

Forecasting Global Food Price Pressure: The Predictive Limits of Upstream Indicators against an Autoregressive Benchmark

Zhang Juan^{1,*}, Lin Yihuan², Shen Xin³

¹ Institute for Mathematical Research (INSPeM), Universiti Putra Malaysia, (Malaysia)

² School of Business and Economics, Universiti Putra Malaysia, (Malaysia)

³ School of Business and Economics, Universiti Putra Malaysia, (Malaysia)

Received: 09/06/2026

Accepted: 19/06/2026

Published: 30/06/2026

Summary: This paper addresses a central tension in global food price forecasting: upstream cost and risk indicators are structurally important for explaining food-price pressure, yet their real-time forecasting value is far less secure. The evidence favours a parsimonious benchmark. Across the out-of-sample evaluation, the univariate AR(1) model delivers the lowest overall forecast risk, while Ridge and XGBoost remain statistically indistinguishable from it under modified Diebold-Mariano tests. More saturated specifications, particularly LASSO and ARDL-Full, lose forecasting power, indicating that additional lagged predictors can introduce estimation noise faster than usable signal in monthly food-price forecasting. Regime-specific results qualify but do not overturn this conclusion. XGBoost records a localized advantage during the Russia-Ukraine commodity shock, consistent with the ability of tree-based models to adapt to nonlinear cross-market stress, but this gain does not persist across the full sample or at longer forecast horizons. The findings separate diagnostic relevance from point-forecasting capacity. Energy, fertiliser, exchange-rate, supply-chain and geopolitical indicators remain valuable for interpreting global food-price pressure, especially during crisis episodes; however, their incremental one-month-ahead forecasting contribution is uneven. For real-time monitoring systems, richer information sets and machine-learning flexibility should be disciplined by transparent autoregressive benchmarks rather than accepted on the basis of economic plausibility alone.

Keywords: Food price forecasting; AR(1) benchmark; Upstream indicators; XGBoost; Forecasting parsimony

* Corresponding author, e-mail: gs66172@student.upm.edu.my.

I- Introduction :

Global food price forecasting has become more difficult as energy, fertiliser, exchange-rate, supply-chain and geopolitical shocks increasingly overlap. The post-COVID rebound and the Russia-Ukraine shock showed that food price pressure is no longer driven by a single commodity channel. Recent evidence indicates that higher energy and fertiliser costs can exert stronger pressure on food prices than export curtailments alone during conflict periods (Alexander et al., 2023). This strengthens the case for treating upstream cost indicators as potential early-warning variables rather than as background macroeconomic controls.

The cost-push mechanism is well documented. Energy prices enter agricultural production through fuel, machinery, irrigation, storage and transport, while natural gas affects nitrogen fertiliser costs and, indirectly, crop production costs. Earlier work linked crude oil prices to global grain prices (Chen et al., 2010), and more recent studies emphasise nonlinear or asymmetric energy-food price transmission (Chowdhury et al., 2021; Alnour et al., 2023). Updated evidence also points to crude oil, natural gas and fertiliser as connected channels in agricultural price dynamics (Bor & Sakarya, 2026; Figiel et al., 2026).

Yet economic relevance is not the same as forecast usefulness. Many studies explain food price movements using oil prices, fertiliser prices, raw material costs, exchange rates or crisis indicators, but fewer ask whether these variables improve real-time forecasts once only lagged information is allowed. Machine learning studies report that global cost indicators can help model world food prices (Ulussever et al., 2023; Han et al., 2023), while crisis-period forecasting work suggests that predictive content may vary across regimes (Toledo & Duncan, 2024). The unresolved issue is whether these gains survive a direct comparison with a simple autoregressive benchmark.

This paper tests that issue using monthly data from 2000M1 to 2026M5. The forecasting target is the monthly log change in the FAO Food Price Index. The predictor set includes Brent crude oil, natural gas, the World Bank Fertilizer Price Index, the US Dollar Index, the Global Supply Chain Pressure Index and the Geopolitical Risk Index. FAO food sub-indices are excluded because they are components of the aggregate index and would create a mechanical forecasting exercise rather than a genuine test of upstream information.

The study is guided by three research questions:

RQ1: Do upstream cost and risk indicators improve one-month-ahead forecasts of global food price pressure relative to an AR(1) benchmark?

RQ2: Do regularised and machine learning models outperform parsimonious time-series models in a short monthly food-price series?

RQ3: Does predictive performance change across crisis and normalisation periods, particularly during the COVID-19 disruption and the Russia-Ukraine commodity shock?

The empirical design uses an expanding-window out-of-sample evaluation over 2019M1–2026M5. It compares parsimonious benchmarks, autoregressive distributed lag models, regularised regressions and XGBoost, with forecast differences assessed through standard accuracy measures and modified Diebold-Mariano tests. Sub-period and forecast-horizon checks are used to examine whether the results are sensitive to crisis regimes or prediction length.

The contribution is practical. Recent forecasting studies have increasingly turned to flexible machine learning models and richer global predictor sets (Ulussever et al., 2023; Zelingher & Makowski, 2024). This paper asks whether that added complexity clears a transparent benchmark in a monthly global food-price setting. For food-price monitoring systems, the point is not whether upstream indicators are economically meaningful, but whether they provide reliable incremental forecasting value once the autoregressive structure of food prices is accounted for.

