Forecasting Foreign Exchange Reserves Through Regime Change: A Multi-Model Comparison for Sri Lanka

Abstract

A systematic comparison of forecasting models spanning classical time-series econometrics, Bayesian methods, regime-switching specifications, and machine learning is conducted for predicting Sri Lanka's gross foreign exchange reserves over a period encompassing the country's 2022 sovereign default. Models are evaluated across five variable sets of increasing dimensionality using rolling-origin backtest frameworks, with forecast accuracy assessed via RMSE, Mean, Absolute Scaled Error, probabilistic calibration, and Diebold-Mariano tests. Markov-Switching VAR models emerge as the dominant specification, achieving 72–76% reductions in RMSE relative to a naïve benchmark during the post-crisis period (2023–2025), with near-perfect probability interval calibration (92–100% coverage at the 95% level). The success of regime-switching specifications reflects the distinct accumulation and depletion phases characteristic of reserve dynamics during crisis episodes. XGBoost with engineered features provides competitive performance on monetary and high-dimensional variable sets, while Bayesian VARs offer robustness across longer forecast horizons. The naïve random walk benchmark remains difficult to beat during the acute crisis period itself, echoing the Meese-Rogoff puzzle in a novel macroeconomic context. These findings suggest that explicitly modeling regime dynamics, rather than assuming parameter stability, is essential for reserve forecasting in crisis-prone emerging markets.

Keywords: Foreign exchange reserves, forecasting, Bayesian VAR, regime switching, machine learning, XGBoost, Dynamic Model Averaging, Sri Lanka, sovereign default, emerging markets

JEL Classification: C52, C53, E47, F31, F37

1. Introduction

Foreign exchange reserves constitute the first line of defence for emerging market economies against balance-of-payments crises, currency runs, and sovereign debt distress. Their adequacy or inadequacy can determine whether a country weathers external shocks or spirals into default. Sri Lanka's experience between 2019 and 2022 provides a stark illustration: gross reserves declined from approximately US$7.6 billion at end-2019 to an estimated US$50 million of usable reserves by April 2022, precipitating the country's first sovereign default since independence in 1948 [1](Athukorala, 2024; Wignaraja, 2024). The crisis triggered cascading failures across the fiscal, monetary, and real sectors, resulting in GDP contraction exceeding 7%, inflation surpassing 50%, and widespread social upheaval that ultimately toppled the sitting government (Weerakoon & Jayasuriya, 2023).

Despite the obvious policy importance of reserve forecasting, the academic literature on this subject remains remarkably thin, particularly when compared to the extensive body of work on exchange rate prediction, inflation forecasting, and GDP nowcasting. The few existing studies tend to rely on single-model frameworks, typically ARIMA or reduced-form VARs, and rarely evaluate performance across the kind of regime change that makes forecasting most consequential. This paper addresses that gap by conducting what is, to the authors' knowledge, the most comprehensive multi-model forecasting comparison applied to emerging market reserve dynamics, evaluated against a dataset that spans stable accumulation, pandemic disruption, sovereign default, and IMF-supervised recovery.

The forecasting framework used, comprises fourteen models drawn from four methodological traditions. The classical econometric approach is represented by ARIMA with exogenous regressors, Vector Error Correction Models (VECM), and their Markov-switching extensions (MS-VAR, MS-VECM). The Bayesian tradition contributes a Bayesian VAR with Minnesota prior (BVAR), estimated via Gibbs sampling with hyperparameter grid search, and Dynamic Model Averaging/Selection (DMA/DMS) following the framework of Raftery, Kárný, and Ettler (2010) as adapted to macroeconomics by Koop and Korobilis (2012). The machine learning category includes XGBoost with extensive feature engineering and Long Short-Term Memory (LSTM) networks with sequence modelling. Naïve benchmarks — random walk and seasonal naïve — serve as the standard of comparison, anchoring the analysis in the tradition established by Meese and Rogoff (1983).

Each model class is evaluated across five variable sets of increasing dimensionality: a Parsimonious set (reserves, trade balance, exchange rate), a Balance of Payments set (adding exports, imports, remittances, tourism earnings), a Monetary set (adding broad money), a PCA-derived set (three principal components extracted from eight indicators), and a Full set (all available predictors). This design allows the effect of model specification to be disentangled from the effect of information content, a distinction that is critical when small samples and structural breaks can cause richer models to overfit rather than improve.

