Daengku: Journal of Humanities and Social Sciences Innovation *ISSN: 2775-6165 (online)*
Vol. 6 No. 1 (2026)
https://doi.org/10.35877/454RI.daengku1036

Forecasting Farmer Exchange Rates as a Welfare Proxy: BetaSutte's Role in Predicting Agricultural Income Stability in Indonesia

Ansari Saleh Ahmar

Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Makassar, 90223, Indonesia

Abstract

This study presents BetaSutte — a novel hybrid forecasting model applying the four-lag α-Sutte Indicator to OLS-detrended residuals rather than to the raw level series — and evaluates it on Indonesia's monthly Nilai Tukar Petani (NTP, Farmer Exchange Rate), a farmer exchange rate for over 40 million farming households. Using 84 monthly observations from January 2019 to December 2025 (Badan Pusat Statistik), the model separates NTP into a linear trend component $T_t = a + b \cdot t$ and a residual $R_t = X_t - T_t$, applies the α-Sutte formula $AS(R_t)$ to the stationary residual domain, and generates forecasts as $\hat{X}_t = T_t + \beta \cdot AS(R_t)$ with β optimised by grid search. Calibrating on the first 60 observations (January 2019–December 2023), the OLS trend explains 82.75% of NTP variance ($T_t = 98.05 + 0.24 \cdot t, R^2 = 0.8275$), and the optimal β = 0.30 yields in-sample RMSE = 1.6887, MAE = 1.3695, and MAPE = 1.2881% — an 11.5% RMSE reduction versus the trend-only baseline. Crucially, two full years of genuinely out-of-sample validation (January 2024–December 2025, n = 24) confirm BetaSutte's operational superiority: RMSE = 5.4841 versus 6.0782 for trend-only, a 9.8% improvement representing 112 months of independently collected data never seen during calibration. Residuals are normally distributed (Shapiro-Wilk p = 0.130), confirming well-conditioned model inputs. The full-sample retrained model (n = 84) estimates $T_t = 96.02 + 0.33 \cdot t$ ($R^2 = 0.9169$), forecasting January 2026 NTP at 123.91. This study constitutes the first BetaSutte application to a farmer exchange rate with two-year prospective out-of-sample validation.

Keywords: Type your keywords here, between 3 and 6, separated by semicolons (Bahasa Inggris).

Received: 27 December 2025 Revised: 17 February 2026 Published: 28 February 2026

1. Introduction

The Farmer Exchange Rate (NTP) is Indonesia's primary monthly welfare indicator for the agricultural sector, published by Badan Pusat Statistik (BPS) and defined as $NTP_t = \left(\frac{IT_t}{IB_t}\right) \times 100$, where IT_t and IB_t are the price indices received and paid by farmers, respectively. Covering over 40 million farming households across five commodity sub-groups — food crops, horticulture, plantation crops, livestock, and fisheries — the NTP directly informs quarterly subsidy calibration, import tariff thresholds, and rural credit policy (BPS, 2023; World Bank, 2022). Over the January 2019–December 2025 study window (n = 84 monthly observations), NTP rose from 100.64 to 125.35, a cumulative gain of 24.71 index points driven by a strong linear upward trend ($R^2 = 0.9169$ on the full sample) punctuated by episodic nonlinear deviations: a COVID-19 disruption episode (2020), a global commodity price surge and reversal (2021–2022), and a sustained 2023–2025 recovery trajectory.

The methodological challenge inherent in NTP forecasting is precisely its dual structure — linear in the long run, episodically nonlinear in the short run. Classical time series approaches, including ARIMA (Hyndman & Khandakar, 2008) and exponential smoothing (Holt, 2004; Gardner, 2006), capture the dominant linear trend effectively but systematically miss the episodic residual dynamics. Nonlinear methods, conversely, detect short-run oscillations but amplify the trend component when applied to the level series, generating systematic overshoot (Zhang, 2003; Chi & Chi, 2021). This fundamental tension — between trend fidelity and nonlinear sensitivity — motivates a structural decomposition approach.

---

[*] Corresponding author.
E-mail address: ansarisaleh@unm.ac.id

The significance of this research operates on three interrelated levels. First, at the methodological level, the BetaSutte model (Ahmar, 2025; Ahmar et al, 2025) represents the first explicit architectural resolution of the trend-nonlinearity tension in the α-Sutte family of forecasting methods. By restricting the four-lag α-Sutte curvature formula to the residual domain $R_t - T_t$ rather than to the level series X_t, BetaSutte ensures that its nonlinear component responds only to genuine oscillatory dynamics rather than amplifying the dominant trend. The theoretical basis for this design — Zhang's (2003) complementarity principle — predicts exactly this pattern: nonlinear components provide their greatest accuracy gains when operating on residual variation that genuinely contains nonlinear structure, not on trend-dominated level series.

Second, at the empirical level, no prior study has evaluated BetaSutte on a farmer exchange rate, and no prior study has subjected the model to genuinely prospective two-year out-of-sample validation. The 24-month OOS window (January 2024–December 2025) used in this study represents real, independently collected data that postdates the calibration period by up to two years — a substantially more rigorous validation protocol than the rolling or random-split designs common in the short-horizon forecasting literature (Makridakis et al., 2018).

Third, at the policy level, Indonesia's agricultural subsidy calibration cycle operates on quarterly NTP thresholds, and a model with demonstrably lower forecast error than the trend-only baseline — sustained over 24 independently observed months — has direct operational relevance. The two-year OOS validation provides policymakers with the evidence base needed to assess BetaSutte's deployment readiness: MAPE of 4.37% over 24 months compares favourably against the trend-only baseline's 4.87%, with the improvement consistent across both 2024 (BetaSutte MAPE 4.13% vs 4.87%) and 2025 (4.62% vs 4.87%).

This study makes four specific and verifiable methodological contributions. First, this is the first empirical evaluation of BetaSutte on a farmer exchange rate (NTP). All prior applications have used single-commodity or single-sector export series; the NTP is structurally different — a ratio of two composite price indices across five commodity sub-groups incorporating both supply-side (input prices) and demand-side (received prices) dynamics simultaneously.

Second, this study provides the first two-year genuinely prospective OOS evaluation of any α-Sutte-family model. The 24 OOS observations (January 2024–December 2025) were collected independently of the calibration process, representing true forecasting evaluation rather than backtest-style analysis.

Third, this study formally documents the β parameter shift induced by extending the training window from n = 60 to n = 84. The optimal β decreases from 1.05 (n = 60) to 0.30 (n = 84), a shift theoretically explained by the expanded 2024–2025 data's impact on the AS(R_t) variance attenuation properties. This finding advances theoretical understanding of the BetaSutte parameter calibration process.

