A previsão de preços melhora o *hedge* na entressafra? Evidências do mercado de boi gordo no Brasil

Does Forecasting Improve Hedging in the Off-Season? Evidence from Cattle Markets

João Guilherme Perotto Fernandes
Luiz de Queiroz College of Agriculture, University of São Paulo (ESALQ/USP). Piracicaba, SP, Brazil.
Nicole Rennó Castro
Luiz de Queiroz College of Agriculture, University of São Paulo (ESALQ/USP). Piracicaba, SP, Brazil.

GT 01 – Mercados agrícolas e comércio exterior

Resumo: Este artigo investiga se a incorporação de previsões de preços às decisões de *hedge* melhora a gestão de risco de confinadores de bovinos durante a entressafra. Propõe-se um sistema em dois estágios, previsão-decisão, no qual redes neurais LSTM geram previsões em múltiplos horizontes (1, 6, 12 e 18 meses) – posteriormente filtradas por controle de taxa de falsos descobrimentos, sintetizadas por modelos fatoriais e combinadas sob o critério de Model Confidence Set (MCS) – que alimentam um agente de aprendizado por reforço do tipo Soft Actor-Critic. Esse último é formulado como um problema de decisão markoviano restrito, com imposição de limites de risco via CVaR e *drawdown* máximo por meio de um mecanismo Lagrangiano primal-dual. O sistema é estimado para o mercado de boi gordo da B3, utilizando validação fora da amostra por janelas expansivas no período de janeiro de 2016 a dezembro de 2024. No estágio de previsão, modelos baseados em LSTM alcançam coeficientes de determinação fora da amostra superiores a 0,96 no horizonte de um mês, enquanto combinações via MCS preservam conteúdo preditivo positivo em horizontes mais longos, nos quais modelos individuais tendem a sobreajustar. No estágio de decisão, o modelo SAC-CMDP apresenta o menor *drawdown* máximo (−11,41 R$/arroba), o maior índice de Sortino e efetividade de hedge superior a 0,90, posicionando-se como o melhor desempenho em um indicador composto de risco que incorpora métricas de cauda, variabilidade interanual e custos de transação. Embora os ganhos médios em relação a uma estratégia estática de *hedge* integral sejam relativamente modestos, observa-se melhora significativa nos episódios adversos, nos quais o agente ajusta dinamicamente o nível de proteção antecipando movimentos de convergência de base. Os resultados indicam que o valor econômico da previsão não se manifesta primariamente na acurácia média, mas na redução de riscos, em linha com a distinção teórica entre funções de perda estatísticas e econômicas.
Palavras-chave: aprendizado por reforço, gestão de risco de cauda, hedge de gado, LSTM, previsão de preços.

Abstract: This paper investigates whether integrating price forecasting into hedging decisions improves risk management for cattle feedlot operators during the off-season, when basis risk and supply concentration peak. We propose a two-stage prediction–decision system in which LSTM networks generate multi-horizon (1, 6, 12, and 18-month) price forecasts – screened via false discovery rate control, compressed through factor models, and combined under Model Confidence Set discipline – that feed a Soft Actor-Critic reinforcement learning agent constrained by CVaR and maximum drawdown through a primal-dual Lagrangian mechanism. The system is estimated on Brazil's B3 live cattle market using expanding-window walk-forward validation over January 2016–December 2024. On the forecasting side, LSTM-based models achieve out-of-sample R^2 above 0.96 at one month and MCS combinations retain positive predictive content at 18 months, where individual models overfit. On the hedging side, SAC-CMDP attains the lowest maximum drawdown (−11.41 R$/arroba), the best Sortino ratio, and hedge effectiveness above 0.90, ranking first in a composite risk score that incorporates tail-risk metrics, inter-annual variability, and transaction costs. The margin over a static full-hedge benchmark is narrow on average metrics but concentrated in the worst episodes, where the agent modulates coverage in anticipation of basis convergence. These results demonstrate that the economic value of forecasting accrues not through average-case accuracy but through tail-risk reduction in the decision layer, consistent with the theoretical distinction between statistical and economic loss functions.
Keywords: cattle hedging, price forecasting, reinforcement learning, LSTM, tail-risk management.

1. Introduction

Price risk is a defining feature of livestock production. A feedlot operator who places cattle on feed commits biological capital to a multi-month horizon over which output prices can shift by magnitudes exceeding the entire operating margin (TOMEK; PETERSON, 2001). Because this timeline is largely irreversible, demand for hedging instruments is persistent – yet basis risk fluctuates with seasonal supply patterns and cattle-cycle regime shifts, undermining conventional risk transfer (LEUTHOLD; JUNKUS; CORDIER, 1989; GARCIA; LEUTHOLD, 2004). The standard approach treats hedge ratio estimation as a static mean-variance problem (EDERINGTON, 1979; MYERS; THOMPSON, 1989). Time-varying extensions via GARCH-family models improve on naïve benchmarks under parameter stability (BAILLIE; MYERS, 1991), but cattle markets routinely violate this assumption: seasonal transitions cause the joint distribution of spot and futures returns to shift in ways that render fixed-parameter ratios stale (ALIZADEH; NOMIKOS; POULIASIS, 2008).

A separate literature has pursued improved price forecasting, recognizing that hedging decisions are conditional on price expectations (TOMEK, 1997). Machine learning methods – particularly LSTM networks (HOCHREITER; SCHMIDHUBER, 1997) – have outperformed linear benchmarks at short horizons (JAISWAL et al., 2021; GU; KELLY; XIU, 2020). Yet forecasting and hedging remain studied in isolation: forecasting models are evaluated by statistical loss functions detached from economic positions, while hedging models take realized prices as given. This separation is consequential: a statistically superior forecast may yield inferior hedging if its errors correlate with elevated basis risk, while a modest forecast may be valuable if it reliably identifies regime transitions (GRANGER; MACHINA, 2006; LEITCH; TANNER, 1991).

This paper bridges the gap by proposing an integrated prediction–decision system for live cattle hedging. An LSTM forecasting component generates multi-horizon price signals that feed a reinforcement learning (RL) agent, which learns an adaptive hedge policy maximizing risk-adjusted performance net of transaction costs, margin expenses, and basis risk penalties. The architecture treats hedging as a sequential decision problem under uncertainty (SUTTON; BARTO, 2018), solved via Soft Actor-Critic (SAC) with entropy regularization (HAARNOJA et al., 2018) and constrained by CVaR and maximum drawdown through a primal-dual Lagrangian mechanism (ROCKAFELLAR; URYASEV, 2000; ALTMAN, 1999).

On the forecasting side, the LSTM module incorporates automated stationarity testing, variable screening under false discovery rate control, dimensionality reduction, and forecast combination via Model Confidence Set (HANSEN; LUNDE; NASON, 2011), producing point forecasts and dispersion signals as state variables. The walk-forward expanding-window design follows temporal integrity best practices (LOPEZ DE PRADO, 2018). On the decision side, we compare RL policies (SAC and PPO) against interpretable baselines – no hedge, full hedge, static 50%, time-to-maturity, and rolling ordinary least squares (OLS) – using tail-risk metrics with formal statistical tests. The empirical application uses Brazil's live cattle market, the world's largest commercial herd, where basis risk, seasonal supply concentration, and limited adoption of risk management tools amplify the stakes.