The evaluation framework is built around a temporally motivated train-validation-test split that isolates three distinct macroeconomic regimes: a pre-crisis training period through December 2019, a crisis-era validation window spanning COVID-19 and sovereign default (January 2020–December 2022), and a post-default test period covering the IMF programme and early recovery (January 2023 onward). Rolling-window backtests with expanding estimation windows supplement the single-split results. Model comparisons are formalised through Diebold and Mariano (1995) tests and the Model Confidence Set procedure of Hansen, Lunde, and Nason (2011).

Three principal findings emerge. First, in the post-crisis recovery period, XGBoost with lag-based feature engineering dominates all alternatives, achieving a Theil U2 statistic of approximately 0.375, a 62.5% improvement over the naïve benchmark. The model's advantage stems from its ability to exploit nonlinear lag interactions in a period characterised by a strong, momentum-driven reserve rebuilding trajectory. Second, over the full test period that includes the crisis, no model consistently outperforms the random walk, extending the Meese-Rogoff puzzle from exchange rates to reserve dynamics. This finding highlights the fundamental challenge of forecasting through structural breaks: models trained on pre-crisis regularities are poorly equipped to predict the speed and severity of reserve depletion during a sudden stop. Third, Bayesian approaches — particularly BVAR with Minnesota shrinkage and DMA with time-varying model weights — offer the best risk-adjusted performance across regimes, avoiding the catastrophic failures that afflict some classical specifications during the crisis while maintaining competitive accuracy during recovery.

Beyond the model horse-race, a component-level forecasting framework is contributed in which individual balance-of-payments sub-accounts (exports, imports, remittances, tourism earnings, portfolio flows, debt service) are forecast separately and then aggregated into reserve trajectories. This bottom-up approach enables scenario analysis — for instance, projecting reserves under alternative assumptions about tourism recovery speed or debt restructuring timelines, that is directly relevant for central bank reserve adequacy assessment and IMF surveillance exercises. The remainder of this paper proceeds as follows. Section 2 reviews the relevant literature on reserve adequacy, macroeconomic forecasting methodology, and the Sri Lankan crisis context. Section 3 describes the data, variable construction, and cleaning pipeline. Section 4 details the econometric and machine learning specifications. Section 5 presents the main forecasting results and statistical comparisons. Section 6 develops the component forecasting and scenario analysis framework. Section 7 discusses robustness checks. Section 8 concludes with policy implications.

2. Background and Literature Review

Section 2 reviews the relevant literature on reserve adequacy, macroeconomic forecasting methodology, and the Sri Lankan crisis context.

2.1 Reserve Adequacy and Early Warning Systems

The question of how much foreign exchange reserves a country should hold has generated substantial theoretical and empirical literature, producing a succession of adequacy metrics rooted in distinct conceptions of vulnerability. The import cover rule, dating to the Bretton Woods era, proposed reserves equivalent to three months of imports as a minimum prudential buffer. The Guidotti-Greenspan rule, developed after the 1990s emerging market crises, stipulated that reserves should cover at least 100% of short-term external debt maturing within one year, reflecting the risk of a sudden stop in capital flows (Greenspan, 1999). Monetary-based metrics calibrated reserves against a fraction of broad money (M2), typically 5–20% depending on the exchange rate regime, to capture the risk of domestic capital flight (Calvo, 1996; Wijnholds & Kapteyn, 2001).

The International Monetary Fund synthesised these partial measures into a composite Assessing Reserve Adequacy (ARA) metric for emerging markets, which weights short-term external debt (30%), broad money (10%), export revenue (10%), and other portfolio liabilities (20%) under a fixed exchange rate regime, with lower weights under flexible regimes (IMF, 2011; 2013; 2015). The ARA framework represented a significant advance in acknowledging multiple sources of balance-of-payments pressure simultaneously. However, it remains essentially a static threshold model: it tells a country whether its current reserve stock is adequate relative to a snapshot of vulnerabilities, but does not forecast the trajectory of reserves or the dynamics of depletion. The forecasting framework developed in this paper complements these assessments by providing forward-looking projections of reserve paths under alternative scenarios, enabling a more dynamic approach to surveillance.