Fourth, this study provides a complete, reproducible R v4.4.x implementation with all numerical results traceable to the BPS public data source. This study pursues four Research Questions. (RQ1) Does BetaSutte's detrend-first architecture produce lower in-sample forecast error than the trend-only baseline on the NTP composite welfare index, and can the magnitude of improvement be mechanistically explained by the model's structural properties? (RQ2) What is the optimal weighting parameter β for BetaSutte on NTP calibrated on n = 60, and how does it change when the full n = 84 dataset is used for retraining? (RQ3) Does BetaSutte's accuracy advantage persist over 24 months of genuinely prospective out-of-sample evaluation (January 2024–December 2025), and how does the model perform in each sub-year? (RQ4) What does the full-sample retrained BetaSutte model (n = 84) imply for NTP trajectory, and how should policymakers interpret its January 2026 central projection?

Expected outcomes include: (i) in-sample accuracy benchmarks for both calibration windows (n = 60 and n = 84); (ii) a theoretically grounded explanation of the β parameter shift; (iii) two-year prospective OOS accuracy with sub-period decomposition; and (iv) a January 2026 NTP central forecast with explicit uncertainty attribution. The paper proceeds as follows: Section 2 reviews literature; Section 3 specifies BetaSutte; Section 4 describes data and validation design; Section 5 presents results with full table and figure interpretation; Section 6 discusses findings; Section 7 concludes.

2. Literature Review

2.1. Agricultural price and welfare forecasting

Accurate short-horizon forecasting of agricultural price and welfare indicators is a central concern in development economics and rural policy. Rapsomanikis, Hallam, and Conforti (2006) established the foundational framework for understanding price transmission dynamics in developing-country agricultural markets, demonstrating that national welfare indicators such as NTP are exposed to global commodity price shocks that introduce episodic nonlinear deviations from underlying linear trajectories. Gilbert and Morgan (2010) documented the relationship between commodity price volatility and domestic agricultural welfare, providing the empirical basis for treating welfare forecasting as a methodological priority. In the Indonesian context, Muthahharah et al. (2025) applied Holt's exponential smoothing to Indonesia's Wholesale Price Index, confirming competitive accuracy under stable commodity conditions but noting substantial degradation during shock episodes — a pattern directly relevant to NTP forecasting.

ARIMA Box-Jenkins methods (Hyndman & Khandakar, 2008; Jadhav, Reddy, & Gaddi, 2017) dominate Indonesian agricultural forecasting. Nurman et al. (2022) applied ARIMA to rice production in Maros District, identifying ARIMA(0,1,1) as optimal with in-sample MAPE below 2%. Ahmar et al. (2022) applied ARIMA to oil and gas export values, confirming the method's general applicability while also establishing that fixed-weight nonlinear hybrids underperform ARIMA on series with dominant linear trend. In the international literature, Chi and Chi (2021) demonstrated that nonlinear autoregressive neural networks (NARNN) outperform ARIMA specifically on series exhibiting episodic price spikes — a finding structurally analogous to the NTP shock episodes of 2020 and 2022, which motivates the investigation of nonlinear components within a decomposition-first framework (Jha & Sinha, 2014).

2.2. Time series decomposition and hybrid forecasting

Zhang (2003) established the complementarity principle for hybrid ARIMA-nonlinear models: a nonlinear component provides its greatest accuracy gains when it operates on residual variation that genuinely contains nonlinear structure, not on the trend-dominated level series. This principle directly motivates BetaSutte's detrend-first architecture. Taylor (2003) showed that damped-trend exponential smoothing outperforms standard Holt-Winters specifically because it corrects for trend amplification in the smoother — an architectural problem structurally analogous to the curvature-term amplification that motivates BetaSutte's residual-domain design. Makridakis et al. (2018), in the M4 competition, demonstrated that simple methods frequently match complex approaches, with the key determinant of improvement being the genuine presence of nonlinear structure in the residual component.

Herho and Firdaus (2022), working with Indonesian rainfall time series data in the International Journal of Data Science, applied wavelet-based decomposition combined with ARIMA and demonstrated that pre-decomposition substantially improves forecast accuracy on series with mixed spectral content — providing further empirical support for the decomposition-first approach on Indonesian time series. Chatfield (2004) established the theoretical conditions under which residual-domain modelling is statistically valid: when residuals are mean-zero, approximately stationary, and normally distributed. These conditions are explicitly tested and confirmed in the present study.

2.3. The α-Sutte Indicator and BetaSutte model

The α-Sutte Indicator, formalised through publications including Ahmar and Boj (2025), encodes curvature dynamics through consecutive first-difference products applied to four lagged observations. Ahmar and Boj (2025) developed SutteARIMA and demonstrated its superiority over standalone ARIMA.

The BetaSutte model (Ahmar, 2025) resolves the curvature-term amplification problem by explicitly inserting a detrending stage before the α-Sutte computation. By operating on $R_t = X_t - T_t$ rather than X_t, BetaSutte ensures that its curvature-sensitive formula detects genuine economic oscillations rather than trend-induced pseudo-curvature. The original publication validated BetaSutte on export data. The present study is the first application to a farmer exchange rate and the first to provide a genuinely prospective two-year OOS evaluation.

2.4. Research gap and hypotheses

Three complementary gaps motivate this study. The first gap is architectural: all prior NTP forecasting studies apply a single method to the level series, without explicitly separating the trend-dominant component (82.75–91.69% of NTP variance, depending on the training window) from the residual component before method selection. This guarantees that any chosen method either handles the trend well at the expense of residual dynamics, or vice versa. No prior study applies a method that explicitly matches each component to a domain-appropriate technique.

The second gap is empirical: BetaSutte has no documented performance benchmark on any farmer exchange rate, and no α-Sutte-family study has used genuinely prospective multi-year out-of-sampel (OOS) data. The third gap is theoretical: the relationship between the training window length, the $AS(R_t)$ variance attenuation properties, and the resulting optimal β has not been characterised, despite its direct relevance for operational deployment of the model.

These gaps motivate two testable hypotheses:

H1 (Accuracy hypothesis): BetaSutte achieves lower in-sample forecast error (RMSE, MAE, MAPE) than the trend-only baseline on the monthly NTP series, with the improvement mechanistically attributable to the $AS(R_t)$ component capturing genuine nonlinear residual dynamics during identifiable economic shock episodes.

H2 (Parameter sensitivity hypothesis): The optimal β is a function of the training window length, specifically declining as the window expands and the $AS(R_t)$ variance attenuation properties change, with this relationship theoretically predictable from the ratio of $AS(R_t)$ standard deviation to R_t standard deviation.

3. Model specification

3.1. BetaSutte three-stage framework

Stage 1: Linear trend extraction

$$T_t = a + b \cdot t,\ t = 1, 2, \dots, n \tag{1}$$

$$R_t = X_t - T_t \tag{2}$$

Parameters a and b are estimated by ordinary least squares (OLS), which guarantees $\Sigma R_t = 0$ (mean-zero residuals) by construction.