The remainder of the paper is organised as follows. Section II describes the data and methodology. Section III reports the empirical results and robustness checks. Section IV concludes.

II-Literature Review :

The literature on food price dynamics offers a strong economic case for upstream indicators, but a weaker case for their real-time forecasting superiority. Energy, fertiliser, exchange rates and logistics costs are repeatedly shown to shape food price movements, especially during global disruption episodes. The 2007–2008 food price crisis already linked oil prices, biofuel demand, exchange-rate movements and input costs to international food price pressure (Abbott et al., 2008; Piesse & Thirtle, 2009). Later work placed these forces in a broader commodity-price framework, showing that oil prices and macroeconomic conditions remain central to global food price variation (Baffes & Dennis, 2013). This literature explains why upstream variables belong in food-price analysis; it does not, by itself, prove that they improve short-horizon forecasts.

The cost-push channel has become more visible after the Russia-Ukraine war. Rising fuel and fertiliser prices intensified food-security risks in import-dependent economies, often through indirect production and trade-cost channels rather than export disruption alone (Carriquiry et al., 2022; Arndt et al., 2023; Lin et al., 2023). Recent commodity-market assessments continue to treat energy and fertiliser prices as major sources of pressure for agricultural markets and food affordability (World Bank, 2026b). This newer evidence strengthens the case for monitoring upstream variables, but it also makes the forecasting problem more demanding: when several shocks move together, separating useful signal from crisis-period co-movement becomes harder.

Crucially, cost transmission is seldom linear, instantaneous or stable. Price-transmission research has long shown that upstream and downstream prices may adjust asymmetrically, especially when positive and negative shocks pass through at different speeds (Frey & Manera, 2007). In energy-food markets, this instability is even sharper. Energy shocks may operate through direct fuel and transport costs, indirect fertiliser costs, and broader volatility channels, with the strength of transmission changing across periods and commodities (Taghizadeh-Hesary et al., 2019; Vatsa & Miljkovic, 2022). Recent studies using nonlinear and dynamic frameworks also suggest that oil, gas

and fertiliser links to agricultural prices are state-dependent rather than fixed over time (Bor & Sakarya, 2026; Figiel et al., 2026). For forecasting, this matters because a variable that is significant in a full-sample explanatory model may still fail when used as a lagged predictor in real time.

The growing use of machine learning in food and commodity forecasting responds to this instability. Flexible algorithms promise to capture nonlinear relationships, interactions and crisis-specific patterns that linear models may miss. Recent applications use tree-based models, support vector methods and hybrid frameworks to forecast food prices, freight-related inflation or agricultural commodity shocks (Charles, 2023; Derindag et al., 2023; Ogol, 2025). Yet the evidence remains uneven. Forecast gains often depend on the commodity, horizon, country and evaluation design. Recent applications remain cautious on the same point: complex models may perform well when the signal is rich, but short agricultural price series remain vulnerable to overfitting, regime instability and benchmark failure.

This benchmark problem is not new. Commodity prices are forward-looking, noisy and exposed to abrupt structural breaks. In such environments, additional predictors can increase estimation variance faster than they improve signal extraction. Oil-price forecasting research has repeatedly shown that richer models often struggle to outperform simple random-walk or autoregressive alternatives in real-time evaluation (Baumeister & Kilian, 2012; Alquist et al., 2013). Broader forecasting evidence reaches a similar conclusion: algorithmic complexity is not a substitute for stable predictive information, especially when the test horizon is short and the sample is limited (Makridakis et al., 2018; Petropoulos et al., 2022).

This study therefore treats upstream indicators as candidate signals, not guaranteed predictors. The literature provides strong reasons to include energy, fertiliser, exchange-rate, supply-chain and geopolitical variables. It also provides strong reasons to doubt that a richer information set or a more flexible algorithm will automatically beat a transparent autoregressive benchmark. The empirical question is whether these indicators deliver incremental out-of-sample accuracy for monthly global food price pressure once the persistence of the FAO Food Price Index itself is properly benchmarked.

III– Methods and Materials:

1.Data and variables

This study employs an ex-ante forecasting framework to evaluate the incremental predictive value of upstream cost and risk indicators for global food price pressure. The dependent variable is the monthly log change in the aggregate FAO Food Price Index:

$$\Delta \log(\text{FPI}_t) = \log(\text{FPI}_t) - \log(\text{FPI}_{t-1}). \quad (1)$$

The aggregate FAO index provides a monthly measure of international food market pressure across major traded food commodities (Food and Agriculture Organization of the United Nations [FAO], 2026). The raw dataset spans 2000M1–2026M5. After temporal differencing and lag

initialization, the effective modelling sample covers 2000M3–2026M5, yielding 315 monthly observations.

The conditioning information set includes six upstream predictors. Supply-side cost pressure is proxied by Brent crude oil prices, natural gas prices and the World Bank Fertilizer Price Index, all obtained from the World Bank Commodity Price Data (World Bank, 2026a). Macro-financial pricing conditions are captured by the US Dollar Index. Non-price market frictions are represented by the Global Supply Chain Pressure Index, which measures global logistics and delivery-time pressure (Benigno et al., 2022), and the newspaper-based Geopolitical Risk Index (Caldara & Iacoviello, 2022).

