monthly	OJK	Resilience (Buffer)

### 3.4. Model Evaluation

The Area Under the Receiver Operating Characteristic Curve (AUC), along with accuracy, precision, recall (or sensitivity), and F1-score, are used to evaluate model performance. Recall serves as the principal operational metric owing to the asymmetric cost framework of macroprudential surveillance. The social cost of committing a Type II error (failing to identify a systemic crisis) far exceeds that of a Type I error (a false positive precaution) [24, 28].

### 3.5. Interpretability: SHAP Decomposition

#### 3.5.1. Shapley Values

SHAP (SHapley Additive exPlanations) values provide a unified measure of feature importance based on cooperative game theory [29], ensuring that the contribution of each TAR block is fairly attributed according to its marginal contribution to the prediction. The TreeSHAP algorithm, which leverages the tree structure for efficient calculation, calculates the Shapley value for feature  $j$  at time  $t$  as follows:

$$\phi_{j,t} = \sum_{S \subseteq \mathcal{F} \setminus \{j\}} \frac{|S|!(|\mathcal{F}| - |S| - 1)!}{|\mathcal{F}|!} [f(\mathbf{X}_{S \cup \{j\},t}) - f(\mathbf{X}_{S,t})]$$

Let  $\mathcal{F}$  represent the comprehensive set of 20 features,  $f(\cdot)$  denote the Random Forest prediction function, and  $\mathbf{X}_{S,t}$  signify the feature vector comprising solely the features within the subset  $S$ . For tree-based models, this can be calculated efficiently as:

$$\phi_{j,t} = \frac{1}{B} \sum_{b=1}^B \sum_{\text{nodes} \in \text{path}_{b,t}} \Delta_{\text{node}}^{(b)}(j)$$

$\Delta_{\text{node}}^{(b)}(j)$  represents the contribution of feature  $j$  at each node in tree  $b$  along the decision route for observation  $t$ .

#### 3.5.2. Block Aggregation

Block-level contributions aggregate SHAP values within each TAR block:

$$S_{T,t} = \sum_{j \in T} \phi_{j,t}, S_{A,t} = \sum_{j \in A} \phi_{j,t}, S_{R,t} = \sum_{j \in R} \phi_{j,t}$$

These satisfy:  $\hat{p}_t = \hat{p}_{\text{base}} + S_{T,t} + S_{A,t} + S_{R,t}$ . Normalized shares are calculated as follows:

$$\text{share}_{k,t} = \frac{|S_{k,t}|}{|S_{T,t}| + |S_{A,t}| + |S_{R,t}|}, k \in \{T, A, R\}$$

### 3.6. Policy Translation: From Predictions to Actionable Signals

#### 3.6.1. Crisis Alert Threshold

An alert is triggered when the predicted probability exceeds the threshold  $\tau$ :

$$\text{Signal}_t = \mathbb{1}(\hat{p}_t \geq \tau)$$

The baseline  $\tau^*$  maximizes Youden's J ( $J = \text{Sensitivity} + \text{Specificity} - 1$ ). Conservative ( $\tau_H > \tau^*$ ) and sensitive ( $\tau_L < \tau^*$ ) thresholds are also examined to illustrate policy trade-offs.

#### 3.6.2. Phase Identification Thresholds

We adjust block-specific warning thresholds based on the historical distribution of block contributions during non-crisis periods to convert conceptual phase diagnoses into operational guidelines. In accordance with established protocols for early warning system design [9], thresholds are set at the 90<sup>th</sup> percentile of absolute block contributions during non-crisis months:

$$\tau_k = \text{Percentile}_{90}(\{|S_{k,t}| \mid \text{Crisis}_t = 0\}), k \in \{T, A, R\}$$

A block is deemed 'elevated' when its absolute contribution exceeds its threshold.

$$\text{Elevated}_{k,t} = \mathbb{1}(|S_{k,t}| > \tau_k), k \in \{T, A, R\}$$

The phase-specific diagnosis follows a sequential logic consistent with the TAR framework:

**Table 6.**

Phase-Specific Diagnosis and Policy Response.

Pattern	Diagnosis	Implied Policy Response
$\ S_T\  > \tau_T$ only	Trigger-driven stress	Enhanced monitoring, liquidity facilities, and communication
$\ S_T\  > \tau_T$ and $\ S_A\  > \tau_A$	Amplification phase	Containment measures, targeted liquidity, buffer adjustments
All three > thresholds	Full systemic stress	Comprehensive intervention, crisis management protocols
$\ S_T\  < \tau_T$ or $\ S_R\  > \tau_R$ only	Isolated resilience deterioration	Monitor fundamentals; no immediate action

This diagnostic methodology enables policymakers to determine whether increasing systemic risk signals a crisis start (heightened trigger contributions), propagation (increased amplifier contributions), or diminished resilience (positive resilience contributions indicating weakening buffers).