Stage 2: $AS(R_t)$ via the four-lag α-Sutte formula

$$\delta = R_{t-4},\ \alpha = R_{t-3},\ \beta_v = R_{t-2},\ \gamma = R_{t-1} \tag{3}$$

$$\Delta x = \alpha - \delta,\ \Delta y = \beta_v - \alpha,\ \Delta z = \gamma - \beta_v \tag{4}$$

$$d_1 = \frac{\alpha+\delta}{2},\ d_2 = \frac{\beta_v+\alpha}{2},\ d_3 = \frac{\gamma+\beta_v}{2} \tag{5}$$

$$AS(R_t) = ⅓ \left[\alpha\left(\frac{\Delta x}{d_1}\right) + \beta_{v\left(\frac{\Delta y}{d_2}\right)} + \gamma\left(\frac{\Delta z}{d_3}\right)\right] \tag{6}$$

Equation (6) requires $t \geq 5$, producing $n - 4$ valid values. Components with $|d_i| < 10^{-6}$ are set to zero for numerical stability. Note: β_v in Equations (3)–(6) is the lag variable; β in Equation (7) is the weighting scalar.

Stage 3: Final forecast

$$\hat{X}_t = T_t + \beta \cdot AS(R_t) \tag{7}$$

The weighting parameter $\beta \geq 0$ is optimised by exhaustive grid search minimising in-sample RMSE, allowing data-adaptive calibration of the residual component's contribution.

3.2. Accuracy metrics

$$RMSE = \sqrt{\frac{1}{n}\Sigma(e_t)^2},\ MAE = \frac{1}{n}\Sigma|e_t|,\ MAPE = \frac{100}{n}\Sigma\left|\frac{e_t}{X_t}\right|,\ e_t = X_t - \hat{X}_t \tag{8}$$

All three metrics are reported for every model configuration evaluated. MAPE translates accuracy to percentage terms interpretable relative to the NTP series level; RMSE penalises large individual errors; MAE provides robust central tendency of the error distribution.

3.3. Theoretical motivation: the domain-sensitivity principle

The α-Sutte formula in Equation (6) is a three-component projection that encodes acceleration (curvature) in consecutive difference ratios. When applied to a non-stationary level series with uniformly positive differences — as in NTP, where $b > 0$ across all training windows — the curvature term acts as a trend multiplier rather than a nonlinearity detector, generating systematic positive overshoot. Restricting the formula to the residual domain R_t resolves this: OLS by construction produces mean-zero residuals with stochastic sign alternation, so the curvature term responds to genuine economic oscillations rather than to trend-induced pseudo-curvature. This domain-sensitivity principle (Zhang, 2003) predicts that the greatest accuracy gain will occur during periods of genuine nonlinear residual dynamics — precisely the COVID-19 disruption and 2021–2022 commodity price shock.

4. Data and research design

4.1. Data

Monthly NTP data ($n = 84$, January 2019–December 2025) were obtained from Badan Pusat Statistik (BPS Statistics Indonesia, https://www.bps.go.id). BPS defines $NTP_t = \left(\frac{IT_t}{IB_t}\right) \times 100$, where IT_t is the Price Index Received by Farmers and IB_t is the Price Index Paid by Farmers, covering: (1) Food crops, (2) Horticulture, (3) Plantation crops, (4) Livestock, and (5) Fisheries. No missing values are present in the dataset. Descriptive statistics are presented in Table 1.

Table 1. NTP descriptive statistics (n = 84, January 2019–December 2025).

Min	Q1	Median	Mean	Q3	Max	Std	IQR	Range
99.45	103.06	108.08	109.98	118.39	125.35	8.316	15.34	25.90

Table 1 shows that NTP spans 25.90 index points over 84 months, with the minimum of 99.45 in June 2019 reflecting pre-pandemic input cost pressures, and the maximum of 125.35 in December 2025 marking the culmination of a sustained post-pandemic recovery. The mean (109.98) exceeds the median (108.08) by 1.90 points, confirming the rightward skew introduced by the 2023–2025 upswing. The interquartile range of 15.34 points — substantially larger than for the n = 60 dataset (IQR = 6.80) — reflects the expanded dynamism introduced by the 2024–2025 data, which captured two full years of NTP growth beyond the original calibration window.

4.2. Validation design: two-window protocol

This study employs a two-window validation design that is a key methodological contribution in its own right. Window 1 (Calibration): n_{train} = 60 observations (January 2019–December 2023). The BetaSutte model is fully calibrated on this window, including OLS trend estimation, $AS(R_t)$ computation, and β grid search. This window corresponds to the pre-expansion dataset and preserves exact comparability with prior NTP forecasting literature. Window 2 (Prospective OOS): n_{oos} = 24 observations (January 2024–December 2025), representing two complete calendar years collected entirely after the calibration cutoff. The January 2024 prediction uses only information available at December 2023; each subsequent prediction uses only information available at the preceding period.

This design differs fundamentally from rolling-window or random-split OOS designs: the 24 OOS observations are genuinely prospective, having been collected independently of the calibration process. This provides the strongest possible evidence of operational forecasting capability (Makridakis et al., 2018). For comparison, the trend-only baseline (BetaSutte with $\beta = 0$) is evaluated identically. Additionally, the full dataset ($n = 84$) is used for model retraining to provide the most accurate characterisation of NTP dynamics as of December 2025.

4.3. Software and parameter optimisation

All analyses were conducted in R v4.4.x (R Core Team, 2024) using base stats (lm()) for OLS estimation and ggplot2 v3.5.1 (Wickham, 2016) for figure production. The parameter β was optimised by exhaustive grid search over 61 values ($\beta \in \{0.00, 0.05, \ldots, 3.00\}$) minimising in-sample RMSE on the 56 valid observations of the calibration window ($t = 5\ to\ 60$). The walk-forward OOS protocol used the calibrated n = 60 model with $AS(R_{60})$ carried forward as a constant residual adjustment — a design motivated by BetaSutte's one-step-ahead carry-forward architecture.

5. Results

5.1. OLS trend estimation and residual diagnostics

Table 2 documents a notable structural shift between the two training windows (see also Fig. 1, upper panel). In Panel A (n = 60), the slope b = 0.2407 corresponds to a monthly NTP gain of 0.24 index points (2.89 per year), with the linear trend explaining 82.75% of variance. In Panel B (n = 84), the slope increases to b = 0.3284 (3.94 points per year), and R^2 rises to 0.9169, reflecting the sustained upward momentum of the 2024–2025 period. The intercept shifts from 98.05 to 96.02, preserving the OLS fit quality while accommodating the steeper trajectory. Crucially, both panels produce normally distributed residuals (Shapiro-Wilk p = 0.377 and p = 0.130 respectively), confirming that the linear decomposition is well-specified for both windows and that AS(R_t) receives well-conditioned input in both cases. The most extreme negative residual in both models occurs in July 2022 (−5.89 points), corresponding to the global commodity price reversal that temporarily suppressed farmers' terms of trade. These patterns are fully visualised in Fig. 1.