All nominal commodity and financial price series enter the specifications in monthly log differences, whereas GSCPI and GPR are retained in levels as index-based pressure measures. FAO food sub-indices are omitted from the predictor set to eliminate trivial accounting identities with the aggregate target and to preserve a genuine test of external predictive information. Table 1 presents the final modelling variables, descriptive statistics and contemporaneous correlations with ($\Delta \log(\text{FPI})$).

Table (1): Descriptive Statistics and Correlations

Variable	Mean	SD	Min	Max	Corr. with $\Delta \log(\text{FPI})$
$\Delta \log(\text{FAO FPI})$	0.003	0.025	-0.143	0.123	1.000
$\Delta \log(\text{Brent crude})$	0.004	0.097	-0.511	0.377	0.425
$\Delta \log(\text{Natural gas})$	0.000	0.162	-0.742	0.641	0.107
$\Delta \log(\text{Fertilizer index})$	0.006	0.067	-0.293	0.301	0.166
$\Delta \log(\text{DXY})$	-0.000	0.018	-0.051	0.060	-0.378
GSCPI	0.047	1.030	-1.579	4.449	0.135
GPR index	110.01	53.37	45.06	512.53	-0.007

Note: Sample period: 2000M3–2026M5. N = 315 for all variables. $\Delta \log$ denotes monthly log-difference. Correlation refers to Pearson contemporaneous correlation with $\Delta \log(\text{FAO FPI})$.

The stationarity diagnostics in Table 2 support this transformation framework. The traded price series satisfy the working stationarity requirement after log-differencing, while GSCPI and GPR are retained as level indices consistent with their role as external pressure and risk measures.

Table (2): Stationarity Diagnostics

Variable	ADF p-value	KPSS p-value	Decision
$\Delta \log(\text{FAO FPI})$	0.000	>0.10	Stationary after log-differencing
$\Delta \log(\text{Brent})$	0.000	>0.10	Stationary after log-differencing
$\Delta \log(\text{Natural gas})$	0.000	>0.10	Stationary after log-differencing
$\Delta \log(\text{Fertilizer})$	0.000	>0.10	Stationary after log-differencing
$\Delta \log(\text{DXY})$	0.000	>0.10	Stationary after log-differencing
GSCPI	0.006	<0.025	Level index retained
GPR index	0.000	<0.025	Level index retained

Note: ADF tests the null of a unit root; KPSS tests the null of stationarity. Price variables are used in monthly log differences. GSCPI and GPR are retained in levels as external pressure and risk indices.

2. Forecasting design

Forecasts are generated using a recursive expanding-window design. The initial training window covers 2000M3–2018M12, and the evaluation period runs from 2019M1 to 2026M5. At each forecast origin, parameters are estimated using only information available up to $t - 1$, and the one-month-ahead forecast is evaluated at t . The estimation window then expands by one observation. This design preserves real-time predictive validity and prevents the use of contemporaneous information.

The parsimonious benchmark is a univariate AR(1) model:

$$\Delta \log(\text{FPI}_t) = \alpha + \phi \Delta \log(\text{FPI}_{t-1}) + \varepsilon_t. \quad (2)$$

Multivariate extensions use the following lagged information structure:

$$\Delta \log(\text{FPI}_t) = \alpha + \sum_{p=1}^P \phi_p \Delta \log(\text{FPI}_{t-p}) + \sum_{j=1}^J \sum_{k=1}^K \beta_{j,k} X_{j,t-k} + \varepsilon_t, \quad (3)$$

where $X_{j,t-k}$ denotes the j -th upstream indicator at lag k . The ARDL-Compact specification uses $P = 1$ and $K = 1$, while ARDL-Full allows lags up to $K = 3$. Ridge and LASSO apply shrinkage to the same richer lag structure to reduce estimation variance in the monthly sample (Hoerl & Kennard, 1970; Tibshirani, 1996). XGBoost is included as a nonlinear benchmark capable of capturing regime-dependent interactions among lagged predictors (Chen & Guestrin, 2016). Its inclusion is not treated as a presumption of superior performance; it is evaluated under the same out-of-sample discipline as the linear specifications.

3. Model evaluation and robustness checks

Forecast accuracy is evaluated using RMSE, MAE, directional accuracy and Theil's (U), with RMSE treated as the primary criterion because large forecast errors are particularly consequential during food price spikes. Theil's (U) is benchmarked against AR(1), allowing each specification to be interpreted relative to the parsimonious baseline.

Pairwise predictive accuracy is assessed using the Diebold-Mariano test (Diebold & Mariano, 1995). Given the finite evaluation window of 89 monthly forecasts, the Harvey-Leybourne-Newbold small-sample correction is applied (Harvey et al., 1997). Tests are reported under both squared-error and absolute-error loss functions.

