v. 4, n. 7, 2024

Journal of
Agricultural
Sciences
Research

PRICING MODELS FOR SALE OPTIONS ON CORN FUTURES CONTRACTS IN THE BRAZILIAN MARKET

Liane Rucinski
Ministry of Planning and Budget
Brasília - DF
http://lattes.cnpq.br/1386322531054517

Rafael Pazeto Alvarenga
``Universidade Federal de Mato Grosso do Sul``
Paranaíba - MS
http://lattes.cnpq.br/5467311850426022

Journal of Agricultural Sciences Research ISSN 2764-0973 DOI https://doi.org/10.22533/at.ed.973472431058

Abstract: In order to assist rural producers in managing market risk, the Federal Government's Minimum Price Guarantee Policy [PGPM] allows subsidizing premiums from sales option contracts in hedging operations for agricultural products. The objective of this study was to analyze the effectiveness of pricing models in evaluating out-of-the-money put options on corn futures contracts in the Brazilian market. The Black, Binomial and Least Squares Monte Carlo [LSM] models were tested, combined with historical, implicit and deterministic volatility forecasters. The premiums obtained by the different models were compared to those actually practiced in the market. The Black model, followed by the Binomial, both combined with implied volatility, presented the smallest deviations in relation to real market premiums, according to the mean absolute percentage error [MAPE] criterion, as well as the smallest dispersions, measured by the square root of the root mean square error [RMSE]. The LSM method underpriced options far out of the money when combined with any of the volatility estimators analyzed. When combined with historical volatility, the models under analysis proved to be less accurate and less precise. The deterministic volatility estimator of the Generalized Conditional Autoregressive Heteroscedasticity model [Garch] presented intermediate performance. The results corroborated the wide use of the Black model, which demonstrated the best performance in the precision and accuracy criteria, among the option pricing models analyzed, especially when associated with implied volatility.

Keywords: options, Binomial model, Black, Monte Carlo least squares, volatility.

INTRODUCTION

Market risk, arising from the volatility of agricultural prices, constitutes a source of uncertainty regarding the income of rural producers.

Market risk management can be carried out in "hedge" operations, when agents assume, in the financial market, a position opposite to that in the physical market, protecting themselves against adverse movements in asset prices. Hedge operations with agricultural derivatives are set up in the off-season, when expectations are formed, and the position is closed at harvest time.

In operations supported by the Minimum Price Guarantee Policy, referred to in Decree-Law Number 79, of 1966, Law Number 8,427, of 1991, provides for the granting of a bonus equivalent to a percentage of the premium paid on the acquisition of option contracts for sale on national or international commodity and futures exchanges.

The option contract is the instrument by which the holder transfers the price risk to the writer, upon payment of a premium, in order to ensure the right, but not the obligation, to buy ("call") or sell ("put") the asset at the established price, at a future date.

The definition of the minimum price considers factors that influence the domestic and foreign markets and production costs. According to the Theory of the Firm, revenue equivalent to the average variable cost ensures the continuity of production in the short term. The break-even point is reached when sales revenue covers the operational cost of production, which also includes fixed costs.

Corn, a "commodity" traded on commodity and futures exchanges, is also a typical family farming crop, used in animal feed and ethanol production. According to the United States Department of Agriculture (Wasde/USDA) Supply and Demand Report, from Mar. 2024, Brazil is the third world producer of corn, with a harvest of 116 million tons, in 2023/24, equivalent to 9.5% of the world total, behind only the United States and China.

BAIDYA and CASTRO (2001) analyzed the convergence of the Binomial model. ARAÚJO and BAIDYA (2004) evaluated the sensitivity of the Monte Least Squares method to the number of simulations, discretization and the regression function. GABE and PORTUGAL (2004) compared implicit and deterministic volatility forecasters. POON and GRANGER (2005) compiled 93 comparative volatility studies. SALIBY et al. (2007) evaluated descriptive sampling in reducing variance in Monte Carlo Simulation in option pricing. SAITO and ROCHMAN (2008) compared numerical options pricing methods.

COELHO et al. (2009) evaluated the performance of the Black Model in options on Arabica coffee futures. TONIN and COELHO (2012) tested numerical methods in pricing options on Arabica coffee futures. SOUZA et al. (2014) analyzed the term structure of the volatility of options on CME Group soybean futures. CHATEAU (2014) and MAIA et al. (2014) evaluated options under skewness and excess kurtosis. SILVA and MAIA (2014) evaluated volatility predictors in soybean futures contracts. MAIA et al. (2014) studied volatility smoothing using the Corrado and Su model. PONTES and MAIA (2017) tested Black's model in pricing options on cattle futures.

Aiming at formulating income guarantee policies for producers, studies are needed on the performance of pricing models for options on agricultural "commodities" traded in the Brazilian market.

The present study analyzed the adherence of options pricing models on agricultural product futures in the Brazilian market, compared to the real premiums observed in the market. Black's analytical model and the numerical methods of the Binomial Tree and Monte Carlo Least Squares were evaluated, combined with historical, implicit and deterministic volatility predictors.

The analysis was conducted on a put option on corn futures, American type, out-of-the-money, listed on B^3.

The study aims to answer the questions: i) Among the models analyzed, which provided the best estimator of market premiums? ii) Is there a significant difference between the theoretical premiums estimated by the models under analysis and the market premiums during the contracting period? iii) Which pricing model presented the best precision and accuracy indicators? iv) Which volatility estimator presented the best precision and accuracy indicators?

MATERIAL AND METHODS

Black's analytical model and the numerical models of the Binomial Tree and Monte Carlo Least Squares were tested, associated with the historical volatility estimators of the asset's log-returns, the instantaneous implied volatility of market premiums and the deterministic volatility of the Autoregressive model. of Conditional Heteroscedasticity [Garch].

The producing region considered was Londrina, in the North of Paraná, where corn is planted from September to December, and harvesting takes place from January to May. Therefore, the put option contract on corn futures expiring in March 2022, at harvest time, was chosen.

Considering the hedge operation lasting between 60 and 90 days, the contracting window was defined in the period of December 16th. 2021 to 15 january 15, 2022, in the initial phase of culture development.