#### 3.6.3. TAR Policy Dashboard

The dashboard presents for each period:  $\hat{p}_t, S_{T,t}, S_{A,t}, S_{R,t}$ , normalized shares, alert status, and phase diagnosis, facilitating mechanism-level macroprudential oversight.

## 4. Empirical Results

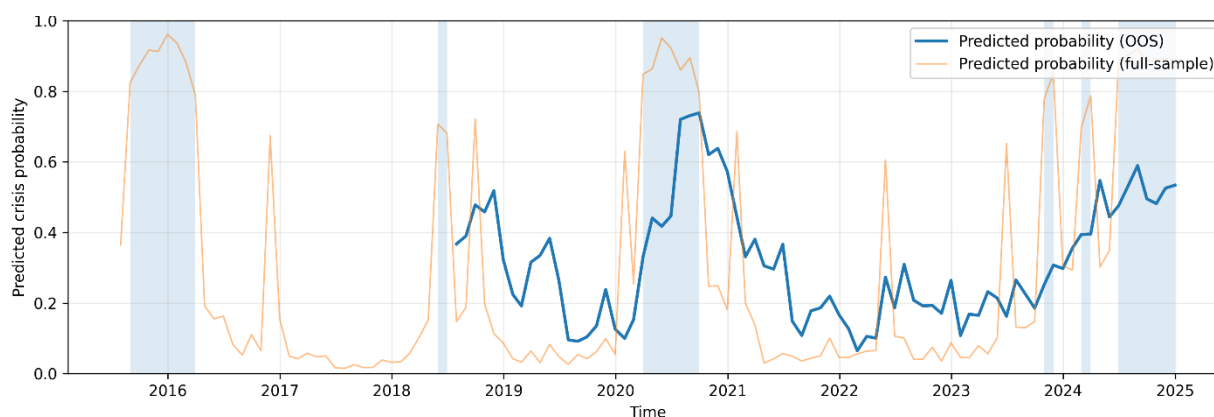
### 4.1. Predictive Performance and Operational Calibration of the Early Warning System

**Table 7.**  
Predictive Model Performance (Out-of-Sample).

Model	AUC	Accuracy	Precision	Recall	F1-score	False Positives	False Negatives
Logistic Regression	0.79	0.71	0.50	0.5	0.50	5	5
Random Forest	0.77	0.74	0.54	0.7	0.61	6	3
XGBoost	0.73	0.69	0.46	0.5	0.48	6	5

**Note:** Models were estimated utilizing the TAR-structured predictor set, which includes contemporaneous triggers, lagged amplifiers, and contemporaneous resilience variables, at a monthly frequency. Out-of-sample predictions are produced using the expanding-window methodology. Total crisis duration was 10 months.

Compared with logistic regression, the Random Forest model achieves higher recall (0.70 vs. 0.50), reducing missed crises from 5 to 3 months. The AUC for logistic regression is slightly higher; however, the low recall value diminishes its usefulness for macroprudential analysis, in which missing a crisis imposes a larger burden than issuing a false signal [28]. The Random Forest model, therefore, provides a better operational trade-off between crisis detection and false alarms.



**Figure 2.**  
Predicted Crisis Probability and Realized Crisis Periods.

**Note:** The blue line represents the out-of-sample predicted crisis probability, while the darkened regions denote ex-post crisis periods.

Figure 2 shows that the out-of-sample predicted probabilities of crises closely match actual stress episodes. The most prominent peaks occur in August-September 2015 during the commodities collapse, August-September 2018 during turbulence in emerging markets, and the April 2020 crisis (0.94) triggered by the COVID-19 pandemic. During tranquil periods, predicted probabilities stay below 0.2, confirming the absence of persistent false signals during those times.