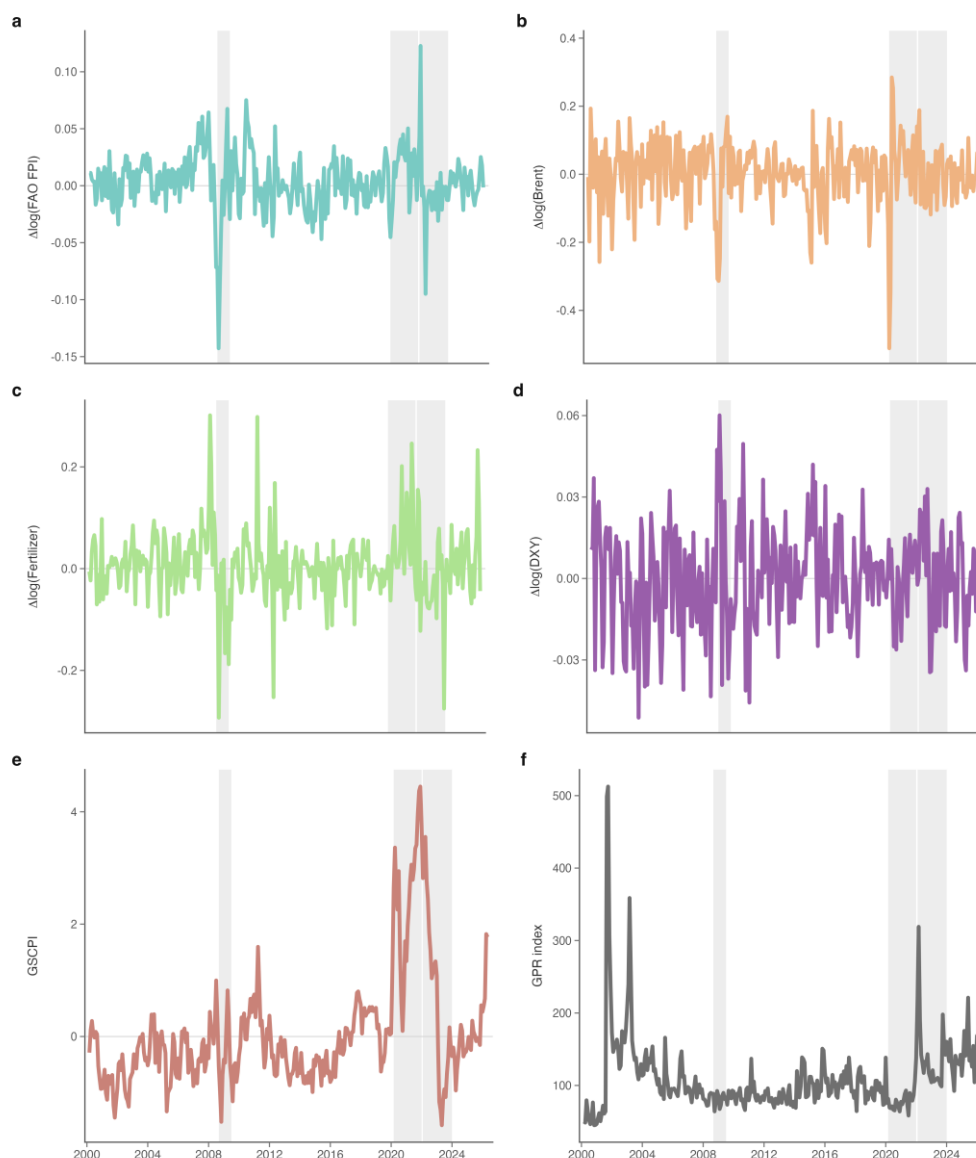
Two robustness checks complement the full-period comparison. First, the evaluation period is partitioned into four regimes: pre-COVID, COVID disruption, the Russia-Ukraine commodity shock and recent normalisation. This tests whether upstream indicators become more informative during crisis episodes. Second, selected models are re-estimated at a three-month forecast horizon to assess whether upstream lag information gains value beyond the one-month setting.

IV- Results and discussion :

1.Out-of-sample forecast accuracy

The historical trajectories in Figure 1 place global food price pressure within the broader movement of upstream macroeconomic indicators. Major shifts in the FAO Food Price Index coincide with energy, fertiliser, supply-chain and geopolitical disturbances, but the lead-lag structure across these series is fluid rather than regular. This temporal misalignment frames the central empirical question: whether structurally plausible upstream indicators can deliver persistent out-of-sample gains once the forecast is conditioned strictly on lagged information and benchmarked against AR(1).

Figure (1): Food Price Pressure and Upstream Indicators



Note: The figure plots monthly changes in the FAO Food Price Index and selected upstream indicators over 2000M3–2026M5. Shaded areas indicate major disruption periods.

The out-of-sample evaluation establishes a restrictive performance ceiling for high-information forecasting. The full-horizon accuracy metrics in Table 3 show that expanding the conditioning predictor set and introducing nonlinear flexibility fail to reduce forecast risk below the univariate benchmark. Over the 2019M1–2026M5 evaluation window, the parsimonious AR(1) specification remains the empirical accuracy frontier, while Ridge and XGBoost operate as near-benchmark competitors without delivering a systematic average reduction in forecast error.