The prices of future contracts, the Esalq corn indicator in the Campinas region, a reference for B^3, and the prices paid to producers in the North of Paraná were collected from the Cepea database for the last 12 months of the contract duration, between the second half of March 2021 and the first half of March 2022. These series are shown in figure 1:

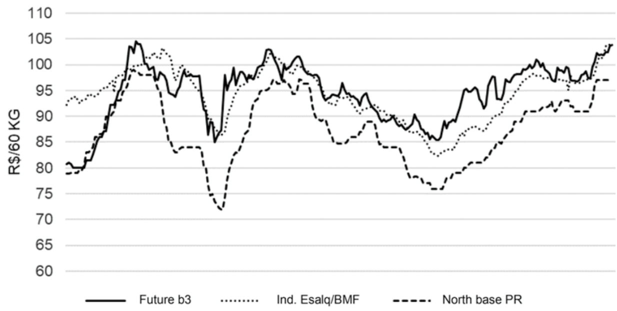

Figure 1: Corn grain prices between March 16, 2021 and March 15, 2022

Source: Original research data

The closing prices of premiums were collected from the Bloomberg platform in the 61 trading sessions throughout the duration of the contract, from December 16th. 2021 to 15 Mar. 2022, for the put option with an exercise price of R$95.00 on the corn futures contract expiring in March 2022 (CRDN2).

The future series were consolidated into a single continuous series, considering the price of the open contract with the nearest maturity, according to the procedure adopted by Souza (2013).

The series of market premiums for the put option with an exercise price of R$95.00 on the corn futures contract expiring in March 2022 is described in the graph shown in Figure 2:

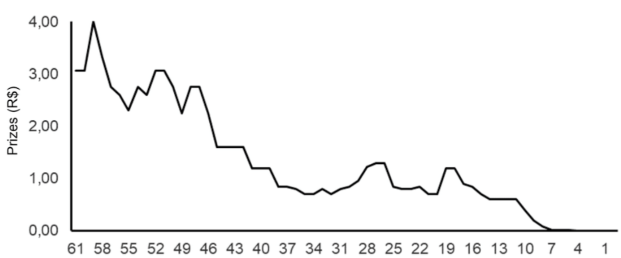

Figure 2: Corn put option premiums K=R$95.00 expiring in March. 2022

Source: Original research data

The log-return series of the underlying futures contract is obtained by the equation (1):

$$r_t = \ln \frac{P_t}{P_{t-1}} \quad (1)$$

where, r_t is the log-return of the asset underlying the option contract on the current date, ln designates the Neperian logarithm, and P_t and P_{t-1} represent the closing prices of the underlying futures on the current date and the previous business day, respectively.

Figure 3 presents the series of log-returns of the corn futures contract in the twelve months prior to expiration, in the period from March 16th. 2021 to March 15, 2022.

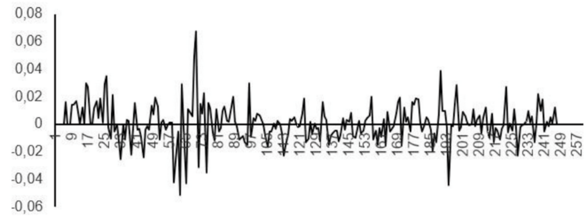

Figure 3: Log-returns of the B[3] corn futures contract expiring in March 2022

Source: Original research data

The base differential (basis) consists of the price discount in the producing region in relation to the spot indicator in the reference region for the Exchange, or the price of the futures contract with the nearest expiration, considered a "proxy" of the indicator in the region. of reference. This difference varies throughout the year between the harvest and off-season periods, which is called basis risk. The base differential is calculated by equation (2):

$$B = S - F \quad (2)$$

where, B is the base differential between the producing region and the reference region, S is the spot price in the producing region, and F is the quote of the futures contract with the nearest expiration or the product indicator in the reference region.

Journal of Agricultural Sciences Research ISSN 2764-0973 DOI https://doi.org/10.22533/at.ed.973472431058

The revenue to be guaranteed to the producer, corresponding to the strike price of the put option, must cover the production cost, plus the maximum basis premium, obtained by equation (3):

$$K = COP + |B| \quad (3)$$

where, K is the exercise price of the put option, COP is the operational cost of production, and B is the base differential between the producing region and the reference place for the exchange, where the goods will be delivered.

The basis was estimated by the maximum difference between the product prices in the North of Paraná and the Esalq corn indicator, surveyed in the Campinas region, a reference for the B[3], in the twelve months prior to the due date on March 15th. 2022. Thus, the base differential considered was R$16.93 per bag.

The operational cost of producing high-tech direct planting corn, according to a Conab survey in March 2022 in Londrina - PR, was R$77.72 per 60 kg bag. Adding the basis premium of R$16.93, the result is R$94.65, the price to be guaranteed to the producer in the operation. Therefore, the exercise price (K) of the put option was chosen as R$95.00 per bag.

This is an out-of-the-money option, with the exercise price being lower than the price of the underlying future during the contract period, December 16th. 2021 to january 15, 2022.

The risk-free discount rate adopted in the calculations was the reference for the Central Bank's Special Securities Settlement and Custody System - Selic, which varied from 9.25% to 10.75% throughout the contracting period.

PRICING MODELS

The Black model (BLACK, 1976) is derived from the well-known Black-Scholes model (BLACK and SCHOLES, 1973) for evaluating European options. However, instead of a cash asset, a futures contract appears as the underlying. Due to its simplicity, it has established itself as the option pricing model most used by traders around the world.

Black's model assumes the same simplifications as the original Black-Scholes model. Considers that underlying prices behave according to a geometric Brownian movement with log-normal distribution, that log-returns follow a normal distribution, independent and identically distributed [iid], assumes constant risk-free interest rate and volatility, exercise only at maturity, absence of arbitration, dividends and transaction costs.

However, the stylized facts of Black's model do not appear in practice. As found by JANKOVÀ (2018), the price series are not exactly log-normal and the log-returns deviate from the Gaussian curve, presenting asymmetry, heavy tails and a large number of "outliers".