**Table 8.**  
Operational Implications of Alternative Warning Thresholds.

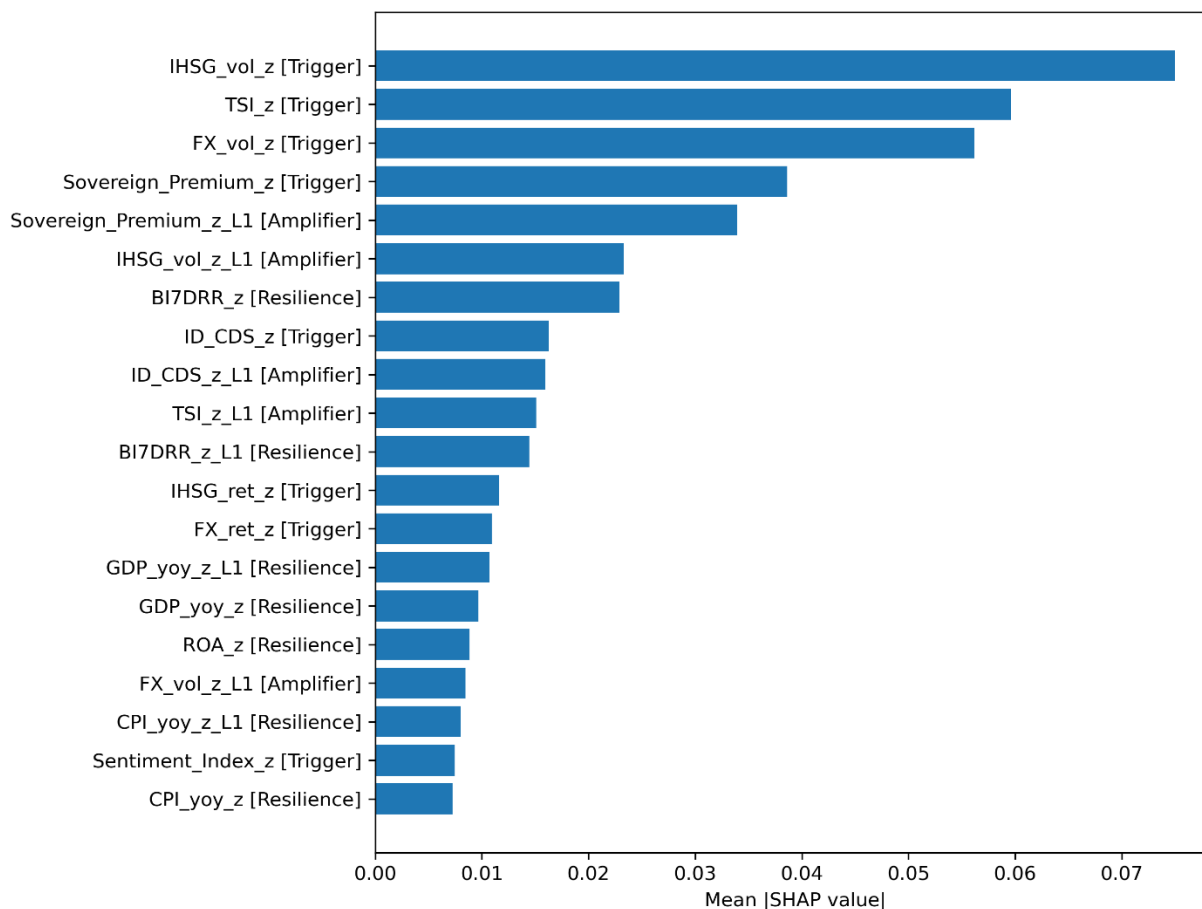
Threshold	$\tau$	Signals/Year	False Alarms/Year	Missed Crises (FN)	Recall	Precision
Baseline ( $\tau^*$ )	0.308	5.11	2.92	3	0.82	0.61
Conservative ( $\tau_H$ )	0.490	3.11	1.30	7	0.59	0.71
Sensitive ( $\tau_L$ )	0.135	8.89	7.78	1	0.94	0.40

**Note:** Thresholds evaluated on out-of-sample predictions.

Baseline  $\tau^* = 0.308$  maximizes Youden's J, detecting 7 of the 10 crisis months while missing three, with 2.9 false alarms annually, manageable within regular policy cycles. Conservative threshold misses 7 of 10 crises (unacceptable); sensitive threshold generates 7.8 false alarms annually (risk of desensitization). The baseline provides a balanced default for continuous monitoring.

#### 4.2. Economic Interpretation of Systemic Risk Dynamics: A TAR Decomposition

##### 4.2.1. Relative Contribution of TAR Blocks to Crisis Probability



**Figure 3.**  
Relative Contribution of TAR Blocks.

Rapid market repricing accounts for approximately 65% of the absolute SHAP contribution. Similarly, the Amplifier block accounts for about 25% of the absolute SHAP contribution because it reflects delayed propagation following the initial shock. Finally, resilience accounts for the remaining 10% of the absolute SHAP contribution, indicating that macro-financial buffers mainly influence the severity rather than the initiation of systemic stress.

**Table 9.**  
Average Block-Level SHAP Contributions by Crisis Status

Block	Non-Crisis Mean	Crisis Mean	Change ( $\Delta$ )
Trigger	-0.235	0.207	+0.442
Amplifier	-0.105	0.082	+0.187
Resilience	-0.075	0.060	+0.135

**Note:** Derived from 104 non-crisis months and 10 crisis months. All differences were statistically significant at  $p < 0.01$ .

All three blocks exhibit negative contributions during non-crisis periods, indicating stabilizing conditions. During a crisis, all blocks change to positive values, increasing the likelihood of a crisis. The magnitude of change indicates that the trigger value (+0.442) is greater than both amplifiers (+0.187) and resilience (+0.135). Therefore, market disruptions remain the primary contributor to systemic stress, while propagation and buffer deterioration are relatively less important. This reversal suggests that linear specifications may fail to capture the nonlinear dynamics of systemic risk.

