RESEARCH ARTICLE Agricultural Science Digest, Volume 46 Special Issue: 129-135 (2026)

Forecasting Minimum Support Price-anchored Mandi Prices for Wheat and Maize in India: A Comparative Machine-learning and Time-series Framework

Sai Prabhakar Balantrapu[1], G. Raghavender Raju[1], Satya Sai Mudigonda[2], Lalith Aditya Kalur[2] 10.18805/ag.D-6567

ABSTRACT

Background: India's minimum support price (MSP) regime creates an administered floor for wholesale (mandi) prices, making accurate price forecasting the analytical foundation for farmer revenue-risk assessment, government procurement planning and the design of price- or revenue-based crop insurance instruments. Wheat and maize present a sharp institutional contrast: wheat is procurement-binding through the Food Corporation of India, while maize receives little procurement support and its price is left largely to private trade. This study benchmarks machine-learning and classical time-series forecasters for the monthly modal mandi price of both crops across a structurally distinct, post-pandemic hold-out period and defines, in measurable terms, what it means for such a series to be MSP-anchored.

Methods: This analysis was carried out at Sri Sathya Sai Institute of Higher Learning (SSSIHL), Prasanthi Nilayam, Andhra Pradesh, India, during 2026. Daily AGMARKNET records for wheat and maize (April 2008-March 2025) were aggregated into 204 monthly All-India modal-price observations per crop-the mean of daily modal prices-with the corresponding MSP reported for reference. Eight algorithms-ridge regression, support vector regression, random forest, extremely randomised trees, gradient boosting, XGBoost, ARIMA and SARIMA-together with a naïve persistence benchmark, were trained on 177 months (April 2008-December 2022) and evaluated out-of-sample on 27 held-out months (January 2023-March 2025) using the mean absolute error, root mean squared error and mean absolute percentage error (MAPE).

Result: Ridge regression on lagged prices and a month indicator dominated every fitted ensemble and, with ARIMA and SARIMA evaluated as static multi-step forecasts, every time-series competitor for both crops, achieving a MAPE of 2.75 per cent for wheat (mean absolute error ₹ 68.97/quintal) and 2.59 per cent for maize (₹ 55.97/quintal). Tree-based learners and support vector regression deteriorated sharply on the trending test window; SARIMA was competitive only for maize. Against naïve persistence, however, ridge's edge was modest for wheat and mixed for maize-persistence matched ridge on MAE and MAPE and narrowly beat it on RMSE-indicating ridge's advantage lies chiefly over the more flexible learners. This is attributed to the near-unit autocorrelation of MSP-anchored series and the inability of tree ensembles to extrapolate beyond their training range, consistent with, though not proving, the institutional asymmetry between procurement-stabilised wheat and procurement-absent maize. These forecasts form the price engine for companion volume-forecasting and procurement-liability studies within the broader actuarial data-science revenue-protection framework.

Key words: Maize, Minimum support price, Price forecasting, Ridge regression, SARIMA, Wheat.

INTRODUCTION

The price a farmer receives at the point of first sale is among the most consequential variables in agricultural economics, shaping household income, the incentive to cultivate and the fiscal exposure of the state under price-support schemes. In India, where close to half the workforce remains engaged in agriculture and marketed surplus flows overwhelmingly through regulated wholesale markets (mandis), price formation is shaped by an administered floor: the minimum support price (MSP) announced annually by the Government of India on the recommendation of the Commission for Agricultural Costs and Prices (CACP), which sets MSP with reference to the cost of cultivation plus a margin (Meena *et al*., 2023). For crops procured at scale by the Food Corporation of India (FCI) and state agencies, the MSP operates as an effective price floor; for crops left largely to private trade, it operates only as a notional benchmark and its procurement

[1]Department of Humanities and Social Sciences (Economics), Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam-515 134, Andhra Pradesh, India.
[2]Centre of Excellence for Actuarial Data Science, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam-515 134, Andhra Pradesh, India.

Corresponding Author: Sai Prabhakar Balantrapu, Department of Humanities and Social Sciences (Economics), Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam-515 134, Andhra Pradesh, India. Email: saiprabhakarbalantrapu@sssihl.edu.in

ORCIDs: 0009-0004-7860-2965; 0000-0002-2093-1019

How to cite this article: Balantrapu, S.P., Raju, G.R., Mudigonda, S.S. and Kalur, L.A. (2026). Forecasting Minimum Support Price-anchored Mandi Prices for Wheat and Maize in India: A Comparative Machine-learning and Time-series Framework. *Agricultural Science Digest*. **46(Special Issue):** 129-135. doi: 10.18805/ag.D-6567.