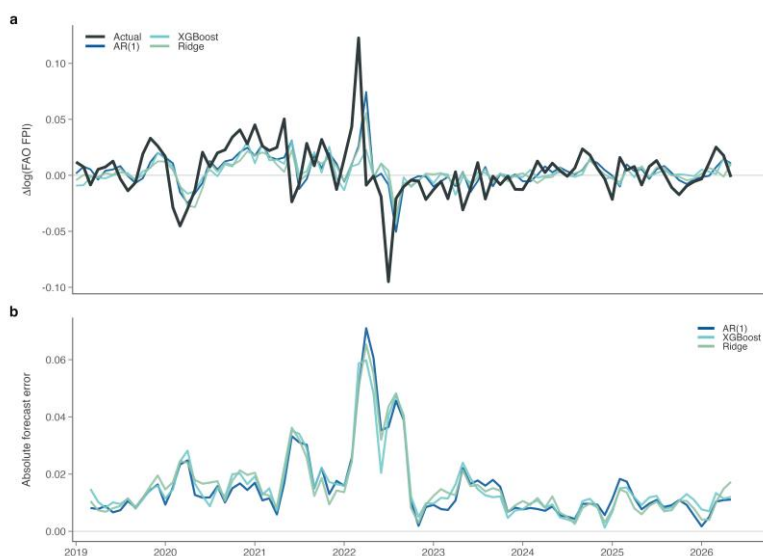
Table (3): Out-of-Sample Forecast Accuracy

Model	RMSE	MAE	Dir. acc. (%)	Theil's U
AR(1)	0.0230†	0.0150†	70.8†	1.000
XGBoost	0.0232	0.0157	61.8	1.010
Ridge	0.0234	0.0157	65.2	1.016
ARDL-Compact	0.0235	0.0158	66.3	1.023
LASSO	0.0239	0.0160	69.7	1.038
ARDL-Full	0.0247	0.0167	64.0	1.073
Naive Mean	0.0248	0.0174	52.8	1.079

Note: Expanding-window forecasts for 2019M1–2026M5 (N = 89). Theil's U = RMSE(model) / RMSE(AR(1)). † indicates the best value in the column.

This resistance to informational expansion presents a surface empirical paradox. The post-Ukraine and 2026 commodity cycles have been marked by closely connected disruptions across energy, fertiliser and agrifood markets, with recent assessments identifying upstream cost shocks as major drivers of agricultural price pressure (World Bank, 2026b; FAO, 2026). Recent ex-post studies similarly emphasise the role of energy, fertiliser and trade-cost channels in shaping food-market stress after 2022 (Alexander et al., 2023; Arndt et al., 2023; Lin et al., 2023). The evidence here does not reject these structural linkages. It shows that their informational content is not easily converted into one-month-ahead predictive gains once the aggregate FAO series is already conditioned on its own history.

Figure (2): Actual and Predicted Food Price Pressure



Note: Panel a compares actual $\Delta \log(\text{FAO FPI})$ with one-month-ahead forecasts from AR(1), XGBoost and Ridge over 2019M1–2026M5. Panel b reports rolling absolute forecast errors.

Figure 2 makes this tracking constraint visible. Across quieter market intervals, AR(1), Ridge and XGBoost forecasts move closely together. During acute price disturbances, all candidate models encounter tracking friction and fail to absorb the sharp discontinuities in the target series. Algorithmic flexibility preserves the competitiveness of XGBoost, but it does not create a persistent statistical separation from the autoregressive baseline.

2.Forecast comparison tests

The Diebold-Mariano evidence in Panel A of Table 4 moves the analysis beyond simple ordinal rankings. The statistically negligible point-estimate margins separating AR(1), Ridge and XGBoost confirm that the strongest multi-variable models function as near-benchmark alternatives rather than superior forecasting systems. Their numerical differences remain too weak to establish predictive dominance under either squared- or absolute-error loss.

Table (4): Forecast Robustness Evidence

Panel A. Modified Diebold-Mariano Tests against AR(1)				
Competing model	SE loss p-value	AE loss p-value	Interpretation	
XGBoost	0.898	0.401	No significant difference	
Ridge	0.711	0.231	No significant difference	
ARDL-Compact	0.254	0.035	AR(1) better under AE loss only	
LASSO	0.033	0.018	AR(1) significantly better	
ARDL-Full	0.041	0.003	AR(1) significantly better	
Naive Mean	0.421	0.060	No significant difference	
Panel B. Sub-period Forecast Performance				
Sub-period	N	Best model	AR(1) rank	Interpretation
Pre-COVID	14	ARDL-Compact	3	Differences negligible
COVID disruption	22	AR(1)	1	AR(1) strongest
Ukraine/energy shock	24	XGBoost	2	XGBoost marginally lower RMSE
Recent normalisation	29	AR(1)	1	AR(1) reasserts advantage