By not recognizing variations in volatility over time, or the "smile" effect, which refers to the increase in volatility as the strike price moves away from the price of the underlying asset, the Black model tends to underprice options deeply within or out of money or long-term maturity. It also tends to underprice American options by not pricing the possibility of early exercise (HULL and WHITE, 1987), in addition to neglecting transaction costs and arbitration.

While in the original Black-Scholes model, only the exercise price is discounted to present value at the risk-free interest rate, in the Black model, both the exercise price and the quotation of the underlying futures contract are discounted to present value.

The call option premium is estimated by the Black model using equation (4):

$$c = e^{-rT}\left(FN(d_1) - KN(d_2)\right) \quad (4)$$

where, F is the quote of the underlying futures contract, K is the exercise price, T is term to maturity, r is the risk-free annual interest rate in continuous compounding, and N indicates the cumulative normal distribution function. The parameters d1 and d2 also depend on volatility: σ, single variable not directly observable. With d1 and d2 given by the equations (5) e (6):

$$d_1 = \frac{\ln(F/K) + (\sigma^2/2)T}{\sigma\sqrt{T}} \quad (5)$$

$$d_2 = \frac{\ln(F/K) - (\sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \quad (6)$$

where, N(d1) represents the "delta" of the call option, that is, the change in the premium depending on the change in the price of the underlying asset, and N(d2) is the probability of exercising the call option, or the probability that the price of the underlying asset exceeds the strike price.

The put option premium is obtained by the "put-call" parity, by equation (7):

$$p + Fe^{-rT} = c + Ke^{-rT} \quad (7)$$

where, F is the price of the underlying future, K is the exercise price, T is term to maturity, r is the risk-free annual interest rate in continuous compounding.

From equations (4) and (7), the premium of the put option at the same exercise price is given by the equation (8):

$$p = e^{-rT}\left(KN(-d_2) - FN(-d_1)\right) \quad (8)$$

where, F is the price of the underlying futures contract, K is the exercise price, T is term to maturity, r is the risk-free annual interest rate, N indicates the cumulative normal distribution function, and the parameters d1 and d2 are given by equations (5) and (6), respectively:

Numerical methods, such as Binomial and Monte Carlo Least Squares, allow pricing of American options, with the possibility of early exercise. However, they present a "tradeoff" between computational cost and convergence, which depends on the complexity of the base function, the discretization of time and the volume of simulations (ARAÚJO and BAIDYA, 2004).

The Binomial Model (1979) or CRR, by COX, ROSS and RUBINSTEIN, consists of a discrete-time approximation of the price movement of the underlying asset, which represents a stochastic process in continuous time.

The model assumes that the price of the underlying asset moves in discrete and uniform time intervals, up or down, according to a binomial distribution. Where p is the risk-neutral probability of the price moving upwards by a specific factor: *u*, and *1-p* is the probability that the price will move below a factor of d.

In the CRR model, the probability p is defined so that the binomial distribution simulates the geometric Brownian movement of the price of the underlying asset, by equation (9):

$$p = \frac{e^{-rt} - d}{u - d} \quad (9)$$

where *u* and *d* are the factors that move the price up and down, respectively, *r* is the risk-free interest rate in continuous capitalization and t is the discrete time interval, measured in years of 252 business days.

The upward movement factor u is given by equation (10):

$$u = e^{\sigma\sqrt{t}} \quad (10)$$

where, σ is the volatility and t are the discretization interval.

The downward movement factor d is the inverse of u, given by the equation (11):

$$d = \frac{1}{u} \quad (11)$$

where u is the upward movement factor.

The binomial tree is formed by "nodes" representing the underlying prices at each discrete time interval, between the valuation date and the option expiration, in each of the possible asset price trajectories.

Figure 4 represents a binomial tree of three-time frames and the prices of the underlying asset at each node:

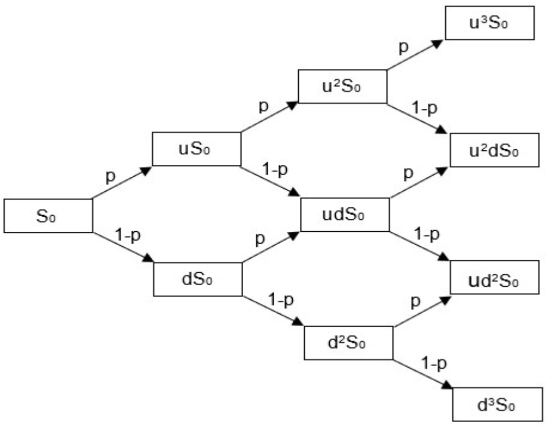

Figure 4: Binomial Tree

Source: Original research data

The recombinant property of the binomial tree, when an upward movement followed by a downward movement is equivalent to a movement downward and then upward, at the same intensities, allows the price to be estimated at each node, in each price trajectory, starting from the valuation date to maturity ("forward"), by equations (12) or (13), respectively, depending on whether the underlying is a spot asset or futures contract:

$$S_n = S_0 u^{nu} d^{nd} \quad (12)$$

$$F_n = e^{-rtn} F_0 u^{nu} d^{nd} \quad (13)$$

where, S_0 or F_0 is the price of the underlying spot asset or futures contract, respectively, on the valuation date, S_n *or* F_n is the price of the underlying spot asset or futures contract at the node considered, *u* and *d, they* are the factors that move the price up or down, respectively, *nu* and *nd, they* refer to the number of price movements up and down, respectively, to the node considered, r is the risk-free interest rate, t is the discretization interval and n designates the discretization interval.

The expected value of the option, at each node, is calculated moving recursively ("backward"), from the final nodes at expiration, until it converges at the first node at the valuation date.

The expected value of the option at each final node, at expiration, will be the intrinsic value, given by equations (14) or (15), for a call or put option, respectively, or zero, if the option ends out of the money:

$$cn = Max(Fn - K, 0) \quad (14)$$

$$pn = Max(K - Fn, 0) \quad (15)$$

where, *K* is the exercise price and F_n, is the price of the underlying at the respective node.

At nodes prior to expiration, the expected value of the American option is the greater of the immediate exercise value and the binomial (or continuation) value, or zero, if the option is out of the money, given by equations (16) or (17):

$$Cn = \text{maximum} \ (cn, Bn, 0) \quad (16)$$

$$Pn = \text{maximum} \ (pn, Bn, 0) \quad (17)$$

where, C_n and P_n designate the value of the call or put option, respectively, at the node considered, c_n and p_n are the immediate exercise values of the call or put option, at the node and B_n is the binomial value at the node.