Submitted: 14-08-2026 **Accepted:** 24-09-2026 **Online:** 08-10-2026

architecture, legal status and fiscal burden remain the subject of active policy debate (Balkrishna, Arya and Singh, 2023). Anticipating where the market price will settle relative to this floor is therefore the analytical foundation of crop revenue protection, procurement budgeting and the design of any index- or revenue-based insurance instrument.

The problem of forecasting policy-anchored agricultural prices is not unique to India and is usefully read against international revenue-insurance practice. In the United States, the Risk Management Agency's Revenue Protection policies under the Federal Crop Insurance Corporation combine futures-market price discovery with yield forecasts to indemnify farmers against a shortfall in revenue relative to a guaranteed level (https://www.rma.usda.gov/revenue-protection), while the European Union's Income Stabilisation Tool under the Common Agricultural Policy compensates farmers whose whole-farm or sector income falls a specified proportion below its recent average, financed through co-funded mutual funds (Severini *et al.*, 2021). What distinguishes the Indian case is the near-complete absence of a liquid futures market for wheat and maize at the mandi level and a correspondingly greater reliance on administratively announced rather than market-discovered price signals. This makes a statistically disciplined forecast of the mandi price itself, rather than of a futures curve, the necessary starting point for any revenue-protection design in the Indian context.

Classical agricultural-price forecasting has rested chiefly on the autoregressive integrated moving-average (ARIMA) framework of Box and Jenkins (1976) and its seasonal extension, SARIMA, which model a series as a linear combination of its own lags and past errors after differencing; both remain in active use for Indian production and price series (Mahajan *et al.*, 2020). These methods are attractive for their interpretability and formal model-selection criteria such as the Akaike information criterion (Akaike, 1974), but they assume linearity and cannot readily exploit exogenous covariates. The past two decades have seen a substantial shift towards machine-learning forecasters: tree ensembles such as random forests (Breiman, 2001) and extremely randomised trees (Geurts, *et al.*, 2006), gradient-boosted trees (Friedman, 2001; Chen and Guestrin, 2016) and support vector regression (Cortes and Vapnik, 1995; Drucker *et al.*, 1997), all of which approximate non-linearities without strong parametric commitment. Yet the large-scale M4 forecasting competition found that statistical methods frequently match or exceed pure machine-learning approaches on economic time series (Makridakis *et al.*, 2018, 2020), a caution directly relevant to the short, trending, policy-anchored series examined here. Ridge regression (Hoerl and Kennard, 1970), which shrinks ordinary-least-squares coefficients through an L2 penalty, trades a small increase in bias for a large reduction in variance and stabilises estimates under the collinearity created by overlapping autoregressive lags-a property that becomes decisive in short, highly autocorrelated samples but is underexplored in the Indian agricultural-price literature, which has tended either to apply ARIMA in isolation or to assert the superiority of complex learners without out-of-sample, out-of-regime testing and rarely against a persistence benchmark.

A substantial body of Indian policy scholarship examines the MSP regime and its uneven crop coverage (Acharya, 2004; Gulati *et al.*, 2013; Chand, 2017; Balkrishna *et al.*, 2023). Its recurring theme is that effective price support is concentrated in wheat and paddy through FCI operations, while coarse cereals, pulses and oilseeds receive weak or sporadic procurement, a pattern also visible at the level of individual mandi markets and crops (Borah, *et al.*, 2025). This asymmetry is the institutional fact that makes a cross-crop forecasting comparison economically meaningful rather than a purely statistical exercise: it raises the question of whether the policy-anchoring of wheat prices, relative to the freer formation of maize prices, manifests as a measurable difference in forecastability. In this study, a price series is described as MSP-anchored when two measurable conditions hold: its month-on-month level exhibits a discrete, near-monotone shift coinciding with the annual MSP revision and its twelve-month lag-spanning exactly one such revision cycle-carries explanatory content comparable to its most recent lags. This is a property of the series being forecast, established empirically, distinct from using MSP itself as a model input; the reasoning is set out in Methods below.