The remainder of the paper proceeds as follows. Section 2 examines the current approach concerning commodity hedging. Section 3 describes the data, the forecasting module, and the RL environment. Section 4 presents empirical results. Section 5 concludes.

2. Literature review

The minimum-variance approach of Ederington (1979) and Johnson (1960) remains the conceptual anchor for commodity hedging. Under joint normality of spot and futures returns, the optimal hedge ratio reduces to the OLS slope coefficient, a result extended by Myers and Thompson (1989) to account for conditional information sets. Empirical work on cattle markets has consistently documented that naïve and static ratios leave substantial residual risk, particularly near seasonal supply transitions (LEUTHOLD; JUNKUS; CORDIER, 1989).

Time-varying hedge ratios estimated via GARCH specifications – bivariate GARCH (BAILLIE; MYERS, 1991), BEKK (ENGLE; KRONER, 1995), and DCC (ENGLE, 2002) – improve in-sample fit by allowing the variance–covariance matrix to evolve, but their out-of-sample gains are sensitive to regime stability (GARCIA; LEUTHOLD, 2004). Alizadeh, Nomikos and Pouliasis (2008) show that Markov-switching models capture structural breaks in freight markets and propose regime-dependent hedge ratios; analogous regime shifts – driven by the cattle cycle and seasonal slaughter concentration – characterize Brazilian cattle markets (SACHS; PINATTI, 2007). More recently, copula-based approaches have relaxed the elliptical dependence assumption underlying GARCH hedging (LAI; CHEN; GERLACH, 2009), yet they still require *ex ante* parametric specification of the dependence structure.

An important practical dimension is basis risk – the wedge between spot and futures at liquidation. Garcia and Leuthold (2004) document that basis variability in livestock markets is higher in the off-season, when supply is concentrated and locational premia diverge. In Brazil, the B3 live cattle contract settles against the CEPEA/ESALQ indicator, and basis dynamics are influenced by the concentration of slaughter among a small number of meatpackers (ROCHELLE, 2000; TONIN; ALVES, 2014). These features imply that a single fixed-parameter model is unlikely to capture the full conditional distribution of hedging outcomes across seasons and cattle-cycle phases.

Even so, forecasting cattle prices has traditionally relied on autoregressive and VAR specifications informed by supply fundamentals – herd dynamics, rainfall, and slaughter seasonality (MARGARIDO et al., 1996; BRESSAN; LIMA, 2002). While these models exploit well-documented mean-reverting and seasonal patterns, they impose linearity and fixed lag structures that limit their ability to capture nonlinear interactions across predictor dimensions.

Machine learning methods relax these constraints. Among tree-based methods, random forests (BREIMAN, 2001) and gradient boosting (CHEN; GUESTRIN, 2016) have shown strong performance in agricultural commodity forecasting by capturing local threshold effects and heterogeneous predictor importance (RIBEIRO; OLIVEIRA, 2011). Among neural architectures, LSTM networks (HOCHREITER; SCHMIDHUBER, 1997) are particularly suited to economic time series with multi-scale temporal dependencies, since their gated units regulate information flow across long lags without suffering from vanishing gradients (GOODFELLOW; BENGIO; COURVILLE, 2016). Gu, Kelly and Xiu (2020) demonstrate that deep learning methods – including LSTMs – deliver significant out-of-sample predictive gains for asset returns in cross-sectional and panel settings. Jaiswal et al. (2021) apply LSTM networks to agricultural price prediction and report persistent improvements over ARIMA benchmarks.

However, forecast evaluation in this literature is almost exclusively based on statistical loss functions – RMSE, MAE, directional accuracy – which are detached from the economic decisions that forecasts are meant to inform. Leitch and Tanner (1991) show that statistical accuracy and economic value can diverge sharply, and Granger and Machina (2006) argue that the mapping from forecast errors to decision payoffs depends on the loss function of the decision-maker rather than the symmetric losses implicit in standard metrics. This gap

between forecast evaluation and decision value is the central motivation for the integrated approach developed in this paper.

A related methodological concern is overfitting and model selection under temporal dependence. High-dimensional predictor sets – enriched with lags, momentum features, and seasonal dummies – can produce spuriously accurate in-sample fits. Shrinkage methods such as lasso (TIBSHIRANI, 1996), ridge (HOERL; KENNARD, 1970), elastic net (ZOU; HASTIE, 2005), and adaptive lasso (ZOU, 2006) impose regularization that trades bias for variance reduction in finite samples. Dimensionality reduction via PCA and factor models (JOLLIFFE, 2002; BAI; NG, 2002; STOCK; WATSON, 2002) provides an alternative compression strategy. The Model Confidence Set (HANSEN; LUNDE; NASON, 2011), combined with the Diebold-Mariano test (DIEBOLD; MARIANO, 1995), offers an inferentially disciplined procedure for retaining only statistically indistinguishable forecasters – an approach we adopt to prevent data snooping (WHITE, 2000) and to produce robust forecast inputs for the decision layer.

2.1 Hedging decisions during off-season

Reinforcement learning (RL) formalizes sequential decision-making under uncertainty by learning a policy that maps states to actions so as to maximize cumulative reward (SUTTON; BARTO, 2018). The foundation is naturally suited to hedging, where the agent must choose a coverage level at each rebalancing date given evolving market conditions, and where the consequences of current decisions depend on future realizations (POWELL, 2011). Early applications of RL to portfolio management adopted tabular Q-learning in discretized state–action spaces (MOODY; WU, 1997; NEUNEIER, 1996), but these approaches scale poorly as the dimensionality of the state grows. Modern deep RL algorithms – Deep Q-Networks (MNIH et al., 2015), Proximal Policy Optimization (SCHULMAN et al., 2017), and Soft Actor-Critic (HAARNOJA et al., 2018) – handle continuous state and action spaces and have been successfully applied to portfolio allocation (JIANG; XU; LIANG, 2017; WANG et al., 2021), options hedging (CAO et al., 2021; BUEHLER et al., 2019), and execution optimization (NING; TRAN; NGUYEN, 2021).

In the hedging domain specifically, Buehler et al. (2019) formulate deep hedging as a risk minimization problem under market frictions, showing that a neural-network-based policy can replicate derivative payoffs more efficiently than delta-hedging when transaction costs and discrete rebalancing are considered. Cao et al. (2021) extend this work by

incorporating CVaR constraints through Lagrangian relaxation, demonstrating that tail-risk-aware policies reduce maximum drawdown without sacrificing mean performance. These results are obtained in derivative settings with well-defined replication targets; the transfer to commodity hedging – where there is no closed-form payoff to replicate and the hedge objective is margin protection rather than payoff replication – remains largely unexplored.