Obstfeld, Shambaugh, and Taylor (2010) provided an influential financial-stability-based framework for optimal reserve holdings, arguing that the relevant scale variable is financial sector liabilities rather than trade flows alone, since modern balance-of-payments crises are driven as much by capital account reversals as by current account deficits. Jeanne and Rancière (2011) formalised the cost-benefit calculus, modelling optimal reserves as the solution to a precautionary savings problem that balances the opportunity cost of holding low-yielding reserve assets against the expected welfare loss from a sudden stop. Both frameworks imply that reserve adequacy is fundamentally forward-looking, dependent on the probability and expected severity of future crises, which is precisely the kind of assessment that requires credible forecasting models.

A parallel literature on early warning systems (EWS) for currency and balance-of-payments crises has developed indicators related to reserve dynamics. The signal extraction approach of Kaminsky, Lizondo, and Reinhart (1998) identified reserve declines as among the most reliable leading indicators of currency crises, alongside real exchange rate overvaluation and export shortfalls. Frankel and Saravelos (2012) confirmed that the pre-crisis level of reserves is the single most robust predictor of crisis incidence across the entire EWS literature. However, these systems are designed to predict binary crisis events rather than to forecast the continuous path of reserves, and they have limited value for the kind of quantitative scenario analysis required in IMF Article IV consultations or central bank internal exercises. The present study addresses this gap by forecasting reserve trajectories — the continuous path rather than the binary event — across a sample that includes an actual sovereign default.

2.2 Forecasting Approaches for Reserve Dynamics

The foundational challenge for any macroeconomic forecasting exercise was established by Meese and Rogoff (1983), who demonstrated that structural exchange rate models based on monetary fundamentals could not outperform a simple random walk in out-of-sample prediction. This finding, subsequently replicated across dozens of currencies and time periods, became one of the most robust negative results in empirical macroeconomics (Rossi, 2013; Cheung, Chinn, & Pascual, 2005). While originally formulated for bilateral exchange rates, the logic of the Meese-Rogoff puzzle applies with equal force to reserve forecasting: reserves are driven by the same balance-of-payments fundamentals — trade flows, capital movements, debt service, central bank intervention — and exhibit similar nonlinearities, regime changes, and measurement challenges. Taylor, Peel, and Sarno (2001) argued that nonlinear mean reversion becomes detectable when deviations from fundamentals are large, suggesting that the puzzle may break down in periods of strong directional trends — a conjecture that the post-crisis recovery results in this paper appear to confirm.

For multivariate macroeconomic systems, the Vector Autoregression approach introduced by Sims (1980) offers a flexible, atheoretical framework that treats all variables as endogenous. When variables share long-run equilibrium relationships, the Vector Error Correction Model of Johansen (1988; 1991) embeds cointegration constraints that can improve forecast accuracy by anchoring short-run dynamics to economically meaningful attractors. In the reserves context, the cointegrating relationship between reserves, trade flows, and the exchange rate provides a natural error-correction mechanism. The Markov-switching extension introduced by Hamilton (1989) and generalised to multivariate systems by Krolzig (1997) allows model parameters to shift between discrete states governed by an unobserved Markov chain — an architecture that is well-suited to reserve dynamics, which exhibit qualitatively different behaviour during accumulation, crisis depletion, and recovery phases. Peria (2002) and Brunetti, Mariano, Scotti, and Tan (2007) applied Markov-switching models to speculative attacks in European and Southeast Asian settings, establishing precedent for their use in emerging market crisis analysis.

Bayesian methods address the curse of dimensionality in VAR models by imposing informative priors that shrink parameter estimates toward a parsimonious benchmark. The Minnesota prior, developed by Litterman (1986) and Doan, Litterman, and Sims (1984), centres the prior on a random walk representation with diminishing influence from distant lags and other variables. Banbura, Giannone, and Reichlin (2010) demonstrated that BVARs with appropriately calibrated Minnesota priors can forecast as accurately as factor models even with dozens of variables, and Carriero, Clark, and Marcellino (2015) showed that simple specifications with fixed hyperparameters often match or exceed more elaborate alternatives. For emerging market applications where data are scarce and structural breaks frequent, the prior acts as a regulariser that stabilises estimates when the effective sample size is small — a consideration directly relevant for the Sri Lankan dataset, where the most data-demanding variable sets begin only in 2012, yielding roughly 95 training observations. Dynamic Model Averaging (Raftery, Kárný, & Ettler, 2010; Koop & Korobilis, 2012) extends Bayesian model averaging to settings where both the best model and the best set of predictors change over time, maintaining a pool of candidate models whose weights update recursively based on recent predictive performance. Leon-Gonzalez and Thi Bich Nguyen (2021) demonstrated the suitability of DMA for forecasting macroeconomic variables in emerging economies including Sri Lanka, finding that the optimal predictor set changed substantially over time — a result that motivates its use here as a meta-model capable of adapting across regime Transitions.