At each node, the binomial value is recursively calculated, which results from the average of the values of the two subsequent nodes weighted by the respective probabilities p and 1-p, according to equations (18) or (19):

$$Bn = e^{-rtn}\left(pC_{n+1_{up}} + (1-p)C_{n+1_{down}}\right) \quad (18)$$

$$Bn = e^{-rtn}\left(pP_{n+1_{up}} + (1-p)P_{n+1_{down}}\right) \quad (19)$$

where, B_n is the binomial value of the option at the node, *r* is the risk-free interest rate in continuous capitalization, t is the discretization interval, (one day, converted to years of 252 business days), p and 1-p are the probabilities that the price of the underlying asset will rise or fall, respectively, C_{n+1up} and $C_{n+1down}$ are the expected values of the call option at subsequent nodes if the price moves up or down, and P_{n+1up} and $P_{n+1down}$ are the expected values of the put option at subsequent nodes if the price rises or falls, n designates the discretization interval.

For European options, as early exercise is not permitted, the expected value in nodes prior to expiration will always be the binomial value.

The Least Squares Monte Carlo Method, "Least Square Monte Carlo" [LSM], proposed by LONGSTAFF and SCHWARTZ (2001), is a discrete-time approximation of the continuous-time Brownian movement of underlying asset prices. The method predicts, using least squares regressions, the continuation value of an American option in the moments before expiration, comparing it to the exercise value at that moment.

The use of larger discretization intervals, simplified polynomial functions and a smaller volume of simulations, in order to reduce the computational cost, harm the convergence of the LSM method (ARAÚJO and BAIDYA, 2004).

The price trajectories of the underlying future, according to a Geometric Brownian Movement, were obtained by Monte Carlo Simulation, using equation (20):

$$F_{n+1} = F_n e^{\left(\left(\mu - ½\sigma^2\right)t + \sigma\varepsilon\sqrt{t}\right)} \quad (20)$$

where, F_{n+1} and F_n are the prices of the underlying asset on the date: $n+1$, and the current date n, respectively, μ is the average of the log-returns over the period, σ is the volatility, t is the discretization interval (one day, converted into years of 252 working days), and ε is a probability belonging to the normal distribution, represented by a random number between 0.001 and 0.999.

Ten thousand price trajectories of the underlying futures contract were simulated. According to the law of large numbers, the Monte Carlo method requires a large volume of simulations for better convergence. However, through sensitivity analysis, ARAÚJO and BAIDYA (2004) concluded that ten thousand price trajectories would be sufficient for convergence, and the subsequent increase in the number of simulations did not result in a significant improvement in the model's accuracy.

Aiming to improve convergence, the technique of variance reduction by antithetical variables was used, through pairs of negatively correlated random numbers, belonging to the same probability distribution.

Therefore, ε is a random number between 0 and 1, the antithetical number will be: $1-\varepsilon$.

The optimal exercise, at time n, on path i, only occurs if the option is in the money and if the value of the anticipated exercise exceeds the expected value of continuation, conditioned on the fact that the option has not yet been exercised. The method is applied recursively, from due date to valuation date.

At expiration, the value of the option in each path i, in which it is exercised, will be the intrinsic value, or zero, if the option ends out of the money, according to equations (21) or (22), for a call or call option. sale, respectively:

$$c_{ni} = Max(F_{ni} - K, 0) \quad (21)$$

$$p_{ni} = Max(K - F_{ni}, 0) \quad (22)$$

where, c_{ni} or p_{ni} is the premium of the call or put option, at expiration, in path i, F_{ni} is the price of the underlying futures contract, at expiration, on path i, K is the exercise price, n identifies the discretization interval, and i identifies the price path of the underlying asset.

In the moments before expiration, the value of the option, at an instant n of a path i, will be the early exercise value, if this is greater than the expected continuation value, or, otherwise, the continuation value, or zero, if the option is out of the money at time n and at subsequent times on path i, given by equations (23) or (24), for call or put options, respectively:

$$c_{ni} = \text{maximum } (0; Se((F_{ni} - K) > E[Y_{ni}|X_{ni}]; (F_{ni} - K);\ e^{-rt}c_{(n+1)i})) \quad (23)$$

$$p_{ni} = \text{maximum } (0; Se((K - F_{ni}) > E[Y_{ni}|X_{ni}]; (K - F_{ni});\ e^{-rt}p_{(n+1)i})) \quad (24)$$

where, $(F_{ni}$ - $K)$ ou (K - F_{ni}) represents the exercise value of the call or put option, at time n of path i, F_{ni} is the price of the underlying future at time n of the way: *i*, *K* is the exercise price, r is the risk-free interest rate in continuous compounding, t is the discretization interval (one day, converted into years of 252 business days), $e^{-rt}c_{(n+1)i}$ or $e^{-rt}p_{(n+1)i}$ is the continuation value of the call or put option, respectively, $c_{(n+1)i}$ or $p_{(n+1)i}$, represents the value of the option at time n+1, on path i, *n* indicates the discretization range and $E[Y_{ni}|X_{ni}]$ is the expected value of continuation at time n, on the way: *i*.

The vector of continuation expected values $E[Y_n|X_n]$ at time n, is obtained by least squares polynomial regression of the vector of continuation values: Y_n, depending on the price vector of the underlying asset X_n.

The regression is performed only on the paths in which the option is in the money at time n. The convergence of the method depends on this.

The vector of the prices of the underlying future, at time n, is given by equation (25):

$$\boldsymbol{X}_{ni} = [F_{ni}]_{i=1}^{m} \quad (25)$$

where, $\boldsymbol{X}_n$ is the vector of underlying prices at time n, and F_{ni} are the coefficients of the underlying prices at time n, on the way: *i*, with *i* ranging from one to *m*.