Two recent developments are directly relevant to our formulation. First, constrained Markov Decision Processes (CMDP) (ALTMAN, 1999) provide a principled mechanism for incorporating hard constraints – CVaR floors, drawdown ceilings – into the RL objective via primal-dual Lagrangian methods, avoiding the *ad hoc* penalty weighting typical of unconstrained reward shaping (CHOW et al., 2017). Second, entropy-regularized policies, as implemented in SAC (HAARNOJA et al., 2018), produce stochastic actions that naturally explore the action space during training and provide built-in robustness to model misspecification – a desirable property when transition dynamics shift across regimes (HAMILTON, 1994; EYSENBACH; LEVINE, 2022).

Despite these advances, no prior work integrates a multi-horizon forecasting approach with an RL hedging agent in commodity markets, conditioning the hedge policy on both market state variables and predictive signals while enforcing realistic frictions and tail-risk constraints. This paper fills that gap.

3. Methodology

The empirical design treats forecasting and hedging as a single identification problem. Forecasts are estimated first, validated out of sample, and only then translated into decision variables. This ordering is essential because hedge policies inherit both the informational content and the uncertainty structure of the forecasting stage.

We provide a comprehensive description of the data employed at each stage of the analysis in the Online Appendix. Specifically, the Appendix details the datasets used in the forecasting stage, as well as those used in the decision stage, where forecasts are mapped into hedging strategies. For each component, we document variable definitions, classification schemes, and corresponding data sources. Importantly, part of the analysis relies on proprietary series from the livestock input group at CEPEA, which are not publicly available. We further report the original data providers and access conditions wherever possible,

ensuring a high degree of transparency and facilitating reproducibility across both the prediction and decision environments.

3.1 Forecasting layer

Let Y_t denote the target level series and h the forecasting horizon. To stabilize long-horizon dynamics and variance, the target is modeled in forward log-difference form:

$$(1)\quad y_t(h) = \ln(Y_{t+h} + a) - \ln(Y_t + a)$$

where a is a strictly positive shift used only when non-positive observations are present. Forecasts are returned to economic level space by

$$(2)\quad \hat{Y}_{t+h|t} = \exp(\hat{y}_t(h) + \ln(Y_t + a)) - a$$

which provides the operational bridge between statistical estimation and economically interpretable decision inputs.

The predictor space is constructed as follows. A lag matrix is created for each candidate predictor $x_{j,t}$ using variable-specific lag depth L_j – three months by default; six for cattle-cycle and rainfall series following Margarido et al. (1996). The target enters through autoregressive terms with short-memory and annual seasonal lags. Momentum and volatility enrichments are added via rolling-window moving averages and standard deviations with windows $w \in \{3, 6\}$. Seasonal controls enter as monthly dummies with Fourier harmonics up to order three. Each series is transformed by a stationarity selection rule over six candidate transformations, accepting the first that satisfies joint ADF rejection and KPSS non-rejection. Missing values are filled causally using training-window medians to prevent look-ahead leakage.

Variable screening is performed by univariate relevance to $y_t(h)$, with multiplicity control via the Benjamini-Hochberg correction (BENJAMINI; HOCHBERG, 1995):

$$(3)\quad q_i = BH(p_i),\ \text{retain } i \text{ if } q_i \leq \alpha$$

After screening, the forecasting design branches into three model families. First, reduction models – PCA and factor projections – compress screened predictors into latent representations (JOLLIFFE, 2002; BAI; NG, 2002). Second, selection models – adaptive lasso, lasso, ridge, elastic net, random forest, and XGBoost – estimate forecasts directly from selected predictors, combining sparse linear shrinkage (TIBSHIRANI, 1996; HOERL; KENNARD, 1970; ZOU; HASTIE, 2005; ZOU, 2006) with nonlinear tree-based methods (BREIMAN, 2001; CHEN; GUESTRIN, 2016). Third, Model Confidence Set filtering retains

only statistically competitive forecasts under bootstrap stepdown elimination (HANSEN; LUNDE; NASON, 2011).

Each candidate forecast is generated by a multi-layer LSTM network with attention, trained via expanding-window walk-forward estimation. The LSTM architecture uses stacked recurrent layers with dropout and layer normalization, an attention mechanism over the lookback sequence, and a dense output head. Training employs the Huber loss, Adam optimizer with gradient clipping, early stopping, and learning-rate reduction on plateau. The expanding-window scheme adds new observations at each step without discarding earlier data, preserving temporal integrity and avoiding look-ahead bias (LOPEZ DE PRADO, 2018).

Forecast comparison uses the Diebold-Mariano statistic (DIEBOLD; MARIANO, 1995):

$$(4)\quad DM(i,j) = \bar{d}(i,j) \cdot \sqrt{T} \cdot (\hat{\omega}(i,j))^{-1/2}$$

where $\bar{d}(i,j)$ is the mean loss differential and $\hat{\omega}(i,j)$ is the long-run variance. Each horizon h contributes three level forecasts to the decision stage:

(5a) $\hat{Y}_{t+h|t}(H)$ – superior forecast under DM ranking among reduction and selection candidates

(5b) $\hat{Y}_{t+h|t}(MCS)$ – Model Confidence Set aggregate among statistically non-inferior models

(5c) $\hat{Y}_{t+h|t}(B)$ – benchmark forecaster

3.2 Economic state representation

The decision state is constructed from variables with direct economic interpretation. Let S_t and F_t denote spot and futures prices. The core transformations are:

$$(6)\quad B_t = S_t - F_t$$

$$(7)\quad b_t = B_t \cdot (|F_t| + \varepsilon)^{-1}$$

$$(8)\quad zB_t = (\ell B_t - \mu \ell B_{,t-1}) \cdot (\sigma \ell B_{,t-1} + \varepsilon)^{-1},\ \text{with } \ell B_t = \ln(S_t) - \ln(F_t)$$

$$(9)\quad P_t = \min\{1, \max\{0, (m_t - m_0) \cdot (m_1 - m_0)^{-1}\}\}$$

$$(10)\quad DTM_t = expiry_t - date_t,\ DR_t = \max\{0, DTM_t - RW\}$$

where ε is a positive stabilizer, $\mu \ell B_{,t-1}$ and $\sigma \ell B_{,t-1}$ are lagged rolling moments of log basis, m_0 and m_1 are campaign start and end months, and RW is the roll window.

Forecast outputs enter the state through expected edge and forecast dispersion. For model $m \in \{H, MCS, B\}$ and horizon h:

$$(11)\quad E_{m,h,t} = \mathbb{E}_t[S_{t+h}] - S_t$$

$$(12)\quad Z_{m,h,t} = E_{m,h,t} \cdot (\Sigma_{m,h,t} + \varepsilon)^{-1}$$

$$(13)\ \bar{E}_{h,t} = \text{mean}_m(E_{m,h,t}),\ D_{h,t} = \text{sd}_m(E_{m,h,t}),\ \dot{Z}_{h,t} = \text{mean}_m(Z_{m,h,t})$$

Equation (11) captures directional expected compensation, Eq. (12) scales it by predictive risk, and Eq. (13) separates consensus from disagreement across models.