This study asks a deliberately narrow, policy-relevant question: which class of forecasting model most reliably predicts the monthly modal mandi price of wheat and maize, one month ahead, across a structurally distinct hold-out period and does it improve materially on a naive persistence benchmark? Three features of the Indian mandi price make this non-trivial. The series are policy-anchored in the sense defined above. The two crops are structurally asymmetric in their relationship to the MSP floor: wheat trades above its MSP in most months because FCI procurement absorbs surplus, whereas maize, lacking comparable procurement, frequently trades below its declared MSP. The post-2022 period constitutes a genuinely out-of-regime test window, with prices rising after the COVID-19 disruption and global commodity shock to historically unprecedented levels. The contribution is threefold: a long, clean, All-India monthly price series for both crops assembled from primary AGMARKNET records; a benchmark of eight forecasting algorithms and a naive persistence baseline, spanning penalised linear models, kernel methods, tree ensembles, gradient boosting and the Box-Jenkins family under a single, fixed train-test protocol and an econometric explanation for why a parsimonious regularised linear model outperforms substantially more flexible learners. The resulting forecasts are the price input to companion studies on market-arrival forecasting and government procurement-liability projection within the wider actuarial

data-science revenue-protection programme; the present study is confined strictly to the price-prediction objective.

MATERIALS AND METHODS

This analysis was carried out at Sri Sathya Sai Institute of Higher Learning (SSSIHL), Prasanthi Nilayam, Andhra Pradesh, India, during 2026. The empirical analysis uses transaction-level records from AGMARKNET, the portal of the Directorate of Marketing and Inspection under the Ministry of Agriculture and Farmers' Welfare, which compiles daily commodity arrivals and minimum, maximum and modal prices reported by regulated markets across India (Directorate of Marketing and Inspection, n.d.). For wheat, three workbooks spanning April 2008 to March 2025 were consolidated, yielding more than 2.4 million daily market-variety records before cleaning; a comparable extract was prepared for maize. After removing records with missing grade information and coercing price fields to numeric type, the daily records were aggregated to a single All-India monthly series per crop. The monthly modal price used throughout is the arithmetic mean of the daily modal-price records reported across all regulated markets nationally within each calendar month-volume-informed in that sense, though not a volume-weighted average. MSP values are reported for reference (Table 1) and do not enter any forecasting model. The resulting panel comprises 204 monthly observations per crop (April 2008-March 2025). Following a fixed chronological protocol, the first 177 months (April 2008-December 2022) form the training set and the final 27 months (January 2023-March 2025)-a structurally distinct, post-pandemic price regime-form the out-of-sample test set.

As the objective is a one-step-ahead operational forecast, predictors were restricted to information available before the forecast month. For the six machine-learning models, lag features for every test month are computed from the actually observed historical price series, so each forecast is a genuine one-step-ahead prediction rather than a recursive forecast built on the model's own earlier outputs. ARIMA and SARIMA are evaluated differently, consistent with standard Box-Jenkins practice: fitted once on the training data, they produce a single dynamic forecast for all 27 test months from that origin, without being re-fed actual test-period observations; the ML and classical-model forecasts, though judged on the same metrics, are thus not generated under identical protocols. For the machine-learning models, the feature vector comprises price lags at one, two, three, six and twelve months plus a calendar-month indicator. MSP itself is not entered as a regressor: revised at most once a year, it is piecewise constant within a marketing season and near-collinear with the twelve-month lag already present; its influence is captured indirectly through the level and trend of the lagged price series it shapes. After lag construction the usable sample reduces to 165 training and 27 test observations; the short lags capture short-run persistence, the six-month lag spans the harvest-to-lean-season gap and the twelve-month lag encodes year-on-year structure including the annual MSP revision.

Nine specifications were estimated under an identical feature set and split. A naïve persistence benchmark, in which next month's forecast equals the current month's observed price, is reported first. Ridge regression minimises the penalised residual sum of squares, with an L2 penalty of $\alpha = 1.0$ (the scikit-learn default) applied identically to both crops; hyperparameters throughout were fixed a priori rather than tuned by cross-validation, a limitation discussed below. Support vector regression uses a radial basis function kernel with $C = 100$ and $\varepsilon = 10$. Random forest and extremely randomised trees each use 300 trees, maximum depth 10 and a minimum of 5 samples per leaf. Gradient boosting uses 300 trees, learning rate 0.05, maximum depth 4 and minimum leaf size 5. XGBoost uses 400 trees, learning rate 0.05, maximum depth 3, minimum child weight 3, 80 per cent row and column subsampling and L2 regularisation of 1.0; its shallower trees and added regularisation relative to gradient boosting reflect the small training sample. ARIMA and SARIMA orders were selected by grid search minimising the Akaike information criterion, after an augmented dickey-fuller test (Dickey and Fuller, 1979) established the order of integration. All models were implemented in Python using scikit-learn, XGBoost and statsmodels; package version numbers were not pinned in the analysis notebooks.