The vector of Continuation values: $\boldsymbol{Y}_n$ is formed by the coefficients of future cash flows discounted at time n, by equations (26) or (27), for call or put options, respectively:

$$\boldsymbol{Y}_{ni} = [e^{-rt}c_{(n+1)i}, 0]_{i=1}^{m} \quad (26)$$

$$\boldsymbol{Y}_{ni} = [e^{-rt}p_{(n+1)i}, 0]_{i=1}^{m} \quad (27)$$

where, *r* is the risk-free interest rate in continuous capitalization, *t is the discretization interval in years*, $c_{(n+1)\,i}$ or $p_{(n+1)i}$, is the value of the call or put option at the next moment: *n+1* on the way: *i*, *n* designates the number of the discretization interval, and *i* is the asset's price path, ranging from 1 to m.

In the regression, a third-degree polynomial was used, defined as a function of the moment in which the coefficient of determination: R^2 stopped growing. As a rule, the higher the polynomial degree of the regression, the better the precision, or the lower the standard deviation of the theoretical premiums (ARAÚJO and BAIDYA, 2004).

The coefficients of the expected continuation values at time n on each path: *i*, were calculated using the regression equation of Y_n in function of X_n. They are only used in regression, the ways: *i* in which the option is in the money at time n, proceeding recursively until the valuation date.

At each time n, the option's expected premium is calculated by taking the average of the option's value across all paths i, ranging from 1 to *m*, by equations (28) or (29), for call or put options, respectively:

$$c_n = \text{average } (c_{ni})i = 1 \ldots\ldots m \quad (28)$$

$$p_n = \text{average } (pn_i)i = 1 \ldots\ldots m \quad (29)$$

where, c_n or p_n is the value of the call or put option at the moment: *n*, c_{ni} or p_{ni} is the value of the option at the moment: *n* on the way: *i*, obtained by equations (23) and (24).

VOLATILITY PREDICTORS

The volatility of financial time series varies according to a term structure, generally decreasing with the reduction of uncertainty as maturity approaches.

Volatility presents an asymmetrical behavior, greater when prices are falling and lower when they are rising, in addition to forming "clusters" (conglomerates), alternating periods of greater or lesser volatility depending on market stability. In the case of options, there is also the "smile" effect, characterized by an increase in the volatility of premiums as the exercise price moves away from the price of the underlying asset.

Due to its simplicity, the most used future volatility predictor is historical volatility, estimated by the standard deviation of log-returns between the valuation date and maturity. However, historical volatility does not capture the term structure or the smile effect.

Old data is less relevant for predicting the future behavior of financial series. Thus, in each trading session, historical volatility was estimated based on the business days remaining until expiration, multiplying by the root of 252 to obtain the annualized volatility, using equation (30):

$$\sigma = \sqrt{\frac{\sum_{t=1}^{n}(r_t-\bar{r})^2}{n-1}}\sqrt{252} \quad (30)$$

where, σ is the annualized historical volatility, *t* is the trading date, ranging from 1 to *n*, *n* is the number of days left until due date, r_t *is the log-return of the asset on the trading date*, and $\bar{r}$ is the average of the log-returns over the period considered.

The instantaneous implied volatility of the option premium incorporates all available information and current market expectations about the behavior of premiums at the next moment, but little explanatory power over the option's maturity period. (GABE and PORTUGAL, 2004).

Due to the non-linearity of Black's formula, it is not possible to invert it in order to analytically obtain the implied volatility of the option premium, requiring the use of numerical methods. (Turitto, 2013).

The implied volatility of the option premium was estimated, by approximation, by the Newton-Raphson iterative algorithm, based on the instantaneous volatility of the market premium at the previous moment and the sensitivity of the option premium to changes in volatility, designated as "Vega"., by the equation (31):

$$\sigma_{i+1} = \sigma_i + \frac{P_m - P_{\sigma i}}{\upsilon_{\sigma i}} \quad (31)$$

where, σ_i e σ_{i+1} are the implied volatilities in the last iteration and the next iteration, P_m is the premium observed in the market, $P_{\sigma i}$ is the premium calculated by Black's equation based on the volatility in the previous instant (σ_i), $\upsilon_{\sigma i}$ is the sensitivity of the option premium to changes in volatility, called "Vega", given by equation (32):

$$\upsilon_{\sigma i} = Fe^{-rt}\sqrt{t}\,N'(d1) \quad (32)$$

where, F is the price of the underlying futures, r is the risk-free interest rate in continuous compounding on an annual basis, t is the unit of time in years, and $N'(d_1)$ is the first derivative of d_1, calculated by the equation (33):

$$N'(d_1) = \frac{e^{\frac{-d_1^2}{2}}}{\sqrt{2\pi}} \quad (33)$$

where, *d1* is obtained from Black's Formula, by equation (5).

The initial estimate of implied volatility (σ_0^2) was obtained from the volatility term structure, as a function of the exercise price (K), on the Bloomberg platform.

Three iterations were performed, considering the rapid convergence of the Newton-Raphson algorithm.

The Autoregressive Conditional Heteroscedasticity model [Garch], proposed by ENGLE (1982) and generalized by BOLLERSLEV (1986), describes the conditional, instantaneous, time-dependent volatility, different from the unconditional volatility, estimated based on the entire series of returns.

The Garch model captures stylized facts from the log-return series, such as volatility clusters, non-Gaussianity, stationarity, linear independence and quadratic autocorrelation, giving greater weight to recent observations. Extensions to the model describe volatility asymmetries (Egarch), or non-stationarity of the data series (Igarch).

The Garch (p, q) model admits that the log-return series is stationary, considers the mean, variance and autocorrelation to be constant, without trends or seasonality, and assumes that the covariances depend only on the lag between observations.

The Garch (1,1) model, with a lag for both the quadratic log-returns (q) and the autoregressive term (p), has proven to be sufficient to explain the volatility of most financial series.

The conditional variance of the Garch model (1,1) is obtained by the equation (34):

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (34)$$

where, σ_t^2, is the conditional variance at the current time, r^2_{t-1}, is the first lag of the quadratic log-return, α is the coefficient of the moving average term, σ^2_{t-1} is the first lag of the conditional variance, β is the coefficient of the autoregressive term and ω is the weighted term, given by the equation (35):

$$\omega = \Upsilon V_L \quad (35)$$

where, V_L is the unconditional, long-run variance, and ϒ is the conditional variance weighting coefficient.