Gradient-boosted decision trees, particularly XGBoost (Chen & Guestrin, 2016), have emerged as strong competitors to traditional econometric models across a range of forecasting problems. Medeiros, Vasconcelos, Veiga, and Zilberman (2021) found that tree-based methods achieved accuracy competitive with or superior to ARIMA, factor models, and penalised regressions for Brazilian inflation prediction. LSTM networks (Hochreiter & Schmidhuber, 1997), designed to capture long-range temporal dependencies, have been applied to financial and macroeconomic time series with mixed results — their performance is sensitive to hyperparameter tuning, sequence length, and training set size, constraints that are particularly binding in macroeconomic applications where monthly data yield at most a few hundred observations. The broader evidence suggests that machine learning methods offer their greatest advantage in stable periods with strong nonlinear patterns but may underperform simpler approaches during structural breaks when the training distribution diverges sharply from the forecast period (Richardson, Mulder, & Vehbi, 2021). A central contribution of this study is the direct comparison between these approaches and their classical and Bayesian counterparts on identical data, variable sets, and evaluation criteria. Rigorous forecast evaluation requires both normalised accuracy metrics and formal statistical tests. The Theil U2 statistic provides a scale-free measure of relative performance against the naïve random walk. The Mean Absolute Scaled Error (Hyndman & Koehler, 2006) offers an alternative normalisation based on in-sample naïve forecast error. The Diebold-Mariano test (Diebold & Mariano, 1995) tests the null of equal predictive accuracy between competing forecasts, while the Model Confidence Set (Hansen, Lunde, & Nason, 2011) identifies the subset of models containing the best performer with a given probability — particularly valuable when fourteen models across five variable sets generate a large number of pairwise comparisons and the risk of spurious findings from multiple testing is substantial.

2.3 The Sri Lankan Reserve Crisis

Sri Lanka's foreign debt expanded from approximately US$11.3 billion in 2005 to US$56.3 billion by 2020, driven by a reorientation of borrowing from multilateral agencies toward international sovereign bond markets beginning with the first ISB issuance in 2007 (Athukorala, 2024). While the headline debt-to-GDP ratio remained apparently manageable at around 42% in 2019, this figure masked a dangerous maturity and currency mismatch: a substantial share of obligations was denominated in foreign currency with bullet maturities, creating periodic refinancing cliffs dependent on continued market access.

The proximate triggers of the crisis were multiple and reinforcing. The 2019 Easter bombings devastated the tourism sector, which had contributed US$4.4 billion in foreign exchange earnings in 62018. COVID-19 then collapsed tourism revenues to 0.8% of GDP in 2020, while simultaneously disrupting remittance flows and export earnings (Weerakoon & Jayasuriya, 2023). A large tax cut implemented in late 2019 sharply contracted government revenue at the moment fiscal buffers were most needed. An ill-conceived ban on chemical fertiliser imports in 2021, intended to conserve foreign exchange, instead deepened import dependence through agricultural failure.
Critically, the Central Bank of Sri Lanka pursued a policy of defending the pegged exchange rate through reserve sales while simultaneously financing fiscal deficits through money creation — a combination that accelerated reserve depletion without addressing the underlying imbalance. Between September 2021 and March 2022, the authorities used reserves to service maturing ISBs and defend the peg, draining gross reserves from approximately US$3.1 billion to under US$2 billion. When the exchange rate was belatedly floated in March 2022, the rupee depreciated by over 80% within weeks. By April 12, 2022, the Finance Ministry announced a suspension of external debt service — the country's first default since independence. Usable reserves had fallen to an estimated US$25–50 million, insufficient to cover even a single week of imports (UNDP, 2022; Wignaraja, 2024).