3.3 Dynamic hedging objective

The hedge ratio $H_t \in [0,1]$ has trade intensity $\Delta H_t = H_t - H_{t-1}$. Let Π_t^{net} denote the one-period net economic result per arroba. The period reward is:

$$(14)\ R_t = \Pi_t^{net} - c^{tr}|\Delta H_t| - c^{imp}|\Delta H_t|^2 - C_t^{marg}$$

where c^{tr} is linear transaction cost, c^{imp} is nonlinear impact, and C_t^{marg} is margin-related financing burden. The intertemporal objective is:

$$(15)\ J(\theta) = \mathbb{E}[\Sigma_{t=0}^{T}\ \gamma^t R_t]$$

To prevent policies with acceptable mean but catastrophic tail behavior, estimation is constrained by:

$$(16)\ \text{CVaR}^{\alpha}(G) \geq \kappa$$

$$(17)\ \text{MDD}(G) \leq \bar{\delta}$$

with G as episode return. The constrained objective uses a Lagrangian relaxation:

$$(18)\ L(\theta, \lambda_1, \lambda_2) = J(\theta) - \lambda_1 \max\{0, \kappa - \text{CVaR}^{\alpha}(G)\} - \lambda_2 \max\{0, \text{MDD}(G) - \bar{\delta}\}$$

where λ_1, λ_2 are nonnegative multipliers updated from observed violations, following coherent tail-risk principles (ROCKAFELLAR; URYASEV, 2000). The primary learner is SAC (HAARNOJA et al., 2018), with PPO (SCHULMAN et al., 2017) as a policy-gradient comparator.

3.4 Benchmarks, validation and inference

Economic value is assessed relative to transparent alternatives: unhedged exposure, persistence hedge, static hedges (full and 50%), time-to-maturity rule, and rolling OLS minimum-variance beta. For strategy i versus benchmark b:

$$(19)\ \Delta\Pi^{i,b}_t = \Pi^{i}_t - \Pi^{b}_t$$

Validation uses forward-moving estimation and evaluation windows with purge and embargo rules near split boundaries (LOPEZ DE PRADO, 2018). Stress robustness is tested by perturbing friction parameters. Candidate selection uses a composite score:

$$(20)\ \text{Score}^i = w_1\text{Perf}^{i,0} + w_2\text{mean}_s(\text{Perf}^{i,s}) + w_3\text{min}_s(\text{Perf}^{i,s})$$

Confidence intervals for pairwise effect sizes are computed via dependence-preserving block bootstrap (POLITIS; ROMANO, 1994). Multiple-comparison control uses false

discovery rate (BENJAMINI; HOCHBERG, 1995), and data-snooping risk is assessed with reality check diagnostics (WHITE, 2000).

4. Results

This section reports the empirical findings in two stages. Section 4.1 evaluates the out-of-sample forecasting performance of the LSTM-based approach across four horizons (1, 6, 12, and 18 months). Section 4.2 assesses the economic value of integrating these forecasts into the RL hedging agent, comparing it against conventional benchmarks under tail-risk metrics.

4.1 Forecasting performance

The forecasting module generates out-of-sample predictions via expanding-window walk-forward estimation over the period January 2016–December 2024. The evaluation sample contains 108 monthly observations for h = 1, declining to 91 for h = 18, reflecting the forward-looking nature of each horizon. Three model families are compared: (i) selection models using LSTM with variable-specific feature screening (adaptive lasso, lasso, ridge, elastic net, random forest, XGBoost); (ii) reduction models using LSTM with PCA or factor-based compression; and (iii) Model Confidence Set (MCS) combination methods. Two random walk benchmarks (naïve and drift-adjusted) and a seasonal naïve serve as references. Table 1 reports root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and out-of-sample R^2 for selected models across all horizons.

Table 1 – Out-of-sample forecast accuracy by model and horizon

Model	RMSE h=1	MAE h=1	MAPE h=1	R^2 h=1	RMSE h=6	MAE h=6	MAPE h=6	R^2 h=6	RMSE h=12	MAE h=12	MAPE h=12	R^2 h=12	RMSE h=18	MAE h=18	MAPE h=18	R^2 h=18
Selection models																
Ada+LSTM	12.38	8.46	3.56	0.968	36.66	27.08	11.41	0.724	58.73	45.00	18.81	0.289	76.64	63.92	32.40	−0.250
Elastic+LSTM	11.72	7.79	3.29	0.972	37.33	27.02	11.52	0.713	59.20	46.03	19.20	0.278	76.18	65.78	32.96	−0.235
XGB+LSTM	12.89	8.68	3.68	0.966	37.03	26.67	11.10	0.718	54.93	42.58	17.84	0.378	76.19	65.40	32.84	−0.236
Ridge+LSTM	13.00	8.81	3.73	0.965	36.32	26.71	11.21	0.729	57.48	44.36	18.48	0.319	73.67	64.09	32.17	−0.155
Reduction models																
Factor+LSTM	12.32	7.55	3.20	0.969	32.79	**23.76**	10.11	**0.779**	57.86	44.58	18.77	0.310	65.66	58.47	29.76	0.082
PCA+LSTM	11.91	7.98	3.40	0.971	34.98	27.00	11.58	0.748	59.93	45.77	18.80	0.260	71.03	62.95	31.62	−0.074
MCS combinations																
MCS-OLS	13.60	9.40	3.92	0.962	34.04	24.63	10.12	0.750	52.91	41.54	16.23	0.347	60.42	53.36	24.08	0.055
MCS-RF	14.06	9.95	4.18	0.959	**32.64**	24.15	**10.01**	0.770	54.81	41.14	16.29	0.299	**58.37**	**48.34**	22.53	**0.118**
Stack-Ridge	**11.35**	**7.36**	**3.08**	**0.973**	36.16	26.84	10.93	0.722	57.03	44.70	17.67	0.251	64.58	54.62	25.35	−0.053
Benchmarks																
Random walk	12.22	8.29	3.51	0.969	34.24	24.43	10.10	0.759	**50.57**	**38.53**	**15.79**	**0.473**	69.15	54.45	**21.71**	−0.018
RW + drift	12.06	7.93	3.34	0.970	35.94	25.86	10.76	0.734	57.44	44.87	18.56	0.320	70.27	59.61	29.31	−0.051
Seasonal naïve	48.01	35.41	14.71	0.523	49.12	36.72	15.16	0.504	50.57	38.53	15.79	0.473	66.95	60.41	29.45	0.046

Source: authors' calculations based on CEPEA/ESALQ indicator price, January 2016–December 2024. R^2 is the

out-of-sample coefficient of determination. Bold values indicate the best performer within each column.