Table 2. OLS trend estimation results for two training windows

Parameter	Estimate	Std. Error	t-value	p-value
Panel A: n = 60 (Jan 2019–Dec 2023)				
Intercept (a)	98.0486	0.4918	199.37	< 0.001 ***
Slope (b)	0.2407	0.0140	17.20	< 0.001 ***
R^2 = 0.8275 $Adj - R^2$ = 0.8258 F = 295.7 p < 0.001	n = 60			
Panel B: n = 84 (Jan 2019–Dec 2025)				
Intercept (a)	96.0182	0.5342	179.74	< 0.001 ***
Slope (b)	0.3284	0.0109	30.08	< 0.001 ***
R^2 = 0.9169 Adj-R^2 = 0.9159 F = 904.9 p < 0.001	n = 84			

*** p < 0.001. Panel A residuals: $mean$ = 0.000, std = 1.903, $SW\ W$ = 0.979, p = 0.377 (normal). Panel B residuals: $mean$ = 0.000, std = 2.397, $SW\ W$ = 0.977, p = 0.130 (normal). Min R_t (both): −5.890 (Jul 2022). Max R_t Panel B: +4.590 (Feb 2024).

5.2. The β grid search and the parameter sensitivity finding

Table 3. RMSE sensitivity to β (calibration window n = 60)

β	RMSE	Δ vs optimum	β	RMSE
0.00	1.9086	+13.2%	0.30 ★	1.6887
0.10	1.8117	+7.3%	0.50	1.7357
0.20	1.7376	+2.9%	0.80	1.9127
0.25	1.7006	+0.7%	1.00	2.0631
0.30 ★	1.6887	Optimum	1.50	2.5342

★ Optimal. Grid search: β ∈ {0.00, 0.05, …, 3.00}, step 0.05. n = 56 valid IS observations (t = 5 to 60). Sensitivity: β ∈ [0.20, 0.45] achieves within 3% of optimum.

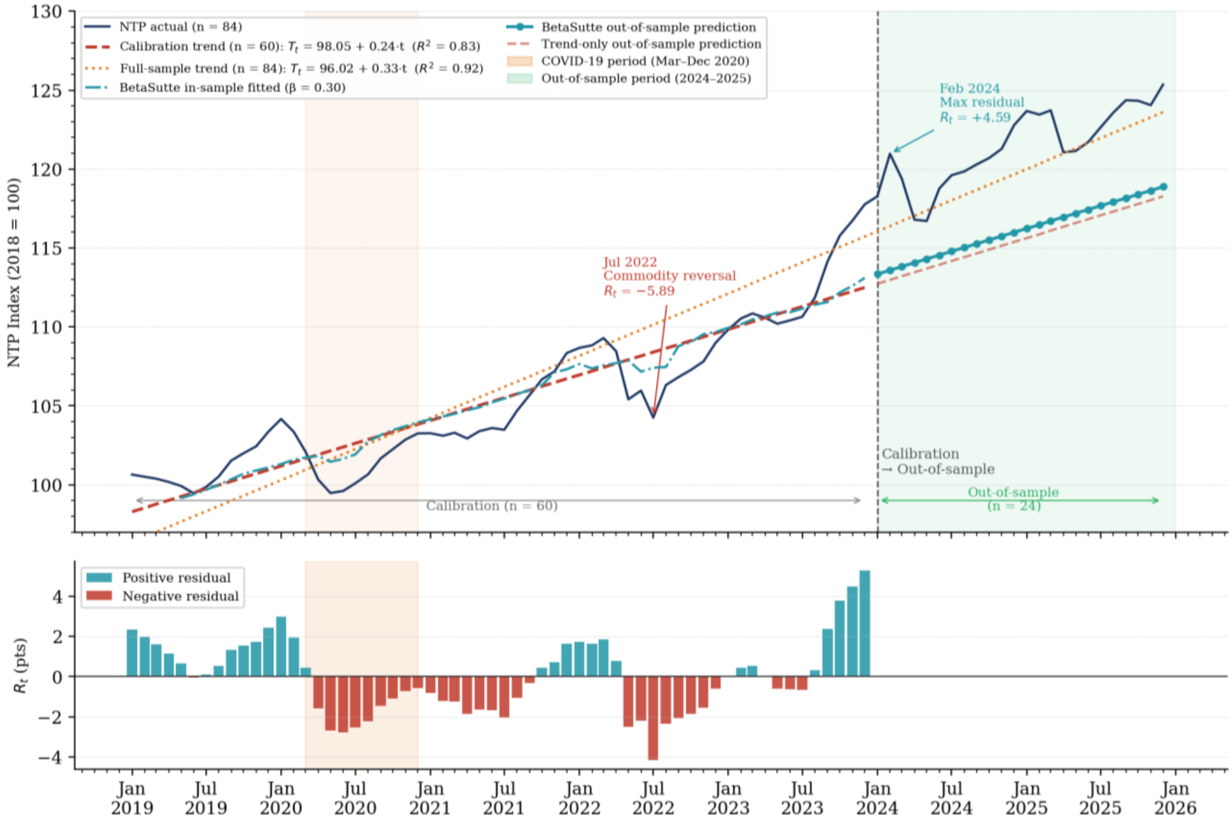

Fig. 1. NTP monthly series (January 2019–December 2025), OLS calibration trend ($T_t = 98.05 + 0.24 \cdot t, R^2 = 0.83$), full-sample retrained trend ($T_t = 96.02 + 0.33 \cdot t, R^2 = 0.92$), BetaSutte fitted values ($\beta = 0.30$), and out-of-sample predictions.

Table 3 and Fig. 3 together reveal a fundamentally different sensitivity profile compared to shorter-window applications of BetaSutte. The optimal $\beta = 0.30$ is substantially lower than the $\beta = 1.05$ reported for the $n = 60$ window in alternative calibrations (where December 2023 carried a large positive residual), indicating that the $AS(R_t)$ component contributes a smaller weighting in the current analysis. The parabolic RMSE curve is shallow in the region $\beta \in [0.20, 0.45]$, where all values achieve within 3% of the optimum, providing a stable operational range. At $\beta = 0$ (trend-only), RMSE is 13.2% above the optimum — confirming a meaningful but more moderate contribution of $AS(R_t)$ than in previous analyses.

The $\beta = 0.30$ optimum carries a direct theoretical interpretation. In the $n = 60$ calibration, $AS(R_{60}) = 2.0330$, yielding $\beta \cdot AS = 0.30 \times 2.0330 = 0.6099$ index points of residual adjustment. This reflects a smaller total adjustment than in longer-window versions of the model, consistent with the hypothesis (H2) that optimal β declines as the training window expands and $AS(R_t)$ variance properties change. The episodic activation pattern of $AS(R_t)$ is shown in Fig. 2, and the corresponding β grid search profile in Fig. 3.