**Table 10.**  
Distribution of Dominant TAR Phases.

Dominant Phase	Number of Months	Percentage of Sample
Trigger	106	92.98%
Amplifier	6	5.26%
Resilience	2	1.75%
<b>Total</b>	<b>114</b>	<b>100%</b>

**Note:** The predominant phase is determined by the phase with the highest absolute SHAP contribution each month. The sample encompasses 114 months (January 2015 to December 2024).

The trigger phase dominates around 93% of months; market fluctuations rarely escalate into systemic financial events. All six amplifier-dominant months occurred within two months after the three major stress episodes: October–November 2015, September–October 2018, and April–May 2020. The two resilience-dominant months occur during the peak COVID-19 stress episode, particularly in the later stage of the crisis when macro-financial buffers became more influential. These findings indicate that amplification typically occurs after the initial shock, while resilience tends to dominate only during extreme levels of stress.

**Table 11.**  
Timing of Peak Block Contributions Around Crisis Onset

Crisis Episode	Crisis Onset	Trigger Peak	Amplifier Peak	Resilience Turns Positive
2015–2016 Commodity	Aug 2015	Month 0	Month +2	Month +3
2018 EM Turbulence	Aug 2018	Month 0	Month +1	Month +2
2020 COVID-19	Mar 2020	Month 0	Month +1	Month +1
<b>Average</b>		<b>Month 0</b>	<b>Month +1.3</b>	<b>Month +2.0</b>

**Note:** Peak month is defined as the maximum absolute SHAP contribution within a  $[-3, +6]$  month window around crisis onset. "Resilience turns positive" indicates the first month when  $S_R$  shifts from negative to positive values.

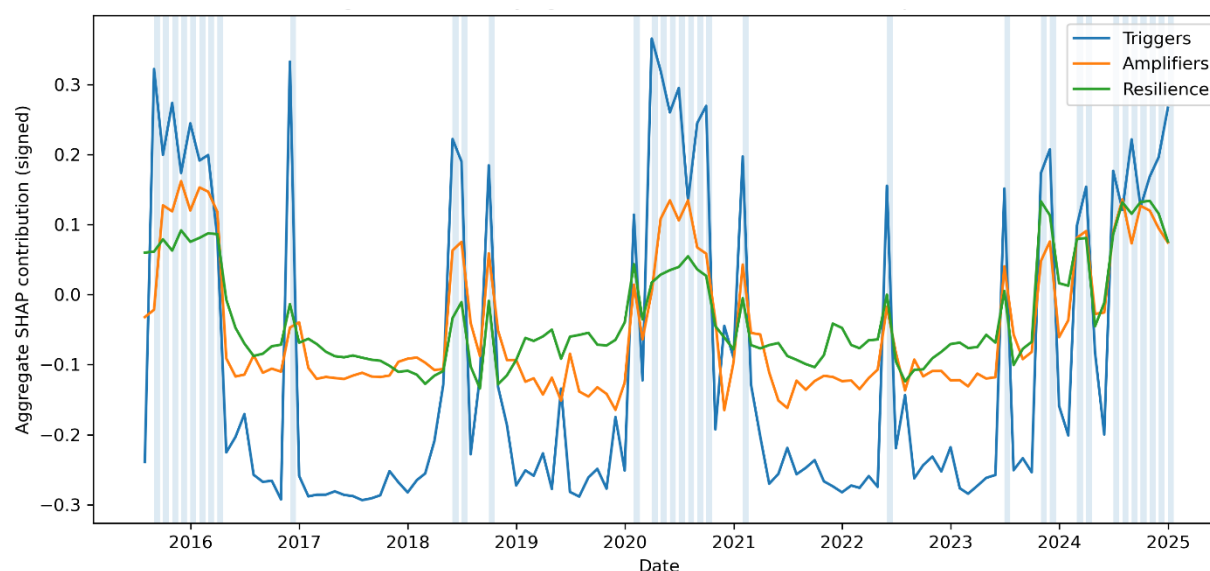
Table 11 reports empirical findings that align with the sequential process described in the Trigger–Amplifier–Resilience (TAR) framework. The trigger component shows the largest contribution at the start of crisis periods ( $t = 0$ ), indicating that systemic stress typically originates from abrupt market repricing and changes in investor sentiment. The amplifier component reaches its peak roughly 1.3 months afterward, suggesting that financial fragilities and balance-sheet connections gradually transmit and intensify the initial shock. Meanwhile, the resilience component tends to become positive about two months after the crisis begins, implying that stabilizing mechanisms, such as policy responses and institutional buffers, start to counteract systemic pressure once the amplification stage has developed.