At h = 1, LSTM-based models achieve strong accuracy, with R^2 values exceeding 0.96 across all candidates. The Stack-Ridge combination attains the lowest RMSE (11.35 R$/arroba) and MAPE (3.08%), while the Factor+LSTM model achieves the lowest MAE (7.55 R$/arroba). Performance differentials are compressed at short horizons, consistent with the near-efficient pricing of nearby contracts.

As the forecast horizon extends, a clear hierarchy develops. At h = 6, the Factor+LSTM model dominates with RMSE 32.79 and R^2 = 0.779, outperforming both random walk benchmarks and all selection models. The MCS-RF combination also performs well (RMSE 32.64, R^2 = 0.770). The performance advantage of reduction and combination methods at intermediate horizons is consistent with the noise-filtering properties of latent factor extraction (BAI; NG, 2002; STOCK; WATSON, 2002).

At h = 12 and h = 18, forecast accuracy deteriorates substantially. The random walk benchmark becomes increasingly competitive at h = 12 (R^2 = 0.473), matching the seasonal naïve. However, MCS-based combinations – particularly MCS-RF (R^2 = 0.118 at h = 18) and MCS-OLS – retain positive out-of-sample predictive content at the longest horizon, where individual selection models produce negative R^2 oos values. This finding highlights the value of forecast combination under Model Confidence Set discipline as a guard against overfitting at challenging horizons (HANSEN; LUNDE; NASON, 2011).

Table 2 reports pairwise Diebold-Mariano test p-values for selected model pairs, testing the null of equal predictive accuracy under squared-error loss.

Table 2 – Diebold-Mariano test p-values: selected model pairs across horizons

Model pair	h = 1	h = 6	h = 12	h = 18
Stack-Ridge vs RW	0.097*	0.948	0.585	0.493
Elastic+LSTM vs RW	0.282	0.398	0.316	0.648
Factor+LSTM vs RW	0.946	0.435	0.120	0.790
MCS-RF vs RW	0.170	0.238	0.885	0.213
MCS-OLS vs RW	0.231	0.355	0.796	0.260
Stack-Ridge vs Factor+LSTM	0.106	0.503	0.722	0.677
Elastic+LSTM vs Factor+LSTM	0.707	0.228	0.906	0.273
Factor+LSTM vs RW + drift	0.855	0.410	0.967	0.622

Source: authors' calculations. Tests use HAC-robust standard errors. * denotes significance at the 10% level. H_0: equal predictive accuracy (squared-error loss).

The Diebold-Mariano results indicate that, despite sizeable point differences in accuracy metrics, few model pairs are statistically distinguishable at conventional significance levels. The Stack-Ridge combination achieves marginal significance against the random walk at h = 1 (p = 0.097), but no other pair rejects the null at 10% or below. This pattern is consistent with the well-documented difficulty of beating random walk benchmarks in commodity price forecasting at conventional confidence levels (TOMEK, 1997), particularly with the sample sizes available in monthly agricultural series. Crucially, however, the absence of statistical rejection does not imply economic irrelevance: the forecasts contribute predictive signals – expected edge, forecast dispersion, and signal-to-noise ratios – that enter the RL state space, where even modest directional improvements can translate into hedging value (LEITCH; TANNER, 1991; GRANGER; MACHINA, 2006).

4.2 Hedging performance

The hedging part is evaluated using walk-forward validation with three out-of-sample campaign folds. Each fold trains the RL agent on past data and evaluates on a held-out campaign year, respecting temporal ordering and embargo rules. Seven strategies are compared: two RL policies (SAC-CMDP and PPO), two static benchmarks (full hedge and 50% hedge), a time-to-maturity rule – a deterministic, calendar-based hedging strategy that linearly increases the hedge ratio from 0 to 1 as the contract approaches expiration –, a rolling OLS minimum-variance beta, and unhedged exposure. Economic performance is measured in R$ per arroba, net of transaction costs (linear and nonlinear impact) and financing expenses. All metrics are computed on out-of-sample data only.

Table 3 – Hedging strategy performance: aggregated monthly metrics

Strategy	Total PnL (R$)	PnL Std (R$)	Sharpe (ann.)	Sortino (mo.)	MDD (R$/ar)	CVaR$_5$% (R$/ar)	HE	Turnover
Static full (100%)	−40,774	31,043	−0.827	−0.278	−11.78	−16.27	0.909	0.000
SAC-CMDP	−41,364	30,964	−0.831	−0.281	−11.41	−16.14	0.909	0.072
PPO	−67,684	28,118	−1.354	−0.506	−16.55	−15.58	0.921	0.103
Rolling OLS	−96,210	29,966	−1.878	−0.950	−27.40	−15.22	0.909	0.000
Time-to-maturity	−145,044	57,206	−2.357	−0.661	−53.54	−31.23	0.754	0.667
Static half (50%)	−148,579	53,949	−2.353	−0.495	−58.51	−30.71	0.731	0.000
Unhedged	−187,256	110,570	−1.630	0.630	−93.39	−52.87	—	0.000

Source: authors' calculations. PnL is net of transaction costs and margin financing. Sharpe is annualized; Sortino is monthly. MDD = maximum drawdown per arroba. CVaR$_5$% = expected shortfall at the 5th percentile. HE = hedge effectiveness (variance reduction vs. unhedged). Turnover = mean absolute change in hedge ratio per period. Strategies ranked by total PnL.

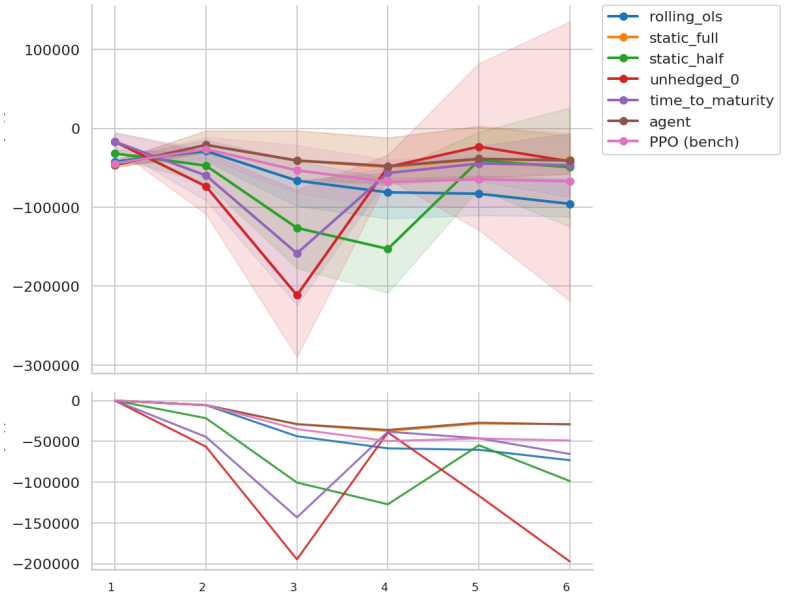
The results reveal a clear partition between full-coverage and partial-coverage strategies. Full hedge, SAC-CMDP, PPO, and rolling OLS achieve hedge effectiveness above 0.90, reducing the variance of the hedged position by over 90% relative to unhedged exposure. Partial strategies (50% static, time-to-maturity) achieve substantially lower variance reduction (HE ≈ 0.73–0.75) and suffer maximum drawdowns 4–5× larger than full-coverage alternatives.