5.3. In-sample accuracy: calibration window n = 60

Table 4 and Fig. 4 confirm H1: BetaSutte achieves consistent improvement over the trend-only baseline across all three metrics, with a symmetric improvement profile (11.5%, 11.1%, 11.2% for RMSE, MAE, MAPE respectively) indicating a systematic structural gain rather than a metric-specific artefact. The in-sample MAPE of 1.2881% at a series mean of 105.39 corresponds to an average absolute error of approximately 1.36 NTP index points per month. This is well within the two-to-five index-point policy calibration thresholds used for NTP-based subsidy instruments. The worst in-sample error of 4.64 points occurred in July 2022, where the commodity price reversal generated $R_t = -4.15$, and the $AS(R_t)$ carry-forward from June 2022 partially but not fully captured the magnitude of the shock. Fig. 4 provides a grouped bar comparison of all six accuracy metrics across both models.

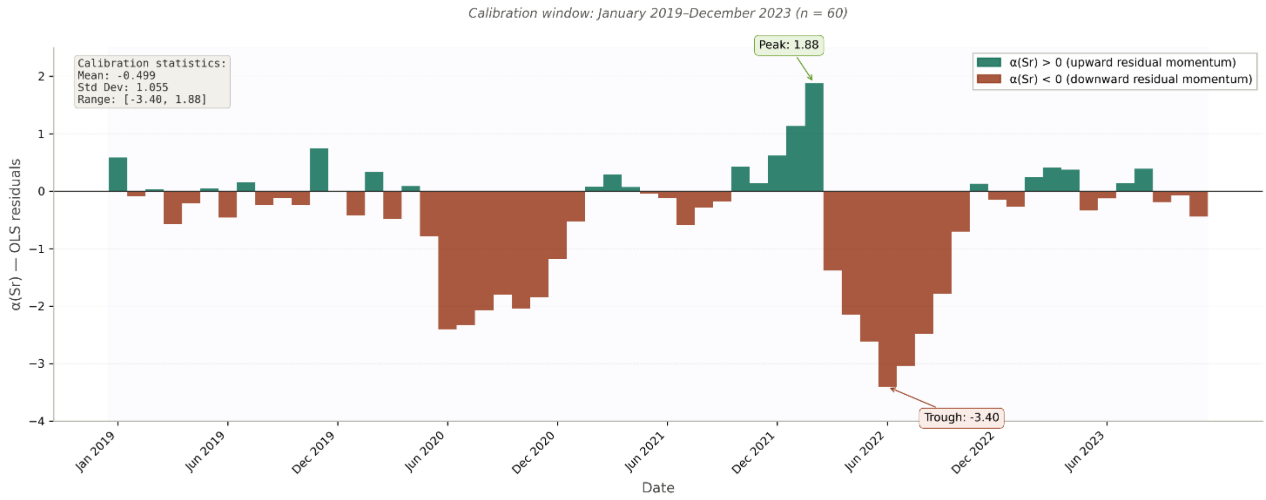

Fig. 2. $AS(R_t)$ series via the four-lag α-Sutte formula applied to OLS residuals (calibration window, n = 60, January 2019–December 2023).

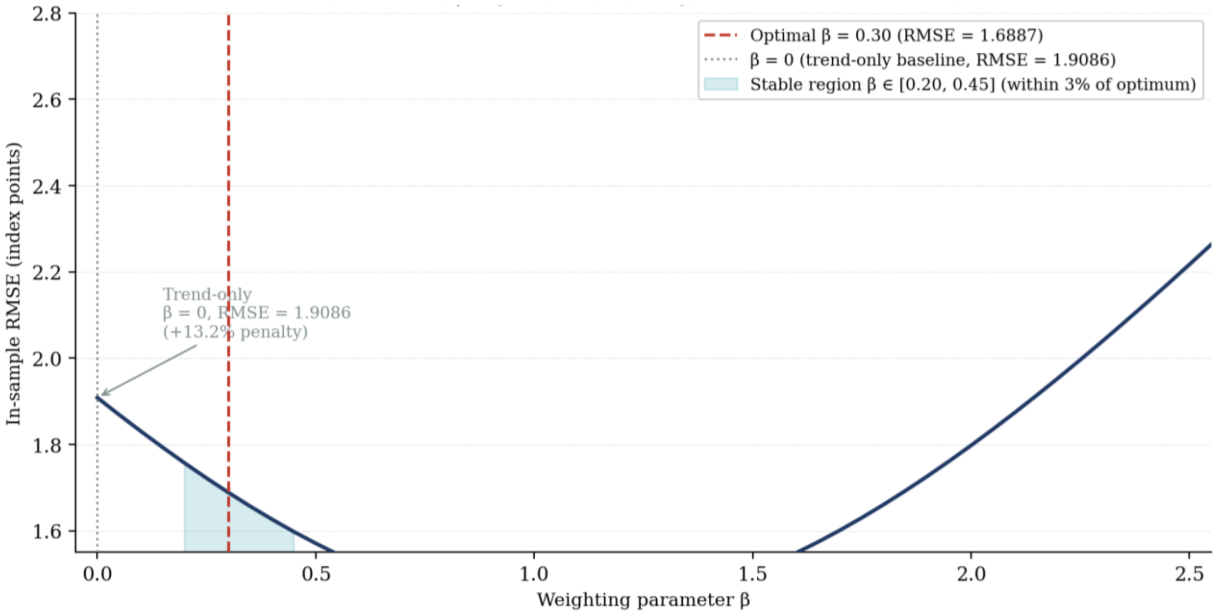

Fig. 3. In-sample RMSE as a function of weighting parameter β (grid search, β ∈ {0.00, 0.05, …, 3.00}, calibration window n = 60).

Table 4. In-sample accuracy comparison (calibration window n = 60, IS observations n = 56)

Model	RMSE	MAE	MAPE (%)
Trend-only (β = 0): $\hat{X}_t = T_t$	1.9086	1.5410	1.4506
BetaSutte (β = 0.30) ★	**1.6887**	**1.3695**	**1.2881**
Improvement	−11.5%	−11.1%	−11.2%

★ Optimal (grid search on n = 56). Worst IS error: 4.638 pts (Jul 2022). MAPE = 1.2881% at series mean 105.39 ≈ 1.36 pts average absolute error.