An IMF Extended Fund Facility of approximately US$2.9 billion was agreed in principle in September 2022 but not formally approved until March 2023, delayed by complex debt restructuring negotiations involving Paris Club members, China, India, and international bondholders. By early 2024, gross reserves had rebuilt to approximately US$4.4 billion, inflation had returned to single digits, and the rupee had recovered roughly a third of its losses.
This chronology presents an extraordinarily challenging dataset for time-series forecasting models. The reserve series exhibits at least three distinct regimes — gradual accumulation (2005–2019), rapid depletion (2020–2022), and rebuilding (2023–present) — each with fundamentally different dynamics, volatility, and relationships to driving variables. The Bai-Perron structural break test confirms multiple significant breaks coinciding with these regime boundaries. Any model trained exclusively on pre-crisis data must extrapolate into a distributional regime it has never observed, while models that include crisis data in their estimation must somehow weight these extreme observations appropriately when forecasting recovery. This setting provides a particularly stringent test of model robustness and motivates the emphasis on regime-switching specifications and adaptive Bayesian methods.

3. Data and Variable Construction

3.1 Data Sources and Sample

The dataset comprises eleven monthly time series spanning January 2005 to 2026, drawn from four primary sources: the Central Bank of Sri Lanka (CBSL) statistical database, the International

Monetary Fund's International Reserves and Foreign Currency Liquidity (IRFCL) template, the Colombo Stock Exchange market statistics, and the Sri Lanka Customs trade returns. All monetary variables are denominated in US dollars to ensure comparability and to avoid conflating real dynamics with exchange rate valuation effects. The sample period encompasses three macroeconomically distinct phases — gradual reserve accumulation (2005–2019), rapid depletion through pandemic disruption and sovereign default (2020–2022), and IMF-supervised recovery (2023–present) — providing the within-sample regime variation that is essential for estimating and evaluating the regime-switching specifications.

[TABLE 1: Summary statistics — means, standard deviations, min, max, ADF statistics, and ACF(1) for all variables.]

3.2 Dependent Variables

Gross official reserves (gross_reserves_usd_m), measured in millions of US dollars, constitute the primary dependent variable. Following the IMF's IRFCL template, gross reserves are decomposed into five components: foreign currency reserves, IMF reserve position, Special Drawing Rights (SDR), monetary gold, and other reserve assets. This decomposition separates actively managed liquid reserves — primarily foreign currency assets — from quasi-fixed components such as gold and SDR allocations whose movements are driven largely by factors outside domestic policy control. Gross reserves serve as the numerator of every major adequacy benchmark, from the import-cover rule to the Greenspan–Guidotti ratio and the IMF's composite ARA metric, making them the natural dependent variable for any study of reserve dynamics (Jeanne & Rancière, 2011; Obstfeld et al., 2010).

Net usable reserves (net_usable_reserves_usd_m) serve as an alternative dependent variable because gross figures can materially overstate a country's true external position. This concern proved consequential in the Sri Lankan case: when the CBSL activated a USD 1.5 billion swap facility with the People's Bank of China, gross reserves remained elevated while the Net Foreign Assets (NFA) of the Central Bank declined by the corresponding amount, since the swap constituted borrowed rather than earned reserves. NFA of Monetary Authorities turned negative in August 2021 — eight months before the April 2022 sovereign default — providing a materially earlier warning signal than gross figures alone (IMF, 2022; CBSL Annual Report, 2021). Including both measures enables evaluation of whether the regime-switching model detects transitions earlier when estimated on a net basis, a question with direct implications for early-warning system design. Reserve change (reserve_change_usd_m), the first difference of the gross reserve series, serves as the primary forecast target. Joint ADF and KPSS testing confirms that gross reserves are integrated of order one, making level-based regression estimation invalid. The first-differenced series is stationary and exhibits informative temporal structure: a negative first-order autocorrelation (ACF(1) = −0.301), indicating mean-reverting intervention behaviour consistent with the CBSL easing off after heavy defence of the rupee in a given month. This mean-reversion dynamic differs materially between accumulation and crisis regimes — precisely the state-dependent behaviour the Markov-switching framework is designed to capture. Reserve levels are recovered from change forecasts by cumulative summation, following standard practice in the time-series forecasting literature (Meese & Rogoff, 1983; Aizenman & Lee, 2007).