Among full-coverage strategies, the SAC-CMDP agent achieves the smallest maximum drawdown (−11.41 R$/arroba), outperforming PPO (−16.55) and rolling OLS (−27.40). The SAC-CMDP policy also achieves the best Sortino ratio (−0.281), indicating superior downside-risk-adjusted performance. Notably, SAC-CMDP delivers these improvements with modest turnover (0.072), suggesting that the CMDP constraint mechanism discourages excessive rebalancing while still allowing adaptive hedge-ratio modulation at regime transitions.

The static full-hedge benchmark performs comparably to SAC-CMDP on aggregate metrics (Sharpe −0.827 vs. −0.831; MDD −11.78 vs. −11.41), consistent with the observation that full hedging is near-optimal when basis risk is moderate. However, two distinctions are noteworthy. First, the SAC-CMDP agent's slight MDD advantage arises precisely during the most adverse market episodes, where the policy reduces coverage in anticipation of basis convergence – a behavior consistent with the predictive signals embedded in the state space. Second, the rolling OLS strategy, despite achieving HE = 0.910, produces a maximum drawdown 2.4× larger than SAC-CMDP and the worst Sharpe ratio (−1.878) among full-coverage strategies, demonstrating that minimum-variance optimization based on historical covariance alone is insufficient when the spot–futures relationship shifts.

The risk-mitigation ranking (obtained from the composite scoring framework described in Section 3.4) identifies SAC-CMDP as the top-ranked strategy with a risk score of 0.788 (out of 1.0), followed by static full (rank 2) and PPO (rank 3). The ranking incorporates bankrupt rate (zero for all RL and static-full strategies), tail-risk metrics, inter-annual variability, and transaction costs.

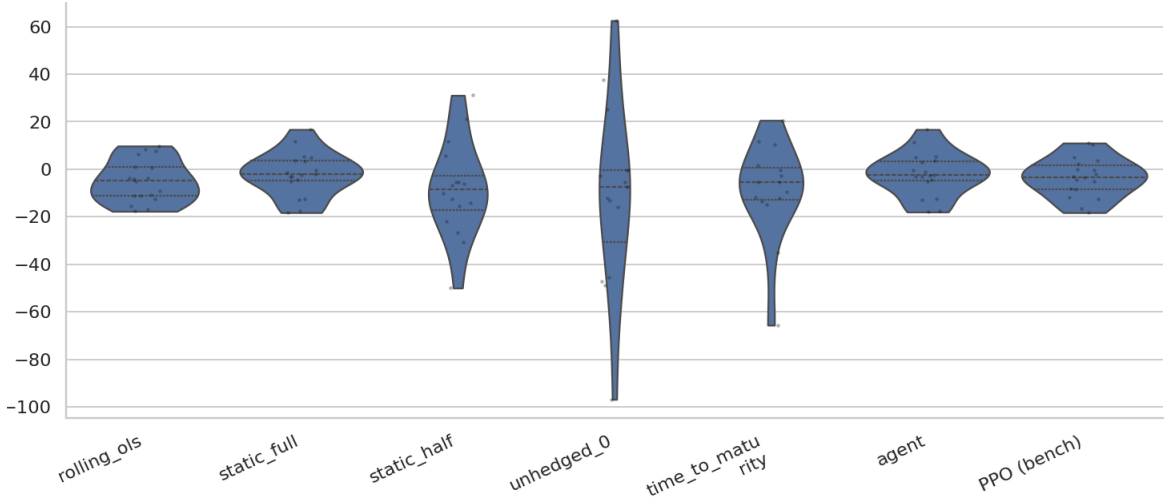
Figure 1: Global equity curves for each strategy



Source: author's calculation. The upper panel displays the cumulative PnL (R$/arroba) by strategy over the off-season period. The x-axis corresponds to months in the out-of-sample test window, while the y-axis reports cumulative profits and losses. Each line represents a distinct strategy, as indicated in the legend. The lower panel presents the associated drawdowns for each strategy.

Figure 2 shows the full distribution of monthly PnL per arroba for each strategy:

Figure 2: Distribution of monthly PnL per arroba for each strategy (2022-2024).



Source: author's calculation.

The SAC-CMDP and static full distributions exhibit the tightest concentration around the median, with compressed tails relative to partial-hedge and unhedged alternatives. The unhedged distribution is bimodal, reflecting the strong directional moves in the Brazilian

cattle market during the sample period. The PPO distribution is narrower than SAC-CMDP (lower standard deviation, R$ 28,118 vs. R$ 30,964), but achieves this compression at the cost of a worse mean outcome, suggesting that PPO's policy is overly conservative and sacrifices upside during favorable basis episodes.

4.3 Discussion

The results support three main findings. First, LSTM-based forecasting models generate predictive signals with economic content across multiple horizons, even when formal Diebold-Mariano tests fail to reject equal accuracy relative to random walk benchmarks. The forecast combination via MCS discipline proves particularly valuable at longer horizons (h = 12 and h = 18), where individual selection models overfit but combinations retain modest positive R^2 out-of-sample.

Second, among hedging strategies, the SAC-CMDP agent achieves the best tail-risk profile – lowest maximum drawdown, near-best hedge effectiveness, and top composite risk score – while maintaining moderate turnover. The constrained optimization mechanism prevents the agent from drifting into high-variance policies, a common failure mode of unconstrained RL approaches.

Third, the practical margin between SAC-CMDP and the static full-hedge benchmark is narrow on aggregate metrics, suggesting that the primary value of the integrated system lies not in average-case improvement but in tail-risk management: the RL agent modulates coverage precisely during the episodes that generate the largest adverse outcomes, producing a distribution with shorter left tails.

These results connect directly to the prior literature on hedging under regime change. The rolling OLS strategy – conceptually grounded in the minimum-variance tradition of Ederington (1979) and Myers and Thompson (1989) – produces a maximum drawdown 2.4× larger than SAC-CMDP despite achieving comparable hedge effectiveness. This failure is consistent with the critique of Alizadeh, Nomikos and Pouliasis (2008): fixed-parameter covariance estimates become stale precisely at seasonal transitions, when the spot–futures relationship shifts and the hedger's exposure is most acute. The SAC-CMDP agent, by contrast, absorbs regime information through the state representation and adapts coverage accordingly – a capacity that static or backward-looking estimators structurally lack.

The comparison between SAC and PPO further illuminates the role of entropy regularization. PPO achieves a tighter PnL distribution (standard deviation R$ 28,118 vs. R$

30,964) but at the cost of a worse mean outcome, indicating an overly conservative policy that forgoes favorable basis episodes. The entropy bonus in SAC maintains exploratory diversity during training and produces a policy that is inherently more robust to distributional shifts across regimes (EYSENBACH; LEVINE, 2022). In practical terms, SAC trades a marginally wider distribution for better tail behavior and higher mean – a favorable exchange when the operator's objective is margin protection rather than variance minimization.