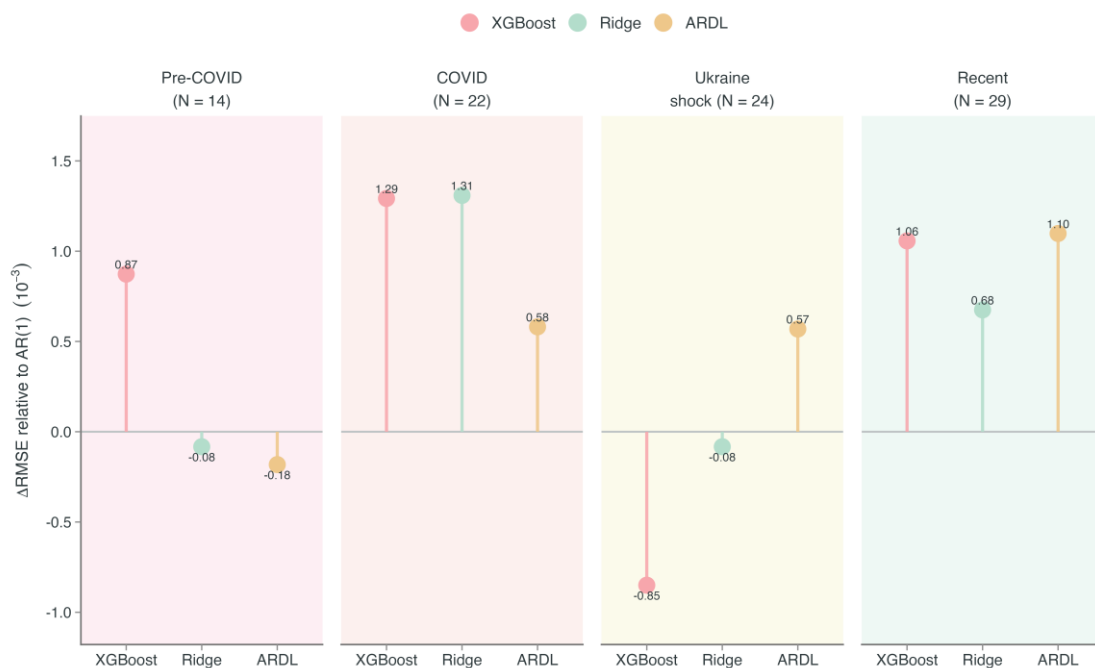
Note: Panel A reports Harvey–Leybourne–Newbold modified Diebold-Mariano tests against AR(1). SE = squared-error loss; AE = absolute-error loss. Panel B summarises sub-period RMSE rankings; sub-period evidence is descriptive because each regime contains a limited number of monthly forecasts.

The parameterisation penalty is clearer for over-extended structures. LASSO and ARDL-Full deteriorate relative to AR(1), implying that unconstrained lag expansion introduces estimation variance faster than usable signal in a constrained monthly sample. This outcome reinforces recent caution in agricultural forecasting: algorithmic sophistication yields predictive utility only when the underlying signal is sufficiently persistent and the validation protocol remains strictly out of sample (Ulussever et al., 2023; Zelingher & Makowski, 2024; Ogol et al., 2025). The intermediate performance of ARDL-Compact supports the same parsimony principle. Restricting the external conditioning layer to a first-order lag reduces the cost of over-parameterisation, but it remains insufficient to dislodge the univariate baseline.

3.Sub-period and horizon robustness

The regime-specific evidence in Panel B of Table 4 introduces localized nuance without undermining the full-sample inference. AR(1) retains its advantage across both the pre-COVID benchmark and the recent normalisation phase. The sole empirical departure emerges during the Russia-Ukraine commodity shock, where XGBoost records a marginal tracking advantage. Figure 3 confirms that this departure is concentrated in the shock period rather than distributed across the evaluation sample.

Figure (3): Sub-period Forecast Performance Relative to AR(1)



Note: The figure reports sub-period forecast performance relative to the AR(1) benchmark. Positive values indicate higher RMSE than AR(1), while negative values indicate lower RMSE.

This localized inversion is structurally coherent. The 2022–2023 episode operated as a compound structural break in which energy, fertiliser, logistics and trade-policy shocks interacted through nonlinear and regime-sensitive channels (Bor & Sakarya, 2026; Figiel et al., 2026). Under these distressed market conditions, the tree-based architecture displays localized adaptability to predictor interactions that rigid linear formulations cannot fully capture. The advantage remains state-dependent, however, and does not establish a durable machine-learning frontier across the broader sample.

Multi-step forward projections lead to the same boundary condition. Extending the forecast horizon to $h = 3$ does not provide external indicators with sufficient delayed-transmission space to overturn the benchmark. Instead, the signal-to-noise ratio weakens across multivariate specifications, and delayed informational recovery does not emerge. The robustness evidence therefore clarifies the

scope of the findings: upstream indicators offer diagnostic value under acute structural stress, but they do not function as persistent predictors across normal and crisis regimes.

4. Interpretation and implications

The results expose a persistent gap between structural relevance and ex-ante forecastability. Energy, fertiliser, exchange-rate, supply-chain and geopolitical indicators remain central to interpreting global food-price pressure. Their weak incremental forecasting performance reflects the temporal irregularity of pass-through rather than economic irrelevance. Upstream shocks do not map onto global food indices through stationary monthly lags; they are filtered through inventories, planting cycles, trade restrictions, subsidies and country-specific pricing frictions.

The machine-learning evidence belongs within this constraint. XGBoost is competitive and locally useful during the Ukraine shock, but it does not deliver statistically reliable full-sample gains. Nonlinear architectures should therefore be treated as state-dependent diagnostic tools rather than autonomous forecasting solutions. In short monthly series, computational flexibility cannot substitute for stable predictive information.

For real-time monitoring, the implication is straightforward. Upstream indicators should remain part of diagnostic early-warning systems, especially when energy, fertiliser, dollar and supply-chain signals move jointly. Their integration into point-forecasting models, however, should be audited against a transparent AR(1) benchmark. Without that benchmark, monitoring systems risk confusing historical explanatory fit with real-time predictive capacity.