All models were assessed on the untouched 27-month test set using the mean absolute error (MAE), root mean

Table 1: Structure of the All-India monthly price dataset.

Attribute	Wheat	Maize
Primary source	AGMARKNET	AGMARKNET
Series span	Apr 2008-Mar 2025	Apr 2008-Mar 2025
Raw daily records (pre-clean)	2,426,543	comparable scale
Monthly observations	204	204
Training months	177 (Apr 08-Dec 22)	177 (Apr 08-Dec 22)
Test months	27 (Jan 23-Mar 25)	27 (Jan 23-Mar 25)
MSP, RMS 2009-10 (wheat)/KMS 2008-09 (maize) (₹/quintal)	1,080	840
MSP, RMS 2025-26 (wheat)/KMS 2024-25 (maize) (₹/quintal)	2,425	2,225

Source: Authors' computation from AGMARKNET records and CACP/DA & FW MSP tables.

squared error (RMSE) and mean absolute percentage error (MAPE), computed on the original price scale. RMSE, which penalises large errors more heavily, was used as the primary ranking criterion; MAPE provides a scale-free measure supporting cross-crop comparison. No information from the test window entered model fitting or hyper-parameter selection at any stage.

RESULTS AND DISCUSSION

Stationarity and model specification

The augmented Dickey-Fuller test failed to reject the unit-root null for both price series-the test statistic is -0.19 for wheat (p = 0.94) and -0.90 for maize (p = 0.79)-confirming non-stationarity in levels and the need for first differencing in the Box-Jenkins models; given the moderate effective sample length (177 training months), this result is treated as indicative rather than definitive, consistent with the generally low power of unit-root tests in short samples. The AIC-optimal specifications are ARIMA (1,1,2) and SARIMA (0,1,2) (0,1,1,12) for wheat and ARIMA (0,1,1) and SARIMA (0,1,2) (0,1,1,12) for maize; in both crops the SARIMA specification achieves a markedly lower AIC than the non-seasonal ARIMA, consistent with an annual seasonal structure tied to harvest and procurement calendars.

Price-forecasting accuracy

Table 2 reports out-of-sample performance across all nine specifications. Ridge dominates every fitted competitor on every metric for wheat, attaining MAE ₹ 68.97/quintal, RMSE ₹ 94.49/quintal and MAPE 2.75 per cent against a test-period mean price of about ₹ 2,517/quintal; the nearest ML competitor, XGBoost, more than doubles this MAPE (6.12%) and support vector regression collapses on the trending window (MAPE 28.25%). Against naïve persistence, however, ridge's margin is modest: persistence alone achieves MAE ₹ 74.23/quintal, RMSE ₹ 98.97/quintal and MAPE 2.95 per cent for wheat, so ridge improves on carrying forward last month's price by roughly 7 per cent on MAE. For maize the comparison is finer-balanced: persistence achieves MAE ₹ 56.94/quintal, RMSE ₹ 71.69/quintal and MAPE 2.64 per cent, essentially matching ridge's MAE (₹ 55.97) and MAPE (2.59%) and narrowly beating it on RMSE (₹ 71.69 versus ₹ 73.25)-a genuine case in which the fitted model does not clearly outperform the simplest possible forecast. SARIMA is the second-best fitted performer for maize (MAPE 5.21%), ahead of every tree-based learner-consistent with, though not proof of, the freer, more seasonally driven formation of maize prices absent large-scale procurement. Fig 1 tracks the ridge forecast against actual prices over the hold-out window. Wheat rose almost monotonically from about ₹ 2,203/quintal in April 2023 to a peak near ₹ 2,926/quintal in January 2025; the ridge forecast tracks this trajectory closely, with monthly errors typically within ₹ 30-80/quintal and the largest misses confined to the two post-harvest turning points, where the one-period autoregressive structure briefly carries the previous month's price forward before re-converging-a transparent, bounded limitation rather than systematic bias. Maize shows comparable tracking around its more volatile, seasonally driven path.