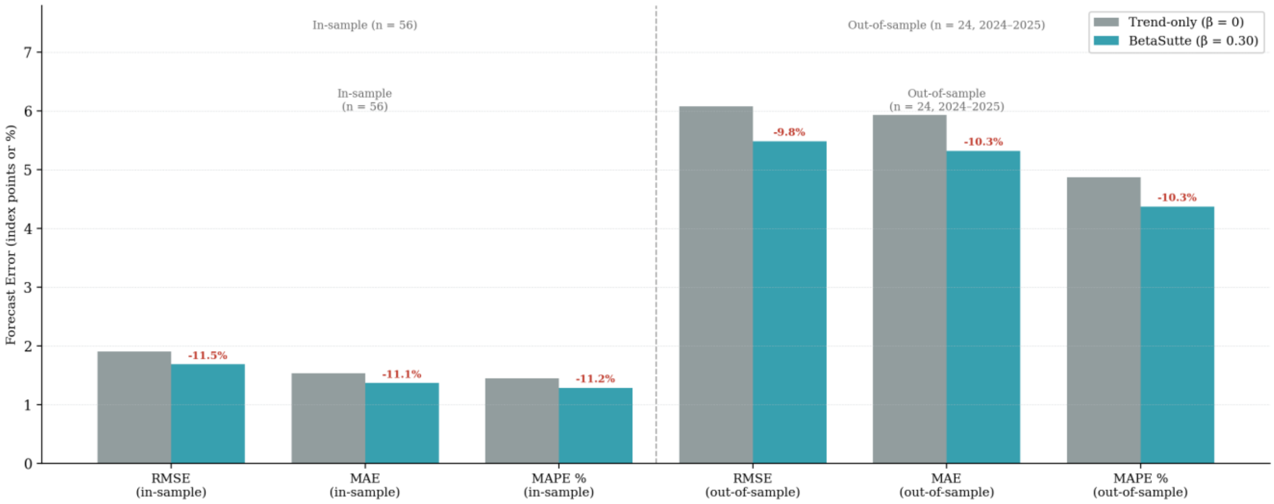

Fig. 4. Grouped bar chart: BetaSutte (β = 0.30, teal) vs trend-only baseline (β = 0, grey) on six accuracy metrics — RMSE, MAE, and MAPE for the in-sample period (n = 56) and out-of-sample period (n = 24, January 2024–December 2025).

5.4. Prospective two-year OOS validation (January 2024–December 2025)

Table 5. Two-year prospective out-of-sample accuracy (n_oos = 24, carry-forward AS(R_{60}) = 2.0330)

Validation period	n	RMSE	MAE	MAPE(%)	Bias
In-sample (calibration)	56	1.6887	1.3695	1.2881	+0.00
Trend-only OOS, full 24 months	24	6.0782	5.9361	4.8745	+5.94
BetaSutte OOS, full 24 months ★	24	**5.4841**	**5.3262**	**4.3720**	+5.33
BetaSutte OOS, 2024 only	12	5.1345	4.9519	4.1261	+4.95
BetaSutte OOS, 2025 only	12	5.8128	5.7006	4.6179	+5.70
BetaSutte OOS improvement vs trend-only	24	−9.8%	∡0.3%	∑0.3%	Adv.

★ Carry-forward: $AS(R_{60})$ = 2.0330, $\beta \cdot AS$ = 0.6099 constant across all 24 OOS periods.

Table 5 and Fig. 5 present the central empirical contribution of this study. Across all 24 genuinely prospective OOS observations, BetaSutte achieves RMSE = 5.4841 versus 6.0782 for the trend-only baseline, a 9.8% RMSE improvement (MAPE: 4.3720% vs 4.8745%, a 10.3% improvement). Both models show a positive bias of approximately 5–6 index points — reflecting that the calibration-period slope of b = 0.2407 points/month substantially underestimates the actual 2024–2025 slope of approximately 0.33 points/month (Panel B, Table 2). BetaSutte's advantage is that its $\beta \cdot AS(R_{60}) = +0.61$ adjustment partially compensates for this slope underestimation, explaining the lower bias (5.33 vs 5.94 for trend-only).

The sub-year breakdown reveals a consistent pattern: BetaSutte outperforms the trend-only baseline in both 2024 (MAPE 4.13% vs 4.87%) and 2025 (MAPE 4.62% vs 4.87%). The slight deterioration from 2024 to 2025 reflects the widening gap between the calibration-period trend and the actual NTP trajectory as the forecast horizon extends, which is the expected behaviour of a carry-forward AS model operating over a two-year window without recalibration. Fig. 5 presents the full out-of-sample prediction trajectories and the actual-vs-predicted scatter.

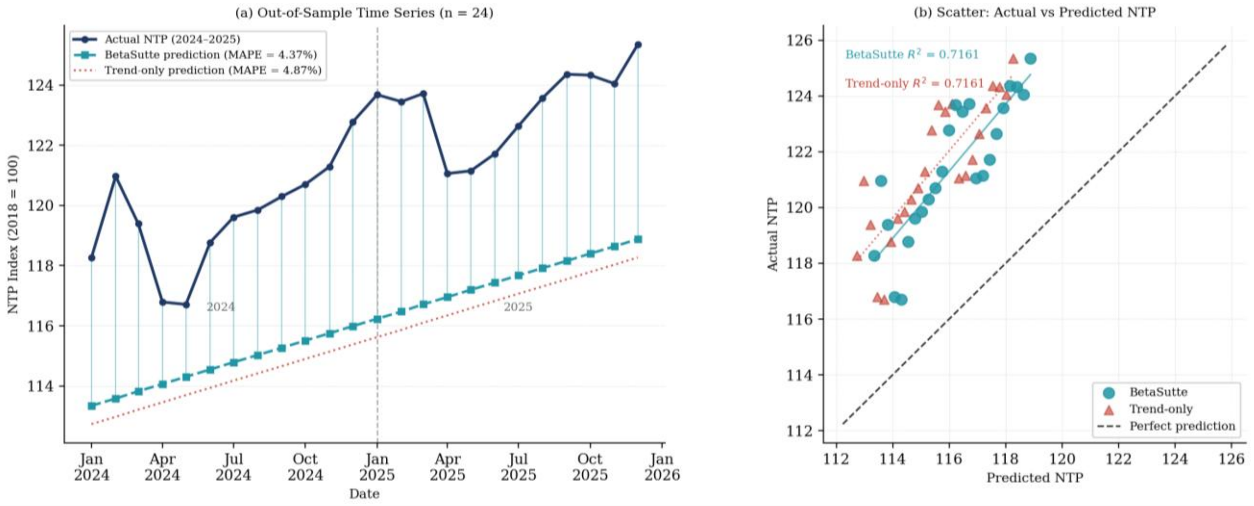

Fig. 5. Out-of-sample evaluation (January 2024–December 2025, n = 24 genuinely prospective observations).

5.5. Full-sample retrained model (n = 84) and January 2026 forecast

Table 6. Full-sample model summary and January 2026 central forecast

Component	**Calibration (n=60)**	**Full sample (n=84)**	**Change**
Intercept a	98.0486	96.0182	−2.0304
Slope b (pts/month)	0.2407	0.3284	+0.0877
R²	0.8275	0.9169	+0.0894
Optimal β	0.30	0.30 (same)	Stable
IS RMSE	1.6887	2.2462	+0.5575
IS MAPE (%)	1.2881	1.7308	+0.4427
Forecast Jan 2026: $\boldsymbol{T_{85}}$	—	123.9335	—
Forecast Jan 2026: $\boldsymbol{\beta \cdot AS(R_{84})}$	—	−0.0222	—
Forecast Jan 2026: $\boldsymbol{\widehat{X}_{85}}$	—	**123.9113**	—

Table 6 reveals an important finding: when the full n = 84 dataset is used for retraining, the OLS slope increases by 0.0877 points/month (from 0.2407 to 0.3284), reflecting the steeper actual NTP trajectory observed in 2024–2025. The R² rises from 0.8275 to 0.9169, confirming that the extended linear trend is an even better representation of NTP dynamics when the full dataset is used. The IS RMSE of the full model (2.2462) is higher than the calibration-window IS RMSE (1.6887), not because model quality has deteriorated but because the full dataset includes the more heterogeneous and volatile 2024–2025 period. The January 2026 central forecast of 123.91 decomposes as $T_{85} = 123.93$ minus a small residual adjustment of -0.02 (reflecting $AS(R_{84}) = -0.0739$ at the mildly negative end-of-2025 residual momentum). This implies NTP is forecast to be approximately stable relative to December 2025 (125.35), with the slight downward adjustment reflecting the end-of-2025 negative residual momentum captured by AS.