3.3 Balance-of-Payments Flow Variables

The current account's foreign exchange generation capacity is represented by four flow variables, each included as a separate series rather than aggregated into summary measures. The disaggregation is driven by the requirements of the MS-VAR framework, which needs to identify independent shocks to each flow component, a requirement that becomes critical when these components exhibit sharply divergent crisis dynamics.

Merchandise exports (exports_usd_m), measured on a Free On Board (FOB) basis, represent the principal recurring channel through which foreign exchange enters the economy. Export receipts augment official reserves when converted through the banking system, making them a first-order determinant of reserve accumulation. The case for separate inclusion is illustrated by the 2020–2022 crisis period: exports grew approximately 14.3 per cent year-on-year in the first half of 2022, even as imports contracted by 26.1 per cent in June 2022 alone — the latter driven by administrative restrictions and foreign exchange shortages rather than symmetric demand adjustment. Aggregating these sharply divergent dynamics into a single trade balance series would obscure the distinct regime-dependent behaviour of each component and weaken the model's ability to identify structural sources of reserve variation across regimes.

Merchandise imports (imports_usd_m), measured on a Cost, Insurance and Freight (CIF) basis, constitute the primary recurring drain on official reserves. Imports are particularly important in a regime-switching context because their behaviour shifts qualitatively across economic states. During accumulation regimes, imports track domestic demand and commodity prices in a relatively predictable fashion. During crisis regimes, import compression occurs through administrative restrictions — import bans, Letter of Credit requirements, and foreign exchange rationing — that are themselves endogenous to the reserve position, creating a feedback loop in which reserve depletion constrains imports, which in turn moderates further depletion. This simultaneity motivates the VAR specification in which all variables are treated as endogenous. Imports also serve as the denominator of the import-cover rule, the most widely cited reserve adequacy threshold in both IMF surveillance and domestic policy discourse (IMF, 2011; 2015).

Worker remittances (remittances_usd_m) are Sri Lanka's single largest non-debt foreign exchange inflow, covering approximately 80 per cent of the annual trade deficit over the past two decades (CBSL Annual Reports). On an annual flow basis, remittances now exceed the entire official reserve stock, implying unusually high reserve-generation efficiency net of the import content that characterises merchandise export receipts. Remittances are modelled as a separate series rather than being subsumed within the BPM6 category of net secondary income for three reasons. First, their magnitude: remittances are the dominant component of secondary income and behave differently from other transfer flows. Second, their structural persistence: the 80 per cent trade-deficit coverage ratio, sustained over two decades, reflects deep determinants — diaspora size, Gulf labour-market demand, and established migration corridors — that provide a relatively stable baseline of foreign exchange supply independent of trade-cycle fluctuations. Third, their distinct crisis dynamics: the 51.6 per cent decline in cumulative remittances during the first half of 2022 relative to 2021 was driven substantially by channel-switching from formal banking to informal hawala and undiyaal networks in response to the parallel market exchange rate premium — a mechanism that operates independently of the trade account and would introduce omitted variable bias if not modelled separately (Obstfeld et al., 2010).

Tourism earnings (tourism_usd_m) constitute a service export under BPM6 accounting and are therefore not captured within merchandise export figures. Their inclusion ensures complete representation of the current account's foreign exchange generation capacity. The case for separate treatment rests on the series' unique structural break pattern: a complete collapse in 2020 due to COVID-19 border closures, followed by a logistic-curve recovery that differs fundamentally from the gradual import compression and export resilience observed in the goods-trade account. STL decomposition further reveals that tourism earnings possess a seasonal strength of 0.593 — substantially higher than merchandise exports at 0.455 — requiring either deseasonalisation prior to estimation or the inclusion of seasonal dummies. This modelling requirement would be obscured if tourism was aggregated with different seasonal trade flows.

3.4 Financial and Monetary Variables

Colombo Stock Exchange net portfolio flows (cse_net_usd_m) track foreign investor participation in Sri Lankan equity markets and serve as the highest-frequency observable proxy for the investor sentiment shifts that precede sudden stops — including bond market exits and credit line withdrawals (Calvo, 1998; Calvo, Izquierdo, & Loo-Kung, 2013). CSE net flows exhibit the lowest first-order autocorrelation in the dataset (ACF(1) = 0.305), confirming their "hot money" character and low serial dependence — precisely the kind of abrupt reversal behaviour that motivates the regime-switching specification. International Sovereign Bond (ISB) spreads would constitute an arguably superior proxy for sudden-stop risk; however, ISB spread data are not available at monthly frequency for the full sample period, making CSE net flows the best available high-frequency measure of capital-flow reversal risk.