Why a linear model wins

The dominance of a penalised linear autoregression over flexible non-linear learners is counter-intuitive only in the abstract; given the structure of the MSP-anchored series defined above, it is broadly expected and three mechanisms plausibly explain it. First, the targets exhibit near-unit first-order autocorrelation-the best single predictor of next month's price is this month's price plus a slow drift-so the data-generating process is close to a regularised AR(1). This is also why ridge's advantage over literal persistence is narrow rather than large (Table 2): Ridge is, in effect, a smoothed, seasonally adjusted version of the same carry-forward logic, not a qualitatively different forecast. Second, the test window lies outside the training range; tree-based models predict by averaging training leaf values and cannot extrapolate beyond the convex hull of the training targets, so they systematically under-predict on a trending

Table 2: Out-of-sample price-forecasting performance for wheat and maize, January 2023-March 2025 (Best model in bold).

Model	Wheat			Maize		
	MAE (₹/q)	RMSE (₹/q)	MAPE (%)	MAE (₹/q)	RMSE (₹/q)	MAPE (%)
Naïve persistence (random walk)	74.23	98.97	2.95	56.94	71.69	2.64
Ridge regression	68.97	94.49	2.75	55.97	73.25	2.59
SARIMA (0,1,2) (0,1,1,12)	171.07	193.75	7.08	107.48	133.62	5.21
ARIMA†	173.46	207.04	7.05	129.72	164.78	5.89
XGBoost	162.84	219.56	6.12	229.03	261.45	10.29
Gradient boosting	168.47	220.21	6.40	155.50	195.56	6.92
Random forest	278.08	335.81	10.55	160.48	198.70	7.14
Extra trees	281.27	339.15	10.68	170.08	206.18	7.58
Support vector regression	724.64	766.25	28.25	473.15	521.43	21.39

Source: Authors' computation on AGMARKNET data. Naïve persistence forecasts each test month as the previous month's observed price. †ARIMA (1,1,2) for wheat; ARIMA (0,1,1) for maize. Metrics on the original price scale.

series, whereas a linear model extrapolates along the fitted slope-this is where ridge's advantage over the other six fitted models becomes large. Third, the training sample is short (165 observations), a regime in which ridge's variance-reduction is decisive while high-capacity learners overfit idiosyncratic fluctuations. The intuition resembles a seatbelt: restrictive in gentle driving, but decisive when the road is short and the surface uncertain, trading a small, known bias for a large reduction in variance when training data cannot support an unconstrained fit.

