

Bangladesh Macro Pulse: A Composite Stress Index from News-Derived Signals

Methodology Note, Version 1.0

Syed Abul Basher
Mangrove Intelligence, Dhaka
mangroveintel.com/stress-index

April 23, 2026

Abstract

This note documents the construction of the *Macro Pulse Composite Stress Index* for the Bangladesh economy. The index condenses structured economic signals — derived from automated classification of Bangladesh-focused news — across sixteen channels into a single daily 0–100 reading of macroeconomic stress. We describe the signal generation pipeline, the per-channel intensity construction, the percentile-rank normalisation strategy chosen for a short training window, and the weighting scheme. We close with an explicit discussion of limitations and a roadmap for anchoring each qualitative channel to high-frequency observable data.

Contents

1	Purpose	1
2	Data	1
2.1	Source: the signal pipeline	1
2.2	Channels	2
2.3	Sign convention	2
3	Construction	2
3.1	Per-channel intensity	2
3.2	Percentile-rank normalisation	3
3.3	Composite index	3
3.4	Stress zones	3
4	Publication schedule	4
5	Limitations	4
6	Roadmap	4
7	Reproducibility	5

1 Purpose

The Macro Pulse Composite Stress Index is intended as a lightweight *early-warning summary statistic* for analysts, investors, and policy observers tracking the Bangladesh economy. It is not a forecast of any particular variable, nor a substitute for Bangladesh Bank’s official monetary and balance-of-payments releases. Its purpose is to compress a large volume of qualitative economic reporting into a single daily reading and to flag *relative* escalation or moderation in macro conditions across sixteen policy-relevant channels.

2 Data

2.1 Source: the signal pipeline

The index is computed from `data/signals.json`, the structured output of the Macro Pulse pipeline. Each signal i in that file comprises, at minimum, the fields

$$(\text{date}_i, \text{channel}_i, \text{direction}_i \in \{\text{tightening}, \text{easing}\}, \text{confidence}_i \in \{\text{low}, \text{medium}, \text{high}\}).$$

These fields are generated by an LLM classifier (Claude Haiku) applied to Bangladesh-focused RSS headlines collected twice daily. The classifier is governed by a fixed prompt documented in `interpret_events.py`; outputs are validated against a controlled vocabulary and cached.

2.2 Channels

Sixteen structural channels cover the main macroeconomic transmission mechanisms for Bangladesh: Foreign Exchange, Exports, Remittances, Food Prices, Credit, Banking, Energy, Monetary Policy, Fiscal, Political Risk, Agriculture, Logistics, Trade Policy, Labor Market, Geopolitical Risk, and Capital Markets. Channel definitions and metadata (lead indicator, time horizon, economic mechanism, expected effects) are fixed in `signal_definitions.py`.

2.3 Sign convention

Signal direction encodes whether the reported condition worsens a macroeconomic constraint (*tightening*; e.g. reserves falling, export orders cancelled, load shedding intensifying, NPL rising) or relaxes it (*easing*; e.g. reserves replenished, rate cut transmitted, food-price stabilisation). We map

$$d_i = \begin{cases} +1 & \text{if direction}_i = \text{tightening} \\ -1 & \text{if direction}_i = \text{easing} \end{cases}$$

so that the per-channel intensity $I_c(t)$ defined below has the property that higher values indicate more macro stress.

3 Construction

3.1 Per-channel intensity

For channel c on date t , the *raw intensity* is

$$I_c(t) = \sum_{\substack{i \in c \\ 0 \leq t - t_i \leq H}} w(\text{conf}_i) \cdot d_i \cdot \exp\left(-\ln 2 \cdot \frac{t - t_i}{\tau}\right), \quad (1)$$

where $H = 60$ days is the trailing window, $\tau = 14$ days is the half-life governing exponential decay, and confidence weights are

$$w(\text{low}) = 1, \quad w(\text{medium}) = 2, \quad w(\text{high}) = 3.$$

The decay term $2^{-(t-t_i)/\tau}$ ensures that a signal’s contribution halves every τ days. Older signals fade continuously rather than drop off sharply at a fixed cut-off.

Choice of τ . A 14-day half-life balances two concerns. Too short ($\tau \leq 7$) produces a twitchy index dominated by the most recent news cycle; too long ($\tau \geq 30$) smooths the series into a trend indicator that reacts slowly to escalation. Fourteen days is also consistent with the typical “lead time” between news-reportable events (e.g. an LC crunch in the press) and observable macroeconomic flow (e.g. monthly remittance settlements).

Role of the outer window H . The 60-day window is redundant given exponential decay (by day 60, a signal’s weight is $2^{-60/14} \approx 5\%$) but is retained for computational predictability and to guarantee that events beyond two months cannot contribute mechanically.

3.2 Percentile-rank normalisation

Raw intensities are not directly comparable across channels because (i) channels differ in their natural news volume (Energy generates many more headlines than Agriculture in a typical week) and (ii) the short training sample precludes stable variance estimation. We therefore map each raw intensity to its empirical percentile rank:

$$S_c(t) = 100 \cdot \hat{F}_c(I_c(t)), \quad (2)$$

where $\hat{F}_c(\cdot)$ is the empirical CDF of $\{I_c(\tau) : \tau \leq t, I_c(\tau) \neq 0\}$, estimated with the mid-rank convention for ties. Days on which channel c has zero activity are assigned $S_c(t) = 50$ (treated as neutral, not as “historically low stress”).