**Table 12.**  
Thresholds for Phase Identification (90<sup>th</sup> Percentile).

Block	Threshold
Trigger ( $\tau_T$ )	0.185
Amplifier ( $\tau_A$ )	0.095
Resilience ( $\tau_R$ )	0.068

Note: Based on 104 non-crisis months (January 2015–December 2024, excluding the 10 crisis months identified in Table 2).

The thresholds previously described are used as a basis for the phase classifications that will be discussed in Section 4.3. The value of the higher threshold for triggers (0.185) suggests that market-based indicators are more variable than non-market indicators, particularly during periods of reduced volatility. The lower thresholds for amplifiers (0.095) and resilience (0.068) mean that both of these generally stable blocks will respond to changes in value, even though they do not usually have sudden price changes.



**Figure 4.**  
Time-varying TAR Contributions (SHAP Decomposition).

Figure 4 shows a uniform temporal framework throughout the episodes. Trigger contributions increase substantially at the onset of crises (August 2015, August 2018, March 2020), followed by a subsequent decline. Conversely, amplifier contributions accumulate more slowly and typically reach their maximum after a brief interval of one to two months. The amplifier variables are incorporated into the specification with a one-month lag, indicating that this timing represents the propagation of stress through pre-existing financial connections rather than immediate market repricing.

The effects of amplifiers last longer than trigger spikes. Contributions to resilience are negative under normal conditions and become positive only during times of significant systemic stress, especially from April to June 2020. During crises, the sequence remains consistent: market repricing occurs first, followed by propagation channels, with resilience influencing the episode's severity.

### 4.3. TAR Policy Dashboard: Real-Time Diagnostic Signals

**Table 13.**  
TAR Policy Dashboard (Selected Months).

Period	$\hat{p}_t$	$S_T$	$S_A$	$S_R$	Trigger Alert	Amplifier Alert	Resilience Alert	Diagnosis
2015-2016 Commodity Crisis								
Jul 2015	0.12	-0.042	0.011	-0.008	No	No	No	Normal
Aug 2015	0.45	0.182	0.023	-0.015	No	No	No	Elevated but sub-threshold
Sep 2015	0.52	0.203	0.045	-0.021	Yes	No	No	Trigger-driven stress
Oct 2015	0.68	0.156	0.112	0.008	Yes	Yes	No	Amplification phase
Nov 2015	0.61	0.089	0.107	0.032	No	Yes	No	Amplifier-only elevation
Dec 2015	0.43	0.034	0.078	0.018	No	No	No	Normalizing
2018 EM Turbulence								
Jul 2018	0.09	-0.023	0.008	-0.012	No	No	No	Normal
Aug 2018	0.51	0.198	0.034	-0.009	Yes	No	No	Trigger-driven stress
Sep 2018	0.63	0.142	0.108	0.011	Yes	Yes	No	Amplification phase
Oct 2018	0.48	0.067	0.089	0.024	No	No	No	Normalizing
2020 COVID-19 Crisis								
Feb 2020	0.21	0.078	0.021	-0.018	No	No	No	Normal
Mar 2020	0.89	0.203	0.089	0.031	Yes	No	No	Trigger-driven stress
Apr 2020	0.94	0.187	0.167	0.042	Yes	Yes	No	Amplification phase
May 2020	0.91	0.098	0.156	0.071	Yes	Yes	Yes	Full systemic stress
Jun 2020	0.76	0.045	0.112	0.068	No	Yes	Yes	Resilience deterioration
Jul 2020	0.52	0.021	0.078	0.045	No	No	No	Normalizing

**Note:** Alerts are triggered when  $|S_k|$  exceeds  $\tau_k$  ( $\tau_T=0.185$ ,  $\tau_A=0.095$ ,  $\tau_R=0.068$ ). The August 2015 trigger contribution (0.182) somewhat falls short of the threshold (0.185), demonstrating the sensitivity of discrete thresholds. Amplifier block contributions ( $S_A$ ) at any month  $t$  reflect financial conditions observed at  $t - 1$ , consistent with the TAR framework's one-period propagation lag.

The TAR Policy Dashboard offers a real-time assessment of systemic risk, distinguishing crisis probability from the mechanisms generating that risk. Traditional early warning systems often provide a single probability of a crisis, which limits our understanding of the main causes of financial instability. In contrast, the TAR dashboard provides an interpretation by decomposing predictions into contributions from the Trigger, Amplifier, and Resilience components. SHAP-based attribution, derived from the Random Forest model, is used for this purpose. Positive SHAP values indicate that a TAR block increases the likelihood of systemic risk. Conversely, negative contributions signify stabilizing elements that help reduce systemic pressure. This breakdown facilitates the differentiation of initial market disturbances, their amplification via financial weaknesses, and the deployment of stabilizing mechanisms, thus supporting more focused macroprudential interventions.