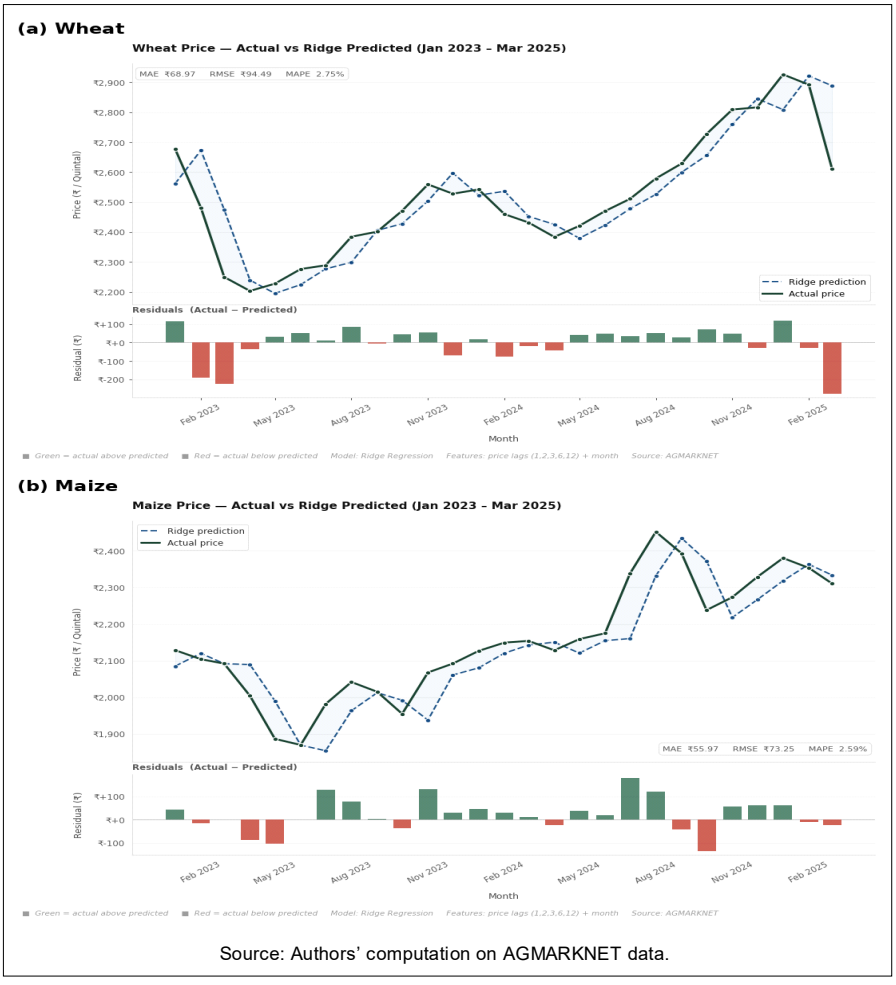
Policy relevance and limitations

Table 3 consolidates the headline results: wheat at 2.75 per cent and maize at 2.59 per cent MAPE, with the classical seasonal model mid-ranked for wheat and second-best for maize. Read alongside the institutional asymmetry documented in the policy literature (Balkrishna *et al.*, 2023), a plausible interpretation is that wheat's price is stabilised by procurement and dominated by persistence, which ridge captures and to which seasonality adds little, whereas maize's price retains stronger

Table 3: Cross-crop summary of best-model price-forecast accuracy.

Metric (Best model)	Wheat	Maize
Selected model	Ridge regression	Ridge regression
MAE (₹/quintal)	68.97	55.97
RMSE (₹/quintal)	94.49	73.25
MAPE (%)	2.75	2.59
Second-best model (by MAPE)	XGBoost	SARIMA
ADF unit-root test, levels	Non-stationary	Non-stationary

Source: Authors' computation.



Source: Authors' computation on AGMARKNET data.

Fig 1: Actual versus ridge-predicted modal mandi price and monthly residuals over the out-of-sample window, January 2023-March 2025: (a) Wheat; (b) Maize.

seasonal dynamics that SARIMA can exploit; the design demonstrates the forecast-performance difference but does not, on its own, establish that procurement policy causes it. For procurement-binding crops such as wheat, a reliable forecast relative to the announced MSP gives FCI and state agencies advance notice of months in which the market may fall to or below the support level. For procurement-absent crops such as maize, the same forecast quantifies the gap between the notional support price and the price farmers actually realise, informing debates on extending effective price support or designing deficiency-payment alternatives. Placed alongside international practice, the exercise also clarifies what India's policy-anchored mandi prices can and cannot substitute for: a disciplined, auditable signal in the absence of a deep futures market, but unlike the futures-based guarantees of the US Revenue Protection programme (https://www.rma.usda.gov/revenue-protection) or the EU's income-based Income Stabilisation Tool (Severini, Zinnanti, Borsellino and Schimmenti, 2021), a single-point forecast rather than a market-implied probability distribution. Two further limitations qualify the evaluation: it rests on a single 27-month hold-out period, so the ranking should be read as evidence from one structural break rather than a general law and hyperparameters were fixed a priori rather than tuned by cross-validation, so the reported ML results represent this fixed specification rather than each model's best achievable accuracy.