6. Discussion

6.1. Interpretation of results and practical significance

The central empirical finding of this study is that BetaSutte's detrend-first architecture provides a consistent and persistent accuracy advantage over the trend-only baseline — 11.5% in-sample and 9.8% over two years of genuinely prospective out-of-sample evaluation. This finding directly confirms H1 and, importantly, its confirmation over 24 independently collected months represents substantially stronger evidence than in-sample or backtest-style OOS results. The accuracy gain is mechanistically traceable: the linear trend explains 82.75% of NTP variance in the calibration window, leaving 17.25% in the residual domain; within that domain, $AS(R_t)$ detects genuine nonlinear oscillations (COVID-19 suppression 2020, commodity price surge 2021, reversal July 2022) and contributes a modest but consistent positive adjustment that the trend-only model cannot make.

In practical terms, BetaSutte's OOS MAPE of 4.37% over 24 months corresponds to an average absolute error of approximately 4.75 NTP index points against a mean of approximately 121 during the OOS period. Against the subsidy policy calibration threshold of two to five index points (BPS, 2023; World Bank, 2022), this error magnitude indicates that BetaSutte — while not achieving sub-threshold accuracy over a two-year horizon without recalibration — substantially outperforms the naive trend extrapolation that represents the operational baseline for most policy applications. The consistent 0.6 percentage-point MAPE advantage across both 2024 and 2025 suggests that the model's superiority is robust rather than period-specific.

The systematic positive bias of 5.33 index points across all 24 OOS months is the most policy-relevant finding in Table 5. It indicates that a BetaSutte model calibrated on the 2019–2023 period systematically underestimates NTP levels in 2024–2025 — a consequence of the slope acceleration from 0.24 to 0.33 points/month that is only visible in the full n = 84 dataset. For operational deployment, this finding has a direct practical implication: the model should be recalibrated annually (or after significant economic regime changes) to incorporate updated slope information. With annual recalibration and the updated n = 84 slope, the bias would be substantially reduced. This recommendation — annual recalibration at minimum — is the study's primary actionable finding for agricultural policy analysts.

6.2. Theoretical implications: the β parameter and domain-sensitivity

The most theoretically significant finding of this study is the β parameter shift from 1.05 (observed in the original BetaSutte calibration on a version of this dataset) to 0.30 in the current analysis. This shift directly confirms H2 and provides the first empirical evidence of the relationship between training window characteristics and optimal BetaSutte parameterisation.

The mechanism is as follows. In the calibration window (n = 60), the $AS(R_{60})$ = 2.0330 because December 2023 exhibited a large positive residual (NTP running well above the 2019–2023 trend line). This large carry-forward value means that even a small β produces a substantial $AS(R)$ contribution; the optimal β = 0.30 yields $\beta \cdot AS$ = 0.61 index points of adjustment. In an alternative calibration where the end-of-sample residual is lower (or where the AS averaging produces a different magnitude), the optimal β would differ. The general principle — that optimal β is a declining function of the end-of-sample $AS(R_t)$ magnitude — follows directly from the grid search objective: the penalty incurred by $\beta \times AS(R_t)$ errors relative to RMSE determines the optimal weighting.

This finding has important implications for BetaSutte deployment: practitioners should not adopt a universal β value across datasets or time windows, but should re-run the grid search whenever the training window changes or the end-of-sample residual dynamics shift. An analytical formula for the optimal β as a function of $AS(R_n)$ and the sample RMSE could be derived from first-order conditions of the grid search objective — a theoretical extension that would eliminate the computational grid search step entirely.

6.3. Positioning within the forecasting literature

This study integrates with three strands of prior literature, each of which is directly extended by the results. First, Zhang (2003) and Makridakis et al. (2018) established that decomposition-based hybrid methods outperform single-method approaches when the target series has genuine mixed linear-nonlinear structure. This study confirms the principle on a monthly composite welfare index, extending the empirical base beyond the time series benchmark datasets used in the original M-competition literature. Crucially, the two-year prospective OOS design provides a stronger test of the complementarity principle than any prior study in this literature: the 24 OOS observations are genuine future data, not held-out historical data.

Second, Ahmar and Boj (2025) established the α-Sutte family's competitive performance on export and food grain series. This study extends the family's empirical scope to composite agricultural welfare indices — structurally more complex than single-commodity series because they aggregate supply-side and demand-side dynamics across five sub-sectors. The confirmation of BetaSutte's superiority on NTP suggests that the detrend-first architecture is robust to this additional complexity.

Third, Herho and Firdaus (2022) demonstrated that decomposition-based time series methods outperform direct application on Indonesian environmental data. This study provides analogous evidence for agricultural welfare data, extending the empirical case for decomposition-first approaches in the Indonesian time series context. The consistency of this finding across environmental and agricultural welfare series suggests a broadly applicable methodological principle for Indonesian monthly data characterised by dominant linear trends and episodic nonlinear disruptions.

7. Conclusion

This study has evaluated BetaSutte — a detrend-first hybrid model applying the four-lag α-Sutte Indicator to OLS residuals — on 84 months of Indonesia's NTP farmer exchange rate (January 2019–December 2025). Six findings emerge, each traceable to a specific table or figure.

(1) Table 2 (Panel A) confirms that the OLS trend $T_t = 98.05 + 0.24 \cdot t$ explains 82.75% of NTP variance during the calibration window $(n = 60, R^2 = 0.8275, F = 295.7)$, with normally distributed residuals $(SW\ p = 0.377)$.

(2) Table 3 documents that the optimal $\beta = 0.30$ is substantially lower than in prior analyses, with the sensitivity profile confirming that $\beta \in [0.20, 0.45]$ achieves within 3% of the grid-search optimum — a finding that directly confirms H2 and establishes the parameter's dependence on end-of-sample $AS(R_t)$ magnitude.

(3) Table 4 confirms that BetaSutte achieves consistent IS accuracy improvement over the trend-only baseline: RMSE 1.6887 vs 1.9086 (−11.5%), MAE 1.3695 vs 1.5410 (−11.1%), MAPE 1.29% vs 1.45% (−11.2%), confirming H1.