The USD/LKR exchange rate (usd_lkr) is included because, under Sri Lanka's managed float prior to the crisis and during the turbulent 2022 transition to a more flexible regime, the exchange rate is fundamentally endogenous to reserve dynamics. Central Bank intervention to defend a peg directly depletes reserves; reserve depletion eventually forces exchange rate adjustment when the defence becomes unsustainable (Obstfeld et al., 2010). Including the exchange rate allows the MS-VAR to endogenously capture the discrete peg-break transition — from approximately LKR 200/USD to over LKR 360/USD — rather than imposing it exogenously through dummy variables.

The exchange rate also introduces a valuation channel: USD/LKR movements affect the rupee-denominated value of reserve assets and external liabilities.
10Broad money (m2_usd_m) captures a fundamentally different risk dimension from trade or debt-based vulnerability measures: the risk of domestic capital flight. M2 proxies the total pool of liquid domestic assets that could potentially be converted into foreign assets, establishing a theoretical upper bound on this domestic drain. Obstfeld et al. (2010) find that the reserves-to-M2 ratio is a statistically significant predictor of currency crises, with explanatory power incremental to the Greenspan–Guidotti ratio — confirming that monetary vulnerabilities represent a distinct risk dimension. Broad money also enters directly into the IMF's ARA composite metric as a risk-weighted component.

The trade balance (trade_balance_usd_m), computed as the difference between FOB exports and CIF imports, serves a complementary role to the disaggregated trade series. It is used for reserve adequacy ratio construction, trend visualisation, and regime characterisation. In model specifications where the trade balance enters directly — for instance, as a candidate cointegrating variable in the VECM framework — exports and imports are excluded as separate regressors to avoid perfect collinearity. The choice between disaggregated and aggregated trade representations is determined by the specific model and hypothesis under examination.

3.5 Variable Set Construction

Five variable sets of increasing dimensionality are defined to disentangle the effect of model specification from the effect of information content:

The Parsimonious set (3 variables) includes reserves, trade balance, and exchange rate. This minimal specification captures the first-order determinants of reserve dynamics under the import-cover and exchange rate defence frameworks.

The Balance of Payments set (7 variables) includes reserves, exports, imports, remittances, tourism earnings, CSE net flows, and exchange rate. Exports and imports replace the trade balance to allow identification of independent shocks; remittances, tourism, and portfolio flows complete the current and financial account representation.

The Monetary set (8 variables) augments the Balance of Payments set with broad money. The addition of M2 introduces the domestic capital flight channel identified by Obstfeld et al. (2010) and captured in the IMF's ARA metric.

The PCA set (3 components) comprises three principal components extracted from the eight indicators in the Monetary set. This specification tests whether dimensionality reduction can preserve informational content while reducing estimation burden — a consideration that is particularly relevant for the BVAR and MS-VAR specifications, where parameter proliferation in small samples is a binding constraint.

The Full set includes all available predictors simultaneously. This specification serves as an upper bound on available information and provides a test of whether the regularisation mechanisms in the Bayesian and ML models — Minnesota shrinkage, tree-based feature selection, LSTM dropout — can effectively manage the dimensionality.

The five-set design is motivated by the well-documented tension in small-sample macroeconomic forecasting between information gains from additional predictors and estimation error from parameter proliferation (Ba■bura et al., 2010; Carriero et al., 2015). By evaluating each model across all five sets, the analysis identifies whether forecast failures stem from model misspecification or from over- or under-parameterisation.

3.6 Stationarity and Seasonal Properties

All series are subjected to joint ADF and KPSS testing to determine the order of integration. Variables confirmed as I(1) are first-differenced for VAR estimation; where cointegrating relationships are identified via Johansen trace and maximum eigenvalue tests, the VECM specification is estimated in levels with error-correction terms. Seasonal properties are assessed via STL decomposition, with seasonal strength coefficients reported for each series. Variables exhibiting seasonal strength above 0.5 — notably tourism earnings at 0.593 — are deseasonalised using X-13ARIMA-SEATS prior to estimation in specifications that do not include seasonal dummies.

overall comments

Language and Style Assessment