CONCLUSION

This study benchmarked eight forecasting algorithms and a naïve persistence baseline for the monthly modal mandi price of wheat and maize across India, using seventeen years of AGMARKNET data and a deliberately out-of-regime hold-out period. A ridge-regularised linear autoregression on price lags and a seasonal indicator dominated all fitted competitors under the evaluation protocol described in Methods, delivering 2.75 per cent and 2.59 per cent MAPE for wheat and maize respectively; its margin over naïve persistence was real but modest for wheat and mixed for maize-persistence matched or narrowly beat ridge on two of three metrics. This qualifies rather than overturns the central finding: Ridge's decisive advantage lies over the more flexible learners and the explanation- near-unit autocorrelation, the inability of tree ensembles to extrapolate on a trending series and the variance-reduction value of regularisation in a short sample-accounts for both patterns together. The crops differ chiefly in the rank of the classical seasonal model-mid-ranked for procurement-stabilised wheat, second-best for procurement-absent maize-a pattern consistent with, though not sufficient to prove, the institutional asymmetry of the MSP regime. Five limitations point to future work: the framework is univariate and exogenous drivers such as global commodity prices, input costs and rainfall may improve turning-point accuracy; the aggregation is national, masking state- and mandi-level heterogeneity; the evaluation rests on a single hold-out period rather than rolling-origin evaluation; hyperparameters were fixed rather than tuned, so reported ML accuracy is a lower bound on what tuning could achieve; and the comparison is point-forecast based, so probabilistic forecasts with calibrated intervals would better serve the actuarial and procurement-liability applications this research programme is directed towards. The processed dataset and analysis code are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENT

This study forms part of the first author's doctoral research programme at the Department of Humanities and Social Sciences (DHSS), Sri Sathya Sai Institute of Higher Learning. No external funding was received for this research.

Disclaimers

The views and conclusions expressed in this article are solely those of the authors and do not necessarily represent the views of their affiliated institutions. The authors are responsible for the accuracy and completeness of the information provided, but do not accept any liability for any direct or indirect losses resulting from the use of this content.

Informed consent

Not applicable. This study is based entirely on secondary, publicly available agricultural market data (AGMARKNET; CACP/DA & FW minimum support price records) and does not involve human participants, animal experimentation, or clinical data.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Acharya, S.S. (2004). Agricultural Marketing in India (4th ed.). Oxford and IBH Publishing.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. **19(6):** 716-723.

Balkrishna, A., Arya, V. and Singh, S.K. (2023). Minimum support price under the aegis of universal basic income: Understanding the implications and way forward: A review. *Bhartiya Krishi Anusandhan Patrika*. **38(3):** 227-237. doi: 10.18805/BKAP640.

Borah, S., Talukdar, U., Barman, U., Deka, N., Baruah, S. and Ahmed, A.A. (2025). Market arrivals and price behaviour of some selected crops: A study from major markets of Assam. *Bhartiya Krishi Anusandhan Patrika*. **40(2):** 180-186. doi: 10.18805/BKAP796.

Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control (Rev. ed.). Holden-Day.

Breiman, L. (2001). Random forests. *Machine Learning*. **45(1):** 5-32.

Chand, R. (2017). Doubling farmers' income: Rationale, strategy, prospects and action plan (NITI Aayog Policy Paper No. 1/2017). NITI Aayog.

Chen, T. and 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).

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*. **20(3):** 273-297.

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*. **74(366):** 427-431.

Directorate of Marketing and Inspection. (n.d.). AGMARKNET: Agricultural Marketing Information Network [Data set]. Ministry of Agriculture and Farmers Welfare, Government of India. https://agmarknet.gov.in.

Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A. and Vapnik, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems*. **9:** 155-161.

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*. **29(5):** 1189-1232.

Geurts, P., Ernst, D. and Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*. **63(1):** 3-42.

Gulati, A., Saini, S. and Jain, S. (2013). Monsoon 2013: Estimating the Impact on Agriculture (ICRIER Working Paper No. 269). Indian Council for Research on International Economic Relations.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*. **12(1):** 55-67.

https://www.rma.usda.gov/revenue-protection.

Mahajan, S., Sharma, M. and Gupta, A. (2020). ARIMA modelling for forecasting of rice production: A case study of India. *Agricultural Science Digest*. **40(4):** 404-407. doi: 10.18805/ag.D-5029.

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *Plos One*. **13(3):** e0194889.

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*. **36(1):** 54-74.

Meena, G.L., Sharma, L., Singh, H., Bairwa, K.C. and Kumar, K. (2023). Cost and return analysis of *kharif* sorghum (*Sorghum bicolor*) in Bhilwara district, Rajasthan. *Agricultural Science Digest*. **43(2):** 196-200. doi: 10.18805/ag.D-5566.

Severini, S., Zinnanti, C., Borsellino, V. and Schimmenti, E. (2021). EU income stabilization tool: Potential impacts, financial sustainability and farmer's risk aversion. *Agricultural and Food Economics*. **9:** 34. https://doi.org/10.1186/s40100-021-00205-4.