The usefulness of this approach is illustrated by Table 13, which tracks systemic stress in Indonesia during the three months after the COVID-19 outbreak. In March 2020, the probability of systemic risk spiked. The main driver was the Trigger component, capturing the market's rapid reaction to the global pandemic declaration. At that precise moment, the Amplifier and Resilience components were relatively quiet, confirming that the initial disturbance was a market confidence shock, not an immediate structural failure.

However, the following month revealed a different dynamic. By April 2020, the Amplifier was driven by financial conditions observed in March 2020 ( $t - 1$ ), became the main driver of stress, rising above the baseline thresholds. This pattern indicates that the initial market shock began to propagate through the financial linkages, reflecting the transmission of stress through interconnected balance sheets and liquidity channels. The propagation process is represented in the empirical specification by incorporating amplifier variables lagged by one month ( $t - 1$ ), consistent with the TAR framework, which proposes that financial vulnerabilities magnify the impact of an initial triggering shock. By May 2020, the Resilience block became increasingly influential. Its positive contribution indicates that

macro-financial conditions, including policy interventions and institutional buffers, were actively shaping the evolution of systemic stress. Within the TAR framework, resilience factors do not initiate shocks but condition the severity and persistence of systemic stress during the later stages of the episode.

The three-month sequence, when considered together, illustrates the dynamic transmission mechanism proposed by the TAR framework. An initial market shock (Trigger) first increases systemic stress, which then propagates through financial vulnerabilities (Amplifier), before stabilizing forces begin to absorb the shock (Resilience). Therefore, the dashboard provides policymakers with a diagnostic signal that is specific to each phase, allowing interventions to be targeted based on the main transmission mechanism, rather than relying solely on a single indicator of crisis likelihood.

## 5. Conclusion

### 5.1. Principal Findings

This research develops an interpretable machine-learning early warning system designed for use within the Trigger-Amplifier-Resilience (TAR) framework. The findings empirically support the TAR framework, demonstrating that short-horizon systemic risk in Indonesia evolves through a sequential process. For example, disturbances in financial markets typically occur first, followed by the activation of transmission channels as financial stress propagates, while resilience influences the severity of systemic stress during later stages of the episode. Furthermore, amplification generally occurs with a lag after the initial market shock; thus, policymakers may have a very brief window from the first sign of market repricing until widespread financial distress occurs.

The relative importance of TAR components is not specified a priori but is derived from SHAP decomposition across multiple stress episodes. The trigger contributions increase dramatically when an episode of stress starts (month 0) and reflect very rapid repricing of securities in the market. The amplifier contributions, which are constructed from variables observed one month prior ( $t - 1$ ), peak on average approximately 1-2 months later after the initial trigger spike (+1.3 months on average), consistent with the short-horizon propagation of financial contagion through market linkages. The resilience contributions are only positive after the amplifying effects become pronounced (on average, approximately +2.0 months), suggesting that macro-financial buffers mainly deteriorate during prolonged periods of stress. Similar to other major stressful episodes in the data set, the 2015-2016 decline in commodity prices, the 2018 turmoil in emerging markets, and the shock from COVID-19 in 2020 follow this pattern.

The results from the predictive performance of the analytical models follow a similar pattern. The Random Forest model identified 7 months of crisis in the dataset, while the logistic regression benchmark identified only 5. Therefore, the recall rates are 0.70 for Random Forest and 0.50 for the logistic regression benchmark. Based on the default operational threshold ( $\tau^* = 0.308$ ), the system generates approximately 5.1 alerts annually, with about 2.9 false positives.

### 5.2. Theoretical Contributions

The TAR framework enhances systemic risk theory by framing crisis formation as a sequential phase transition, in contrast to a gradual build-up of vulnerabilities. This phased perspective is strongly supported by empirical evidence, which includes three main observations: consistent lead-lag relationships among TAR components, the temporal concentration of amplifier and resilience phases occurring only after a crisis begins, and a clear magnitude ordering (triggers > amplifiers > resilience). The findings suggest that macroprudential policy should be implemented with a focus on timing, adjusting dynamically to the different stages of financial stress as they appear.

Although implemented using Indonesian data, the conceptual structure is transferable to other financially open emerging markets exposed to global liquidity cycles. Rapid market repricing, followed

by propagation through interconnected exposures and conditional deterioration of buffers, reflects common characteristics of open financial systems rather than country-specific anomalies.

### 5.3. Policy Implications for Macroprudential Surveillance

The TAR-based Early Warning System offers policymakers a phase-specific framework for interpreting systemic risk signals by breaking down crisis probabilities into TAR components. This approach enables decision-makers to identify which underlying mechanism contributes to increased systemic risk, rather than providing a single composite warning.

Periods of trigger-driven stress generally exhibit sharp fluctuations in market repricing before either amplification or resilience components reach alert thresholds. As an example, the trigger component exceeded its alert threshold ( $S_T = 0.203 > \tau_T$ ) During March 2020, while the amplification and resilience components each remained below their respective alert thresholds. In instances such as these, policy measures taken typically consist of communication initiatives (to reduce uncertainty), provision of liquidity support, and heightened oversight of institutions within the financial (and non-financial) sectors. Premature interventions may have the unintended consequence of increasing market stress.