The limitations follow from the same empirical logic. The aggregate FAO Food Price Index acts as a cross-category smoothing filter, which can dilute commodity-specific predictive signals. National food inflation dynamics may also diverge from the global index because domestic prices are mediated by currency regimes, subsidies, buffer stocks and trade policy. Finally, the crisis sub-periods are necessarily short. They are useful for detecting state-dependent anomalies, but not for confirmatory regime-specific inference.

Overall, upstream cost and risk indicators remain meaningful for interpreting global food-price pressure, but their incremental one-month-ahead forecasting value is uneven. In the monthly global setting examined here, AR(1) remains the benchmark that richer information sets must first overcome.

V- Conclusion:

This paper evaluated the predictive boundary of upstream cost and risk indicators in forecasting global food price pressure. Using an expanding-window out-of-sample framework over 2000M3–2026M5, the analysis compared parsimonious time-series benchmarks, ARDL specifications, regularised regressions and tree-based machine learning models. The forecasting target was the monthly log change of the aggregate FAO Food Price Index, conditioned on lagged energy prices, fertiliser costs, the US dollar, global supply-chain pressure and geopolitical risk.

The empirical evidence supports a clear principle of parsimony. Expanding the information set does not systematically dislodge the univariate AR(1) benchmark. Ridge and XGBoost remain competitive, but their forecast differentials are statistically indistinguishable from AR(1) over the full evaluation window. More saturated structures, particularly LASSO and ARDL-Full, perform less favourably, indicating that unconstrained lag expansion can inject estimation noise faster than it adds usable signal in data-constrained monthly forecasting.

Regime-specific evidence qualifies this result without overturning it. XGBoost records a localized advantage during the Russia-Ukraine commodity shock, suggesting that tree-based models can adapt to nonlinear cross-market interactions under acute structural stress. This advantage remains transient. Once macro-commodity conditions normalise, the full-sample dominance of the autoregressive benchmark reappears. The multi-step horizon check leads to the same conclusion: lengthening the forecast trajectory weakens the signal-to-noise ratio across multivariate designs rather than revealing delayed predictive gains from upstream indicators.

These findings separate diagnostic relevance from point-forecasting performance. Energy, fertiliser, exchange-rate, supply-chain and geopolitical indicators remain indispensable for contextual food-price monitoring. Their inclusion in forecasting systems, however, should be disciplined by transparent time-series benchmarks. Without such a benchmark, early-warning systems risk confusing historical explanatory fit with genuine real-time predictive capacity.

The empirical boundaries of this study also point to useful extensions. The aggregate FAO Food Price Index operates as a cross-category smoothing filter, potentially diluting commodity-specific predictive signals. Domestic food inflation dynamics may diverge from global food prices because national markets are mediated by subsidies, exchange-rate regimes, buffer stocks and trade restrictions. Future research could therefore extend the design to individual commodity indices, country-level food inflation, futures-market indicators or higher-frequency logistics data. The broader lesson remains straightforward: in monthly global food price forecasting, model complexity must earn its place through verified out-of-sample gains, not through theoretical plausibility alone.

Referrals and references:

1. Alexander, P., Arneth, A., Henry, R., Maire, J., Rabin, S., & Rounsevell, M. D. A. (2023). High energy and fertilizer prices are more damaging than food export curtailment from Ukraine and Russia for food prices, health and the environment. *Nature Food*, 4, 84–95. <https://doi.org/10.1038/s43016-022-00659-9>
2. Chen, S.-T., Kuo, H.-I., & Chen, C.-C. (2010). Modeling the relationship between the oil price and global food prices. *Applied Energy*, 87(8), 2517–2525. <https://doi.org/10.1016/j.apenergy.2010.02.020>
3. Chowdhury, M. A. F., Meo, M. S., Uddin, A., & Haque, M. M. (2021). Asymmetric effect of energy price on commodity price: New evidence from NARDL and time frequency wavelet approaches. *Energy*, 231, Article 120934. <https://doi.org/10.1016/j.energy.2021.120934>
4. Alnour, M., Altıntaş, H., & Rahman, M. N. (2023). Unveiling the asymmetric response of global food prices to the energy prices shocks and economic policy uncertainty. *World Development Sustainability*, 3, Article 100083. <https://doi.org/10.1016/j.wds.2023.100083>
5. Bor, Ö., & Sakarya, B. (2026). Nonlinear trends in energy and agriculture price links: A Fourier-based time series analysis. *Agricultural and Resource Economics Review*. Advance online publication.