Percentile-rank has three properties we want for this application:

1. It is distribution-free and robust to the small sample size.
2. It is automatically self-calibrating as history accumulates.
3. It is bounded in $[0, 100]$ without clipping.

3.3 Composite index

The composite stress index $S(t)$ is a weighted average of per-channel scores:

$$S(t) = \frac{\sum_c \omega_c \cdot S_c(t)}{\sum_c \omega_c}, \quad S(t) \in [0, 100]. \quad (3)$$

Channel weights ω_c reflect structural importance to the Bangladesh macroeconomy, informed by the external-sector orientation of the economy and the historical frequency with which each channel has acted as a binding constraint:

Channel	ω_c	Channel	ω_c
Foreign Exchange	15	Fiscal	6
Exports	12	Political Risk	6
Remittances	10	Agriculture	4
Food Prices	10	Logistics	3
Credit	8	Trade Policy	1
Banking	8	Labor Market	1
Energy	8	Geopolitical Risk	1
Monetary Policy	7	Capital Markets	0
<i>Total weight = 100</i>			

Capital Markets is retained in the signal taxonomy but receives zero weight in the composite; it is tracked separately via a hard-data anchor (DSEX and portfolio flow data). Trade Policy, Labor Market, and Geopolitical Risk receive low weight because their signals are frequently second-order or already captured by the FX, Energy, and Exports channels.

3.4 Stress zones

For interpretive clarity the $[0, 100]$ composite is binned into three zones:

Zone	Range	Interpretation
Green	$S(t) \leq 35$	Stress historically muted
Amber	$35 < S(t) < 65$	Stress near historical median
Red	$S(t) \geq 65$	Stress historically elevated

These thresholds are heuristic and are expected to be revisited once the index has accumulated at least twelve months of history and the underlying signal distributions have stabilised.

4 Publication schedule

The pipeline runs twice daily on a scheduled GitHub Action (`.github/workflows/collect.yml`, 00:00 and 12:00 UTC). Each run recomputes `data/signals.json`, recomputes `data/stress_index.json`, and commits both artefacts. The public website reads the JSON directly; no manual step intervenes between data generation and publication.

5 Limitations

Transparency about limitations is an essential part of the product, not a disclaimer appended to it.

1. Classification noise. The underlying signal classifications are produced by a large language model operating on headline text only. Misclassifications occur, particularly around (a) the direction of news containing conflicting cues (“crisis averted”, “growth despite headwinds”), (b) very short or elliptical headlines, and (c) signals that should cross-tag across channels. Prompt iteration and cache cleaning are ongoing.

2. Source selection. The pipeline ingests a fixed set of RSS feeds, weighted toward major Bangladesh English-language outlets (*The Daily Star*, *Business Standard*, *Financial Express*, *Prothom Alo* English) and international wires. Bangla-only reporting is under-represented. No single source should be treated as authoritative; the index’s validity depends on the aggregate.

3. Headline truncation. Classification operates on headlines, not full articles. Magnitude is not extracted: a headline stating “reserves fall” and one stating “reserves fall to twenty-year low” currently receive identical tagging subject only to confidence-level differences.

4. Short training sample. As of the methodology version date the pipeline has accumulated approximately seven weeks of continuous history. The percentile-rank normalisation is therefore calibrated against a very short distribution. Readings will shift as the baseline widens; readers should interpret the absolute level with corresponding caution during this initial period. Relative movement — direction and magnitude of change from week to week — is considerably more reliable than the absolute level.

5. Weighting is structural, not estimated. Channel weights ω_c are prior beliefs informed by the structure of the Bangladesh economy, not estimates from a supervised calibration exercise. A data-driven re-weighting exercise — regressing the index on realised outcomes such as quarterly reserves change, DSEX drawdowns, or reported GDP surprises — will become feasible once twelve to eighteen months of history have accumulated.

6. Publication lag is zero. Unlike official statistics the index has no publication lag, but this also means it contains no revision cycle. Readings for past days will not change; the percentile-rank baseline grows but previous entries are fixed.

6 Roadmap

Planned improvements (in order of priority):

1. **Hard-data anchors.** Each qualitative channel to be paired with one observable high-frequency series — weekly gross reserves (FX), DSEX and portfolio flows (Capital Markets), Tk/USD interbank rate (FX), TCB retail food prices (Food Prices), call-money rate (Banking), monthly EPB exports (Exports), monthly BB remittance inflow (Remittances).
2. **Track record page.** A curated archive of past episodes with the state of the index in the two to six weeks preceding, including honest documentation of misses.
3. **Classifier audit.** Periodic manual review of a random sample of classifications (target 200 per quarter) with a published accuracy metric.
4. **Empirical re-weighting.** Once twelve months of history is available, channel weights are to be revised via a supervised procedure targeting observable stress outcomes.
5. **Bangla-language ingestion.** Extend the collector to Bangla-only outlets (*Prothom Alo* Bangla, *Bonik Barta*, *Daily Kaler Kantho*).

7 Reproducibility

The full pipeline is open source. The repository at <https://github.com/SyedBasher/bangladesh-early-war> contains the classification code, channel definitions, computation script, and this methodology note. The canonical implementation of Equations (1)–(3) is in `scripts/compute_stress_index.py`. The signal corpus itself, from which the index is derived, is published at `data/signals.json`.

Contact

Correspondence on methodology, licensing, or institutional access should be directed to Syed Abul Basher at Mangrove Intelligence.

This methodology note will be versioned. Material revisions are flagged in the header; minor edits (typography, phrasing) are incorporated silently.