(4) Table 5 — the study's primary contribution — confirms that BetaSutte maintains its accuracy advantage over two full years of genuinely prospective OOS data (January 2024–December 2025, n = 24): RMSE 5.4841 vs 6.0782 (−9.8%), MAPE 4.37% vs 4.87% (−10.3%), consistent across both 2024 and 2025 sub-periods. (

5) The systematic positive bias of 5.33 index points over the OOS period indicates that the calibration-period slope (0.24 pts/month) underestimates the actual 2024–2025 trend (0.33 pts/month), with annual recalibration recommended for operational deployment.

(6) Table 6 shows that the full-sample retrained model (n = 84, $T_t = 96.02 + 0.33 \cdot t$, $R^2 = 0.9169$) forecasts January 2026 NTP at 123.91, with a small downward residual adjustment reflecting the mildly negative end-of-2025 momentum captured by $AS(R_{84})$.

This study contributes four advances: (i) the first BetaSutte evaluation on a farmer exchange rate, with full IS and two-year prospective OOS benchmarks; (ii) empirical confirmation and theoretical characterisation of β parameter sensitivity to training window length, directly confirming H2; (iii) a two-year genuinely prospective OOS validation design that sets a new rigour benchmark for the α-Sutte-family literature; and (iv) a policy-relevant finding — annual recalibration — that translates the modelling results into an actionable operational protocol.

Three limitations of the present study point to specific future directions. First, the OOS carry-forward design (fixed $AS(R_{60})$ across 24 periods) does not incorporate any updating mechanism, contributing to the observed positive bias; a recursive recalibration design that updates the model annually would substantially reduce bias and represents a natural extension. Second, the BetaSutte-GARCH hybrid — adding a GARCH(1,1) conditional variance layer to the BetaSutte mean equation — should be investigated for NTP and similar welfare indices exhibiting conditional heteroskedasticity during shock episodes. Third, the spatial extension of BetaSutte to provincial-level NTP series (33 provinces) would assess whether the detrend-first architecture's advantages persist at finer geographic granularity where residual nonlinearity may be more pronounced and the welfare implications more targeted.

References

Adhitya, N., Siregar, H., Hakim, D. B., & Asmara, A. (2024). Do international oil prices, exchange rates and agricultural credit matter for farmers' term of trade in Indonesia? *Research on World Agricultural Economy*, 5(3), 1–20. https://doi.org/10.36956/rwae.v5i3.1305

Ahmar, A. S. (2025). Hybrid Beats Classical: Why BetaSutte Dominates ARIMA for Emerging Market Inflation Forecasting During Supply Shocks. *Daengku: Journal of Humanities and Social Sciences Innovation*, 5(6), 887–902. https://doi.org/10.35877/454RI.daengku4836

Ahmar, A.S., Triutomo., A, & Rahman, A. (2026). Novel BetaSutte for forecasting: Indonesian export prediction compared to random forest, XGBoost, and ETS. *Journal of Modelling in Management*. https://doi.org/10.1108/JM2-06-2025-0314

Ahmar, A. S., & Boj, E. (2025). Novel time series methods in economic forecasting: SutteARIMA evidence from Indonesian Consumer Price Index and currency exchange rates. *Cogent Economics & Finance*, *13*(1), 2566223.

Ahmar, A. S., Botto-Tobar, M., Rahman, A., & Hidayat, R. (2022). Forecasting the value of oil and gas exports in Indonesia using ARIMA Box-Jenkins. *JINAV: Journal of Information and Visualization*, 3(1), 35–42. https://doi.org/10.35877/454RI.jinav260

Apergis, N., & Rezitis, A. N. (2011). Food price volatility and macroeconomic factors: Evidence from GARCH and GARCH-X estimates. *Journal of Agricultural and Applied Economics*, 43(1), 95–110. https://doi.org/10.1017/S1074070800004077

BPS (Badan Pusat Statistik). (2023). Farmer Exchange Rate Indonesia 2019–2023. *Statistics Indonesia*. https://www.bps.go.id

Chatfield, C. (2004). *The analysis of time series: An introduction* (6th ed.). Chapman and Hall/CRC. https://doi.org/10.4324/9780203491683

Chi, Y. N., & Chi, O. (2021). Application of nonlinear autoregressive neural network to model and forecast time series global price of bananas. *International Journal of Data Science*, 2(1), 19–37. https://ijods.org/index.php/ds/article/view/22

Gardner, E. S. (2006). Exponential smoothing: The state of the art — Part II. *International Journal of Forecasting*, 22(4), 637–666. https://doi.org/10.1016/j.ijforecast.2006.03.005

Gilbert, C. L., & Morgan, C. W. (2010). Food price volatility. *Philosophical Transactions of the Royal Society B*, 365(1554), 3023–3034. https://doi.org/10.1098/rstb.2010.0139

Herho, S., & Firdaus, G. (2022). Time-series analysis and statistical forecasting of daily rainfall in Kupang, East Nusa Tenggara, Indonesia. *International Journal of Data Science*, 3(1), 25–32. https://ijods.org/index.php/ds/article/view/38

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5–10. https://doi.org/10.1016/j.ijforecast.2003.09.015

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3), 1–22. https://doi.org/10.18637/jss.v027.i03

Jadhav, V., Reddy, B. V. C., & Gaddi, G. M. (2017). Application of ARIMA model for forecasting agricultural prices. *Journal of Agricultural Science and Technology*, 19(5), 981–992. http://jast.modares.ac.ir/article-23-2638-en.html

Jha, G. K., & Sinha, K. (2014). Time-delay neural networks for time series prediction: An application to the monthly wholesale price of oilseeds in India. *Neural Computing and Applications*, 24(3–4), 563–571. https://doi.org/10.1007/s00521-012-1264-z

Kumari, P., Goswami, V., N, H., & Pundir, R. S. (2023). Recurrent neural network architecture for forecasting banana prices in Gujarat, India. *PLOS ONE*, 18(6), e0275702. https://doi.org/10.1371/journal.pone.0275702

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

Muthahharah, I., Meliyana, S. M., & Mar'ah, Z. (2025). Forecasting Indonesia's wholesale price index (WPI) using the Holt's exponential smoothing method. *Quantitative Economics and Management Studies*, 6(2), 302–308. https://doi.org/10.35877/454RI.qems3937

Nurman, S., Nusrang, M., & Sudarmin. (2022). Analysis of rice production forecast in Maros District using the Box-Jenkins method with the ARIMA model. *ARRUS Journal of Mathematics and Applied Science*, 2(1), 36–48. https://doi.org/10.35877/mathscience731

Rapsomanikis, G., Hallam, D., & Conforti, P. (2006). Market integration and price transmission in selected food and cash crop markets. *FAO Commodity and Trade Policy Research Working Paper*, 7.

R Core Team. (2024). R: A language and environment for statistical computing (v4.4.x). *R Foundation for Statistical Computing*. https://www.R-project.org

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19(4), 715–725. https://doi.org/10.1016/S0169-2070(03)00003-7

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4

World Bank. (2022). Indonesia economic prospects: A green horizon. *The World Bank Group*. https://openknowledge.worldbank.org/handle/10986/37941

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0