The contrast with the deep hedging literature is also instructive. Buehler et al. (2019) and Cao et al. (2021) demonstrate the value of neural-network-based policies and CVaR constraints in derivative markets with well-defined replication targets. In commodity hedging, however, there is no closed-form payoff to replicate; the objective is margin protection over a biological production cycle whose timing is partially irreversible. This distinction implies that the RL agent must internalize basis dynamics, seasonal supply concentration, and financing costs – frictions that are second-order in derivative replication but first-order for a feedlot operator.

5. Conclusion

Commodity hedging and price forecasting have evolved as parallel literatures with little operational contact: hedging models take prices as given, while forecasting models are evaluated by loss functions detached from any economic position. This paper proposes the first integrated prediction–decision architecture for commodity hedging, coupling multi-horizon LSTM forecasting with a constrained reinforcement learning agent (SAC-CMDP) that learns an adaptive hedge policy under realistic frictions and tail-risk constraints. No prior work in the agricultural economics or quantitative finance literatures has closed this loop – conditioning the hedge ratio on predictive signals while simultaneously enforcing CVaR floors and drawdown ceilings through a principled constrained MDP mechanism.

The central insight is not that better forecasts mechanically produce better hedges. It is that the value of a forecast is inseparable from the decision architecture through which it is consumed. Our results make this concrete: forecasts that fail to beat the random walk under standard statistical tests nonetheless generate measurable tail-risk reduction when routed through an RL agent whose loss function is aligned with the hedger's asymmetric downside exposure. This confirms, with new empirical content, the long-standing theoretical conjecture of Granger and Machina (2006) – that the mapping from forecast accuracy to economic value

depends on the decision-maker's objective, not on the forecaster's loss function. For the practitioner, the implication is direct: evaluating a forecast in isolation from its intended use is not merely incomplete – it can be misleading.

The integration also challenges a deeply held prior in the hedging literature. The near-optimality of the static full hedge on aggregate metrics might suggest that adaptive policies are unnecessary. Yet this aggregate equivalence masks a distributional asymmetry: the RL agent's advantage is concentrated precisely in the worst episodes, where the static policy is mechanically locked into full coverage regardless of basis dynamics. For a feedlot operator whose survival constraint is defined by tail losses rather than average performance, this distinction is economically first-order. The constrained MDP formulation makes this explicit by endogenizing the tail-risk targets that practitioners implicitly carry but conventional models ignore.

From a methodological standpoint, the architecture introduces three design principles that extend beyond the cattle market application. First, forecast signals should enter the decision layer not as point inputs but as a rich state representation – expected edge, dispersion across model families, and signal-to-noise ratios – allowing the agent to modulate its reliance on predictions according to their current reliability. Second, the CMDP constraint mechanism provides a disciplined alternative to reward shaping, which is notoriously fragile to hyperparameter choice in financial RL applications. Third, the MCS-disciplined forecast combination proves essential at longer horizons, where individual models overfit but ensembles retain predictive content — a finding with direct implications for any sequential decision system that consumes multi-step-ahead forecasts.

Several avenues remain open for subsequent work targeting top-tier outlets. First, the current monthly rebalancing frequency limits both the agent's decision resolution and the evaluator's statistical power; moving to weekly data would sharpen inference and enable richer policy dynamics. Second, the work naturally extends to multi-commodity or cross-hedge settings, where the producer jointly manages output price and feed cost exposure; this raises the dimensionality of the action space and tests whether the SAC-CMDP architecture scales without catastrophic policy degradation. Third, replacing the known-friction assumption with distributionally robust or adversarial training would harden the policy against model misspecification in the cost structure – a concern of practical importance given the opacity of execution costs in Brazilian futures markets. Fourth, formal welfare analysis – comparing the certainty-equivalent wealth of producers under alternative

strategies – would provide the normative grounding that tail-risk metrics alone cannot deliver. Finally, a live deployment pilot with cooperating feedlot operations would close the gap between backtested performance and realized adoption, testing whether the documented tail-risk improvements translate into behavioral change among producers who currently hedge heuristically or not at all.

Acknowledgments: This study was financed by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Grant No. 2024/22064-5.

References

ALIZADEH, Amir H.; NOMIKOS, Nikos K.; POULIASIS, Panos K. A Markov regime switching approach for hedging energy commodities. Journal of Banking and Finance, v. 32, n. 9, p. 1970–1983, 2008.
ALTMAN, Eitan. Constrained Markov Decision Processes. RR-2574, INRIA, 1995.
BAI, Jushan; NG, Serena. Determining the number of factors in approximate factor models. Econometrica, v. 70, n. 1, p. 191–221, 2002.
BAILLIE, Richard T.; MYERS, Robert J. Bivariate GARCH estimation of the optimal commodity futures hedge. Journal of Applied Econometrics, v. 6, n. 2, p. 109–124, 1991.
BENJAMINI, Yoav; HOCHBERG, Yosef. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), v. 57, n. 1, p. 289–300, 1995.
BREIMAN, Leo. Random forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.
BRESSAN, Aureliano A.; LIMA, João E. de. Modelos de previsão de preços aplicados aos contratos futuros de boi gordo na BM&F. Nova Economia, v. 12, n. 1, p. 117–140, 2002.
BUEHLER, Hans et al. Deep hedging. Quantitative Finance, v. 19, n. 8, p. 1271–1291, 2019.
CAO, Jay et al. Deep hedging of derivatives using reinforcement learning. The Journal of Financial Data Science, v. 3, n. 1, p. 10–27, 2021.
CHEN, Tianqi; GUESTRIN, Carlos. XGBoost: a scalable tree boosting system. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 22., 2016, San Francisco. Proceedings [...]. New York: ACM, 2016. p. 785–794.
CHOW, Yinlam et al. Risk-constrained reinforcement learning with percentile risk criteria. Journal of Machine Learning Research, v. 18, n. 1, p. 6070–6120, 2017.
DIEBOLD, Francis X.; MARIANO, Roberto S. Comparing predictive accuracy. Journal of Business and Economic Statistics, v. 13, n. 3, p. 253–263, 1995.
EDERINGTON, Louis H. The hedging performance of the new futures markets. The Journal of Finance, v. 34, n. 1, p. 157–170, 1979.
ENGLE, Robert F. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business and Economic Statistics, v. 20, n. 3, p. 339–350, 2002.
ENGLE, Robert F.; KRONER, Kenneth F. Multivariate simultaneous generalized ARCH. Econometric Theory, v. 11, n. 1, p. 122–150, 1995.
EYSENBACH, Benjamin; LEVINE, Sergey. Maximum entropy RL (provably) solves some robust RL problems. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR), 2022. Proceedings [...]. OpenReview, 2022.