Due to the stationarity of log-returns, the sum of the weighting parameters of the Garch model (1,1) corresponds to unity, according to equation (36):

$$\alpha + \beta + \Upsilon = 1 \quad (36)$$

where, α is the coefficient of the moving average term, β is the coefficient of the autoregressive term and γ is the weighting coefficient of the unconditional variance term in the Garch (1,1) model.

From equations (35) and (36), the unconditional variance is obtained by equation (37):

$$V_L = \frac{\omega}{1-\alpha-\beta} \quad (37)$$

where, α is the coefficient of the moving average term, β is the coefficient of the autoregressive term and ω is the unconditional variance term.

The parameters: α, β and ω are adjusted by likelihood, so that the calculated values and those extracted from the conditional variance: σ_t^2, results as close as possible.

The known elements: α and β, the parameter ϒ is obtained by equation (36).

The likelihood function, to be maximized, is given by the equation (38):

$$L_{\omega,\alpha,\beta} = \frac{1}{\sigma_t\sqrt{2\pi}} \exp\left(\frac{r_t^2}{2\sigma_t^2}\right) \quad (38)$$

where, σ_t is the conditional volatility, σ_t^2 is the conditional variance and r^2_t is the quadratic log-return of the underlying asset.

According to GABE and PORTUGAL (2004), conditional volatility σt is a "one step ahead" predictor of instantaneous volatility. Annualized conditional volatility is obtained: σ_{ty}, by the equation (39):

$$\sigma_{ty} = \sigma_t\sqrt{252} \quad (39)$$

where σ_t is the conditional volatility. And the unconditional, long-term, annualized volatility is obtained by the equation (40):

$$\sigma_y = \sqrt{252 V_L} \quad (40)$$

Where V_L, is the unconditional, long-run variance.

PERFORMANCE METRICS

Performance metrics allow classifying different models according to accuracy and precision criteria.

Accuracy refers to the position deviation between the estimated values in relation to the observed values, measured by the Mean Absolute Error [MAE] and the Mean Absolute Percentage Error [MAPE], defined by equations (41) and (42), respectively:

$$MAE\ (y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (41)$$

$$MAPE\ (y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| 100\% \quad (42)$$

where, y_i *is the observed value*, $\hat{y}_i$ *is the estimated value*, e n is the number of observations.

Accuracy measures the dispersion between estimated and observed values, measured by the Root Mean Square Error [RMSE], given by the equation (43):

$$RMSE\ (y, \hat{y}) = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(y_i - \hat{y}_i)^2} \quad (43)$$

where, y_i *is the observed value*, $\hat{y}_i$ *is the estimated value, and* n is the number of observations.

RESULTS AND DISCUSSION

Table 1 presents the observed premiums and the theoretical premiums estimated by the models under analysis, throughout the contracting period:

Table 2 shows the results of the tests applied to the estimated premiums:

The Shapiro-Wilk and Jarque-Bera tests rejected the normality of the log-return series for a significance level "α" of 0.05. The non-Gaussianity of log-returns is not captured by the Black model, but only by the deterministic volatility models of the Garch family.

The Durbin-Watson test rejected the hypothesis of independence of the log-return series, for a significance level "α" of 0.05. The Ljung-Box test indicated weak linear dependence of log-returns, weak significant first-order autocorrelation and absence of significant higher-order autocorrelation. These conditions are compatible with the application of the Garch model.

The Bartlett test rejected the hypothesis of constant variance of log-returns, indicating the heteroscedasticity of the series over time. Constant variance is a simplification assumed by the Black model and historical volatility, which gives equal weight to all observations. The term structure of volatility is captured by implicit, or instantaneous, and deterministic volatility estimators, which give greater weight to recent observations.

The Augmented Dickey-Fuller [ADF] test rejected the presence of a unit root, concluding that the series was stationary. According to PINHO et al. (2017), log-return series are generally stationary, although price series are normally non-stationary. If the assumption of stationary returns was violated, the use of the integrated Igarch extension would be required.

The Shapiro-Wilk and Jarque-Bera tests rejected the hypothesis of normality of premiums observed in the twelve months prior to maturity. The normality hypothesis was rejected for the Black models combined with historical volatility, Binomial combined with implied volatility and LSM combined with deterministic volatility.

The Bartlett test rejected the homogeneity of variances between the series of observed real premiums and theoretical premiums estimated by the Black, Binomial and LSM models, combined with historical, implicit and deterministic volatilities, in all cases.

If the assumptions of normality and homogeneity of variances are violated, parametric tests such as Student's "t" and Snedecor's "F" lose significance, requiring the application of non-parametric tests.

The Mann-Whitney and Kruskal Wallis

Date	Real prize	Black			Binomial			LSM		
		σ_H	σ_{VI}	Garch	σ_H	σ_{VI}	Garch	σ_H	σ_{VI}	Garch
1	3,07	4,61	3,45	4,54	1,43	3,09	2,09	0,00	0,00	0,00
2	4,00	5,31	3,92	5,27	1,44	2,91	2,00	0,03	0,00	0,03
3	3,33	4,87	3,52	4,92	1,43	2,71	1,93	0,07	0,00	0,13
4	2,75	2,98	1,26	2,99	1,42	2,52	1,88	0,12	0,02	0,21
5	2,60	2,30	2,30	3,31	1,41	2,31	1,84	0,17	0,19	0,22
6	2,30	1,94	2,21	2,60	1,39	2,17	1,82	0,27	0,19	0,29
7	2,75	2,00	2,35	2,47	1,38	2,10	1,80	0,27	0,19	0,36
8	2,60	3,89	4,60	4,20	1,38	2,17	1,83	0,36	0,26	0,58
9	3,07	3,44	2,64	5,04	1,36	2,37	1,90	0,36	0,26	0,59
10	3,07	3,52	3,15	4,64	1,36	2,57	2,00	0,42	0,26	0,74
11	2,75	2,78	2,29	3,65	1,33	2,61	2,13	0,42	0,29	0,75
12	2,75	1,58	1,57	2,41	1,35	2,66	2,24	0,45	0,48	0,81
13	2,75	1,54	2,35	2,80	1,31	2,70	2,23	0,48	0,48	0,90
14	2,75	1,59	2,77	2,60	1,26	2,70	2,22	0,50	0,48	1,21
15	2,75	1,29	2,39	2,11	1,25	2,67	2,19	0,53	0,63	1,22
16	1,60	1,10	2,03	1,85	1,24	2,67	2,17	0,60	0,63	1,19
17	1,60	1,07	1,58	1,74	1,23	2,69	2,16	0,60	0,63	1,20
18	1,60	1,09	1,61	1,70	1,21	2,71	2,12	0,85	0,73	1,21
19	1,60	1,08	1,59	1,64	1,19	2,65	2,09	0,85	0,73	1,22
20	1,20	0,82	1,28	1,29	1,16	2,58	2,04	0,85	0,77	1,23
μ	2,49	2,44	2,44	3,09	1,33	2,58	2,03	0,41	0,36	0,70
Median	2,68	1,97	2,32	2,70	1,35	2,65	2,07	0,42	0,28	0,75
σ	0,70	1,41	0,91	1,28	0,09	0,25	0,15	0,26	0,26	0,45
CV	0,28	0,58	0,37	0,41	0,07	0,10	0,07	0,64	0,73	0,65