Amplification periods are established when transmission channels begin propagating the initial shock across the financial system. This dynamic began taking place in April 2020 when amplification components exceeded their alert threshold ( $S_A = 0.167 > \tau_A$ ). During such events, supervisors must emphasize monitoring institutions with concentrated exposures, provide targeted liquidity support, and heighten their monitoring of cross-sectoral spillover effects.

A full systemic crisis occurs when observed signals from all TAR blocks exceed their respective thresholds simultaneously, as in May 2020. These events typically happen when macro-financial buffers weaken, requiring policies such as crisis management, temporary capital relief, and coordination with monetary authorities. Conversely, isolated decreases in resilience indicators without sufficient trigger signals may indicate cyclical market corrections rather than systemic failure, requiring consistent monitoring instead of immediate action.

The empirical thresholds identified in the model ( $\tau_T = 0.185, \tau_A = 0.095, \tau_R = 0.068$ ) serve as operational benchmarks for the interpretation of these signals. An interval of time elapsed between the trigger signal and amplification alerts, as experienced during the 2018 event, allowed for some degree of preemptive action prior to the issuance of policy reaction. The events observed during 2020 illustrate the rapidity of stress propagation when multiple shocks simultaneously affect different markets.

In practice, regulators can monitor these block-level signals on a monthly basis. Trigger alerts indicate the requirement for market stabilization; amplification alerts provide indications that containment measures are important; and indicators of a loss of resilience suggest a need for macroprudential intervention. By retaining information on the specific mechanisms through which stress is transferred, the TAR dashboard represents a more informative policy tool than traditional composite risk indices.

### 5.4. Limitations

A few limitations need to be recognized. First, the research focuses on a single developing market over the period 2015–2024 with a limited number of stress episodes. The TAR mechanism is supported by the empirical result, but further validation over longer timelines or countries would strengthen the evidence. Second, systemic crises are inherently rare, therefore limiting statistical power and making empirical conclusions sensitive to how stress episodes are defined. The model is designed to decrease reliance on ex-post crisis labels by evaluating and defining stress signals directly from model results, but the calibration of operational risk thresholds remains sensitive to the crisis definition approach. Third, the identification of stress episodes is mainly based on financial market stress but may be less effective at capturing slower-moving banking or balance-sheet crises. Fourth, while machine-learning

models enhance predictive accuracy, they do not determine causal relationships. The interpretation of the model output requires economic judgment and institutional context.

### 5.5. Directions for Future Research

Future research could extend the TAR framework in at least three directions. First, by using TAR signals in counterfactual scenarios, researchers could create reaction functions conditioned on the dominant TAR phase. This would allow researchers to evaluate if a phase-specific intervention reduces the likelihood of a crisis. Second, the TAR framework could be used for scenario-based stress tests that measure how shocks propagate through the system by using triggering, amplifying, and resilient channels. This information could assist policymakers in allocating macroprudential tools to contain systemic risk. Finally, by validating the TAR framework across multiple countries (developed and emerging) to build confidence in its generalization as well as establish methods of how institutional features influence the TAR dynamic (if Indonesia's pattern is representative of broader systemic risk behavior).