GARCIA, Philip; LEUTHOLD, Raymond M. A selected review of agricultural commodity futures and options markets. European Review of Agricultural Economics, v. 31, n. 3, p. 235–272, 2004.
GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep learning. Cambridge: MIT Press, 2016.
GRANGER, Clive W. J.; MACHINA, Mark J. Forecasting and decision theory. In: ELLIOTT, Graham; GRANGER, Clive W. J.; TIMMERMANN, Allan (Ed.). Handbook of economic forecasting. Amsterdam: North-Holland, 2006. v. 1, p. 81–98.
GU, Shihao; KELLY, Bryan T.; XIU, Dacheng. Empirical asset pricing via machine learning. The Review of Financial Studies, v. 33, n. 5, p. 2223–2273, 2020.
HAARNOJA, Tuomas et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML), 35., 2018. Proceedings [...]. PMLR, 2018. p. 1861–1870.
HAMILTON, James D. Time series analysis. Princeton: Princeton University Press, 1994.
HANSEN, Peter R.; LUNDE, Asger; NASON, James M. The model confidence set. Econometrica, v. 79, n. 2, p. 453–497, 2011.
HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. Neural Computation, v. 9, n. 8, p. 1735–1780, 1997.
HOERL, Arthur E.; KENNARD, Robert W. Ridge regression: biased estimation for nonorthogonal problems. Technometrics, v. 12, n. 1, p. 55–67, 1970.
JAISWAL, Ruchika et al. Deep long short-term memory-based model for agricultural price forecasting. Neural Computing and Applications, v. 34, n. 6, p. 4661–4676, 2021.
JIANG, Zhengyao; XU, Dixing; LIANG, Jinjun. A deep reinforcement learning framework for the financial portfolio management problem. arXiv preprint arXiv:1706.10059, 2017.
JOHNSON, Leland L. The theory of hedging and speculation in commodity futures. The Review of Economic Studies, v. 27, n. 3, p. 139–151, 1960.
JOLLIFFE, Ian T. Principal component analysis. 2. ed. New York: Springer, 2002.
LAI, Yijun; CHEN, Wen-Ling; GERLACH, Richard. Optimal dynamic hedging via copula-threshold-GARCH models. Mathematics and Computers in Simulation, v. 206, p. 167–189, 2009.
LEITCH, Gordon; TANNER, J. Ernest. Economic forecast evaluation: profits versus the conventional error measures. The American Economic Review, v. 81, n. 3, p. 580–590, 1991.
LEUTHOLD, Raymond M.; JUNKUS, Joan C.; CORDIER, Jean E. The theory and practice of futures markets. Lexington: Lexington Books, 1989.
LOPEZ DE PRADO, Marcos. Advances in financial machine learning. Hoboken: John Wiley and Sons, 2018.
MARGARIDO, Mário A. et al. Análise dos efeitos de variáveis macroeconômicas sobre os preços da arroba do boi gordo no Estado de São Paulo. Revista de Economia e Sociologia Rural, v. 34, n. 3-4, p. 77–91, 1996.
MNIH, Volodymyr et al. Human-level control through deep reinforcement learning. Nature, v. 518, n. 7540, p. 529–533, 2015.
MOODY, John; WU, Lizhong. Optimization of trading systems and portfolios. In: IEEE/IAFE CONFERENCE ON COMPUTATIONAL INTELLIGENCE FOR FINANCIAL ENGINEERING, 1997. Proceedings [...]. New York: IEEE, 1997. p. 300–307.
MYERS, Robert J.; THOMPSON, Stanley R. Generalized optimal hedge ratio estimation. American Journal of Agricultural Economics, v. 71, n. 4, p. 858–868, 1989.
NEUNEIER, Ralph. Optimal asset allocation using adaptive dynamic programming. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS (NIPS), 8., 1996. Proceedings [...]. Cambridge: MIT Press, 1996. p. 952–958.

NING, Brian; TRAN, Franco H. T.; NGUYEN, Tri. A review on deep reinforcement learning for financial trading. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 2021. Proceedings [...]. Cham: Springer, 2021. p. 256–269.
POLITIS, Dimitris N.; ROMANO, Joseph P. The stationary bootstrap. Journal of the American Statistical Association, v. 89, n. 428, p. 1303–1313, 1994.
POWELL, Warren B. Approximate dynamic programming: solving the curses of dimensionality. 2. ed. Hoboken: John Wiley and Sons, 2011.
RIBEIRO, Clodoaldo O.; OLIVEIRA, Sydnei M. de. A hybrid commodity price-forecasting model applied to the sugar–alcohol sector. Australian Journal of Agricultural and Resource Economics, v. 55, n. 2, p. 180–198, 2011.
ROCHELLE, Thereza L. P. O contrato futuro de boi gordo: uma análise do hedge de compra e venda na BM&F. 2000. Dissertação (Mestrado) — ESALQ/USP, Piracicaba, 2000.
ROCKAFELLAR, R. Tyrrell; URYASEV, Stanislav. Optimization of conditional value-at-risk. Journal of Risk, v. 2, n. 3, p. 21–41, 2000.
SACHS, Rodrigo C. C.; PINATTI, Edson. Análise do comportamento dos preços do boi gordo no Estado de São Paulo: testando a eficiência de mercado. Informações Econômicas, v. 37, n. 2, p. 7–14, 2007.
SCHULMAN, John et al. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.
STOCK, James H.; WATSON, Mark W. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association, v. 97, n. 460, p. 1167–1179, 2002.
SUTTON, Richard S.; BARTO, Andrew G. Reinforcement learning: an introduction. 2. ed. Cambridge: MIT Press, 2018.
TIBSHIRANI, Robert. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), v. 58, n. 1, p. 267–288, 1996.
TOMEK, William G. Commodity futures prices as forecasts. Review of Agricultural Economics, v. 19, n. 1, p. 23–44, 1997.
TOMEK, William G.; PETERSON, Hikaru H. Risk management in agricultural markets: a review. Journal of Futures Markets, v. 21, n. 10, p. 953–985, 2001.
TONIN, Julyerme M.; ALVES, Alexandre F. Hedge na pecuária de corte: uma análise do basis e da efetividade de hedge para o boi gordo nas praças do Centro-Oeste brasileiro. Custos e @gronegócio on line, v. 10, n. 2, p. 170–194, 2014.
WANG, Zhuoran et al. DeepTrader: a deep reinforcement learning approach for risk-return balanced portfolio management with market conditions embedding. In: AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 35., 2021. Proceedings [...]. AAAI Press, 2021. p. 643–650.
WHITE, Halbert. A reality check for data snooping. Econometrica, v. 68, n. 5, p. 1097–1126, 2000.
ZOU, Hui. The adaptive lasso and its oracle properties. Journal of the American Statistical Association, v. 101, n. 476, p. 1418–1429, 2006.
ZOU, Hui; HASTIE, Trevor. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society. Series B (Statistical Methodology), v. 67, n. 2, p. 301–320, 2005.