6. Figiel, S., Gajda, J., & Kufel-Gajda, J. (2026). Effect of global energy price shocks on dynamics of world agricultural and food prices. *Agriculture*, 16(9), Article 945. <https://doi.org/10.3390/agriculture16090945>
7. Ulussever, T., Ertuğrul, H. M., Kılıç Depren, S., Kartal, M. T., & Depren, Ö. (2023). Estimation of impacts of global factors on world food prices: A comparison of machine learning algorithms and time series econometric models. *Foods*, 12(4), Article 873. <https://doi.org/10.3390/foods12040873>
8. Han, X., Yuan, T., Wang, D., Zhao, Z., & Gong, B. (2023). How to understand high global food price? Using SHAP to interpret machine learning algorithm. *PLOS ONE*, 18(4), Article e0284280. <https://doi.org/10.1371/journal.pone.0284280>
9. Toledo, P., & Duncan, R. (2024). Forecasting food price inflation during global crises. *Journal of Forecasting*, 43(4), 1087–1113. <https://doi.org/10.1002/for.3061>
10. Zelingher, R., & Makowski, D. (2024). Investigating and forecasting the impact of crop production shocks on global commodity prices. *Environmental Research Letters*, 19(1), Article 014026. <https://doi.org/10.1088/1748-9326/ad0dda>
11. Abbott, P. C., Hurt, C., & Tyner, W. E. (2008). What's driving food prices? Farm Foundation Issue Report. <https://doi.org/10.22004/ag.econ.37951>
12. Piesse, J., & Thirtle, C. (2009). Three bubbles and a panic: An explanatory review of recent food commodity price events. *Food Policy*, 34(2), 119–129. <https://doi.org/10.1016/j.foodpol.2009.01.001>
13. Baffes, J., & Dennis, A. (2013). Long-term drivers of food prices (Policy Research Working Paper No. 6455). World Bank.
14. Carriquiry, M., Dumortier, J., & Elobeid, A. (2022). Trade scenarios compensating for halted wheat and maize exports from Russia and Ukraine increase carbon emissions without easing food insecurity. *Nature Food*, 3, 847–850. <https://doi.org/10.1038/s43016-022-00622-0>
15. Arndt, C., Diao, X., Dorosh, P., Pauw, K., & Thurlow, J. (2023). The Ukraine war and rising commodity prices: Implications for developing countries. *Global Food Security*, 36, Article 100680. <https://doi.org/10.1016/j.gfs.2023.100680>
16. Lin, F., Li, X., Jia, N., Feng, F., Huang, H., Huang, J., Fan, S., & Ciais, P. (2023). The impact of Russia-Ukraine conflict on global food security. *Global Food Security*, 36, Article 100661. <https://doi.org/10.1016/j.gfs.2022.100661>
17. World Bank. (2026a). World Bank commodity price data: The Pink Sheet. World Bank. Retrieved June 9, 2026, from <https://www.worldbank.org/en/research/commodity-markets>
18. World Bank. (2026b). Commodity Markets Outlook: April 2026. World Bank. <https://openknowledge.worldbank.org/>
19. Frey, G., & Manera, M. (2007). Econometric models of asymmetric price transmission. *Journal of Economic Surveys*, 21(2), 349–415. <https://doi.org/10.1111/j.1467-6419.2007.00507.x>
20. Taghizadeh-Hesary, F., Rasoulinezhad, E., & Yoshino, N. (2019). Energy and food security: Linkages through price volatility. *Energy Policy*, 128, 796–806. <https://doi.org/10.1016/j.enpol.2018.12.043>
21. Vatsa, P., & Miljkovic, D. (2022). Energy and crop price cycles before and after the global financial crisis: A new approach. *Journal of Agricultural Economics*, 73(1), 220–233. <https://doi.org/10.1111/1477-9552.12454>
22. Charles, D. (2023). The lead-lag relationship between international food prices, freight rates, and Trinidad and Tobago's food inflation: A support vector regression analysis. *Green and Low-Carbon Economy*, 1(2), 94–103. <https://doi.org/10.47852/bonviewGLCE3202797>
23. Derindag, O. F., Chang, B. H., Gohar, R., Wong, W.-K., & Bhutto, N. A. (2023). Food prices response to global and national factors: Evidence beyond asymmetry. *Cogent Economics & Finance*, 11(1), Article 2187128. <https://doi.org/10.1080/23322039.2023.2187128>
24. Ogot, B. O., Omondi, E., Olukuru, J., Muriithi, B., & Senagi, K. (2025). Predicting food prices in Kenya using machine learning: A hybrid model approach with XGBoost and gradient boosting. *Frontiers in Artificial Intelligence*, 8, Article 1661989. <https://doi.org/10.3389/frai.2025.1661989>
25. Baumeister, C., & Kilian, L. (2012). Real-time forecasts of the real price of oil. *Journal of Business & Economic Statistics*, 30(2), 326–336. <https://doi.org/10.1080/07350015.2011.648859>
26. Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). Forecasting the price of oil. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 2, pp. 427–507). Elsevier. <https://doi.org/10.1016/B978-0-444-53683-9.00008-6>
27. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), Article e0194889. <https://doi.org/10.1371/journal.pone.0194889>
28. Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P.,

- Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., ... Ziel, F. (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, 38(3), 705–871. <https://doi.org/10.1016/j.ijforecast.2021.11.001>
29. Food and Agriculture Organization of the United Nations. (2026). FAO Food Price Index. Retrieved June 9, 2026, from <https://www.fao.org/worldfoodsituation/foodpricesindex/en/>
30. Benigno, G., di Giovanni, J., Groen, J. J. J., & Noble, A. I. (2022). A new barometer of global supply chain pressures. Federal Reserve Bank of New York Liberty Street Economics.
31. Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4), 1194–1225. <https://doi.org/10.1257/aer.20191823>
32. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
33. Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
34. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
35. Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. <https://doi.org/10.1080/07350015.1995.10524599>
36. Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2), 281–291. [https://doi.org/10.1016/S0169-2070\(96\)00719-4](https://doi.org/10.1016/S0169-2070(96)00719-4)