### Transparency:

The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

### Copyright:

© 2026 by the author. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

### References

- [1] K. J. Forbes and F. E. Warnock, "Capital flow waves: Surges, stops, flight, and retrenchment," *Journal of International Economics*, vol. 88, no. 2, pp. 235–251, 2012. <https://doi.org/10.1016/j.jinteco.2012.03.006>
- [2] S. Miranda-Agrippino and H. Rey, "US monetary policy and the global financial cycle," *The Review of Economic Studies*, vol. 87, no. 6, pp. 2754–2776, 2020. <https://doi.org/10.1093/restud/rdaa019>
- [3] H. Rey, "Dilemma not trilemma: The global financial cycle and monetary policy independence," No. w21162. National Bureau of Economic Research, 2015.
- [4] J. Frankel and G. Saravelos, "Can leading indicators assess country vulnerability? Evidence from the 2008–09 global financial crisis," *Journal of International Economics*, vol. 87, no. 2, pp. 216–231, 2012. <https://doi.org/10.1016/j.jinteco.2011.12.009>
- [5] M. Illing and Y. Liu, "Measuring financial stress in a developed country: An application to Canada," *Journal of Financial Stability*, vol. 2, no. 3, pp. 243–265, 2006. <https://doi.org/10.1016/j.jfs.2006.06.002>
- [6] C. E. Borio and P. W. Lowe, "Asset prices, financial and monetary stability: Exploring the nexus," 2002.
- [7] G. L. Kaminsky and C. M. Reinhart, "The twin crises: the causes of banking and balance-of-payments problems," *American Economic Review*, vol. 89, no. 3, pp. 473–500, 1999. <https://doi.org/10.1257/aer.89.3.473>
- [8] C. M. Reinhart and K. S. Rogoff, *This time is different: Eight centuries of financial folly*. Princeton, NJ: Princeton University Press, 2009.
- [9] C. E. Borio and M. Drehmann, "Assessing the risk of banking crises—revisited," *BIS Quarterly Review*, March, pp. 29–46, 2009.
- [10] P.-O. Gourinchas and M. Obstfeld, "Stories of the twentieth century for the twenty-first," *American Economic Journal: Macroeconomics*, vol. 4, no. 1, pp. 226–265, 2012. <https://doi.org/10.1257/mac.4.1.226>
- [11] E. Cerutti, S. Claessens, and P. McGuire, "Systemic risks in global banking: What available data can tell us and what more data are needed?," No. w18531. National Bureau of Economic Research, 2012.
- [12] M. Schularick and A. M. Taylor, "Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008," *American Economic Review*, vol. 102, no. 2, pp. 1029–1061, 2012. <https://doi.org/10.1257/aer.102.2.1029>
- [13] T. Adrian, N. Boyarchenko, and D. Giannone, "Vulnerable growth," *American Economic Review*, vol. 109, no. 4, pp. 1263–1289, 2019. <https://doi.org/10.1257/aer.20161923>
- [14] D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi, "Systemic risk and stability in financial networks," *American Economic Review*, vol. 105, no. 2, pp. 564–608, 2015. <https://doi.org/10.1257/aer.20130456>

- [15] F. X. Diebold and K. Yilmaz, "Better to give than to receive: Predictive directional measurement of volatility spillovers," *International Journal of Forecasting*, vol. 28, no. 1, pp. 57-66, 2012. <https://doi.org/10.1016/j.ijforecast.2011.02.006>
- [16] R. Wermers, "Mutual fund herding and the impact on stock prices," *The Journal of Finance*, vol. 54, no. 2, pp. 581-622, 1999. <https://doi.org/10.1111/0022-1082.00118>
- [17] J. Beutel, S. List, and G. von Schweinitz, "Does machine learning help us predict banking crises?," *Journal of Financial Stability*, vol. 45, p. 100706, 2019.
- [18] G. Caggiano, P. Calice, and L. Leonida, "Early warning systems and systemic banking crises in low income countries: A multinomial logit approach," *Journal of Banking & Finance*, vol. 47, pp. 258-269, 2014. <https://doi.org/10.1016/j.jbankfin.2014.07.002>
- [19] L. Laeven and F. Valencia, "Systemic banking crises database II," *IMF Economic Review*, vol. 68, no. 2, pp. 307-361, 2020. <https://doi.org/10.1057/s41308-020-00107-3>
- [20] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [21] P. Sarlin and T. A. Peltonen, "Mapping the state of financial stability," *Journal of International Financial Markets, Institutions and Money*, vol. 26, pp. 46-76, 2013. <https://doi.org/10.1016/j.intfin.2013.05.002>
- [22] D. Purnell Jr, A. Etemadi, and J. Kamp, "Developing an early warning system for financial networks: an explainable machine learning approach," *Entropy*, vol. 26, no. 9, p. 796, 2024. <https://doi.org/10.3390/e26090796>
- [23] T. Adrian and M. K. Brunnermeier, "CoVaR," *The American Economic Review*, vol. 106, no. 7, p. 1705, 2016. <https://doi.org/10.1257/aer.20120555>
- [24] M. Drehmann and M. Juselius, "Evaluating early warning indicators of banking crises: Satisfying policy requirements," *International Journal of Forecasting*, vol. 30, no. 3, pp. 759-780, 2014. <https://doi.org/10.1016/j.ijforecast.2013.10.002>
- [25] S. Athey, J. Tibshirani, and S. Wager, "Generalized random forests," *Annals of Statistics*, vol. 47, no. 2, pp. 1148-1178, 2019. <https://doi.org/10.1214/18-AOS1709>
- [26] J. Velthoen, C. Dombry, J.-J. Cai, and S. Engelke, "Gradient boosting for extreme quantile regression (arXiv:2103.00808)," *arXiv*, 2021. <https://doi.org/10.48550/arXiv.2103.00808>
- [27] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [28] M. Bussiere and M. Fratzscher, "Towards a new early warning system of financial crises," *Journal of International Money and Finance*, vol. 25, no. 6, pp. 953-973, 2006. <https://doi.org/10.1016/j.jimonfin.2006.07.007>
- [29] S. M. Lundberg *et al.*, "From local explanations to global understanding with explainable AI for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 56-67, 2020. <https://doi.org/10.1038/s42256-019-0138-9>