Table 1: Premiums estimated by pricing models and volatility estimators
Source: Original Survey Results
Note: Put option on corn futures. March 2022.
Exercise price: K = 95.

Shapiro Wilk and Jarque Bera				Mann Whitney and Kruskal Wallis			
H_0 normal H_1 not normal	Black	Binomial	LSM	H_0 medians = H_1 medians ≠	Black	Binomial	LSM
History	H_1	H_0	H_0	Histórica	H_0	H_1	H_1
Implicit	H_0	H_1	H_0	Implícita	H_0	H_0	H_1
Garch	H_0	H_0	H_1	Garch	H_0	H_0	H_1

Table 2: Hypothesis tests: estimated premiums inrelation to observed premiums.
Source: Original research results

non-parametric tests rejected the hypothesis of equality of medians between the series of premiums estimated by the LSM model, combined with historical, implicit and Garch volatility estimators, and the series of premiums observed in the contracting period, with significance "α" of 0.05. The same occurred in relation to the Binomial model, combined with historical volatility. The equality of the medians was not rejected for the theoretical premiums of the Black model, combined with any of the volatility estimators, and Binomial, associated with Garch volatilities.

Table 3 presents the performance metrics of the pricing models combined with volatility models, compared to the observed premiums:

Table 4 classifies the pricing models combined with volatility estimators according to the accuracy and precision criteria.

Black's model, combined with historical and implied volatility estimators, presented the best performance in the accuracy and precision criteria. With deterministic volatility, it achieved the best performance in the precision criterion and intermediate in accuracy.

TONIN and COELHO (2012) verified the good performance of the Black model in comparison to numerical methods in evaluating options with different maturity periods and degrees of "moneyness". COELHO et al (2009) concluded that the Black model was adhered to in evaluating out-of-the-money options.

The Binomial model, combined with historical or implied volatility estimators, showed intermediate performance in the accuracy and precision criteria. With deterministic volatility, it surpassed the Black and LSM models in the precision criterion, with intermediate performance in terms of accuracy.

COELHO et al (2009) and SAITO and ROCHMAN (2008) verified good results regarding accuracy and precision indicators for the Binomial model.

The LSM method underpriced the premiums, in any case, and performed poorly in the accuracy and precision criteria, with the historical, implied and Garch volatility estimators. In the case under analysis, the low temporal discretization, the low number of simulations and the simplified polynomial function imposed severe conditions on the use of the LSM method.

As the LSM method uses for least squares regressions only the price trajectories in which the option is in the money, its convergence was impaired in the case under analysis, due to the reduced number of remaining trajectories. Conversely, SAITO and ROCHMAN (2008) confirmed the good performance of the method, despite the high computational cost.

The simplifications assumed by pricing models undermine their adherence. The hypothesis of fully efficient markets does not arise in practice, there is always asymmetry of information and arbitrage and theoretical risk-neutral interest rates normally differ from those practiced by agents depending on the risk of the underlying asset.

Table 5 classifies volatility estimators according to accuracy and precision criteria.

Implied volatility, associated with the Black and Binomial models, resulted in the best accuracy and precision indicators, but presented poor performance when applied to the LSM method.

POON and GRANGER (2005) concluded that implied volatility is the best predictor of future volatility. GABE and PORTUGAL (2004) and SOUZA et al (2014) verified the efficiency of implied volatility in short-term forecasting. TONIN and COELHO (2012) found the good performance of implied volatility compared to historical volatility.

Model	Volatility	MAE	MAPE	RMSE
		R$	%	R$²
Black	History	0,751	31,124	0,871
Black	Implicit	0,396	15,773	0,631
Black	Garch	0,661	23,437	0,922
Binomial	History	1,188	44,126	1,323
Binomial	Implicit	0,551	26,392	0,673
Binomial	Garch	0,738	28,407	0,859
LSM	History	2,104	79,819	2,279
LSM	Implicit	2,147	81,604	2,333
LSM	Garch	1,793	64,747	2,079

Table 3: Performance of pricing and volatility models

Source: Original Survey ResultsNote: Put option on the corn futures contract with expiration date in March 2022. (CRDN2)

Accuracy: MAE / MAPE				Accuracy: RMSE			
Volatility	Black	Binomial	LSM	Volatility	Black	Binomial	LSM
History	1	2	3	History	1	2	3
Implicit	1	2	3	Implicit	1	2	3
Garch	1	2	3	Garch	2	1	3

Table 4: Performance of pricing models

Source: Original Research Results

Rating 1: GoodNote 2: Intermediate

Grade 3: Unsatisfactory

Accuracy: MAE / MAPE				Accuracy: RMSE			
Prizes	History	Implicit	Garch	Prizes	History	Implicit	Garch
Black	3	1	2	Black	2	1	3
Binomial	3	1	2	Binomial	3	1	2
LSM	2	3	1	LSM	2	3	1

Table 5: Performance of volatility estimators

Source: Original Research Results

Rating 1: Good

Note 2: Intermediate

Grade 3: Unsatisfactory

The deterministic volatility of the Garch model applied with the LSM method showed good performance in accuracy and precision, and intermediate performance when associated with the Binomial model. With Black's model, it demonstrated intermediate performance in accuracy and poor performance in precision.

GABE and PORTUGAL (2004) concluded that the volatility of the Garch model is the best predictor of future volatility throughout the option's maturity, as it captures volatility "clusters". POON and GRANGER (2005) found that Garch volatility proved to be an adequate predictor of future volatility over the term of the option. However, PONTES and MAIA (2017) concluded that deterministic volatility overpriced premiums in the Black model.

Historical volatility, associated with the Binomial model, showed poor performance in the accuracy and precision criteria, but intermediate performance when applied to the LSM method. With Black's model, it demonstrated poor performance in accuracy and intermediate performance in precision.

PONTES and MAIA (2017) found that historical volatility underpriced the theoretical premiums of the Black model. TONIN and COELHO (2012) found a significant difference between the premiums calculated using historical volatility and the observed premiums, particularly for options very in or out of the money, as they do not capture the smile effect and the term structure of volatility. POON and GRANGER (2005) concluded that historical volatility can be a satisfactory predictor of future volatility for at-the-money or close-to-the-money options when data is widely available.

The lack of liquidity of contracts listed on B^3 amplifies the fluctuation in premiums, lengthening the difference between the best purchase offer and the lowest sales price ("bid-ask spread"), negatively affecting the performance of volatility models (NAVARRO et al, 2022).

CONCLUSIONS

The Black model, combined with implied volatility, presented the best performance according to the precision and accuracy criteria, the Binomial model obtained an intermediate performance and the LSM model presented a weaker performance.

The wide dissemination of Black's model impacts investors' decisions and ends up influencing real market premiums.

Better adherence to the Binomial model was expected, as it prices the possibility of early exercise of American options.

The LSM method did not prove to be adherent in pricing out-of-the-money options, underpricing the premiums, considering that the regressions are only carried out on the trajectories in which the option is exercised, harming the convergence of the method in the case under analysis.

As expected, historical volatility underpriced out-of-the-money option premiums by not capturing the smile effect nor the term structure of volatility.

Implied volatility presented the best indicators of accuracy and precision, deterministic volatility from the Garch model achieved intermediate performance and historical volatility demonstrated weaker performance.

Better performance was expected from the deterministic volatility predictor of the Garch model, as it could capture volatility "clusters".

The main contribution of the study was to demonstrate that the theoretical premiums of the Black model, especially when combined with implied volatility, presented the best performance in the precision and accuracy criteria compared to numerical methods, corroborating its wide use.

Further studies will be able to deepen knowledge about the pricing of options in the Brazilian market in different situations, focusing on purchase and sale options, different agricultural "commodities", pricing models and volatility forecasters, involving sensitivity analyzes of parameters such as degree of "moneyness" and time until maturity.

REFERENCES

ARAÚJO, R.; BAIDYA, T. 2004. **Avaliação de opções reais através do método dos mínimos quadrados de Monte Carlo.** XXIV Encontro Nacional de Engenharia de Produção – Enegep, Florianópolis, SC, Brasil, 2004, Anais, p. 2248-2255.

BAIDYA, T.; CASTRO, A. 2001. **Convergência dos modelos de árvores binomiais para avaliação de opções.** Pesquisa Operacional, Vol 21, nº 1, 17-30, jun. 2001

BLACK, F.1976. **The pricing of commodity contracts**. Journal of Financial Economics 3,167-179.

BLACK, F.; SCHOLES, M. 1973. **The pricing of options and corporate liabilities.** Journal of Political Economy, 81, 637-654.

BOLLERSLEV, T. 1986. **Generalized autoregressive conditional heteroskedasticity.** Journal of Econometrics, 31: 307–327.

CHATEAU J. 2014. **Valuing european put options under skewness and increasing excess Kurtosis.** Journal of Mathematical Finance, 2014,4,160-177.

COELHO, A.; PINHEIRO, S.; FERREIRA, F. 2009. **A fórmula de Black precifica corretamente as opções de compra sobre futuros agropecuários no brasil? Uma aplicação para o caso do café arábica.** Pesquisa e Debate, SP, vol. 20, nº 2(36) p. 299-315, 2009.

COX, J.; ROSS, S.; RUBINSTEIN, M. 1979. **Option pricing: a simplified approach.** Journal of Financial Economics, no 7, 1979, p. 229-263.

ENGLE, R. F. 1982. **Autoregressive Conditional Heteroskedasticity with estimates of the variance of United Kingdom ination.** Econometrica, 50: 987–1007.

GABE, J.; PORTUGAL, M. 2004. **Volatilidade Implícita Versus Volatilidade Estatística: Um exercício utilizando opções e ações da Telemar S.A.** Revista Brasileira de Finanças, Rio de Janeiro, v. 2, n. 1, p. 47-73, 2004.

HULL, J.; WHITE, A. (1987). **The Pricing of Options on Assets with Stochastic Volatilities.** The Journal of Finance XLII.

JANKOVÀ, Z. 2018. **Drawbacks and limitations of Black-Scholes model for option princing.** Journal of Financial Studies and Research. Vol. 2018. Article ID 179814.

LONGSTAFF, F. A.; SCHWARTZ, E. S. 2001. **Valuing american options by simulation: a simple least-squares approach.** Review of Financial Studies, 14(1): 113-147.

MAIA, V.; FIGUEIREDO, A.; KLOTZLE, M. 2014. **Smooting the Volatility Smile using the Corrado-Su Model.** In; V Congresso Nacional de Administração e Ciências Contábeis – adCont 2014. Anais. Disponível em <http://www.academia.edu>. Acesso em 02 fev. 2022.

NAVARRO, T.; NUNES, V.; ALFANO, V.; HU, O. 2022. **Análise dos modelos de Black e Scholes para precificação de opções no mercado financeiro brasileiro.** Trabalho de Conclusão de Curso de Engenharia de Produção. Universidade Presbiteriana Mackenzie. https://dspace.mackenzie.br/handle/10899/29216. Acesso em 07 jul. 2022.

PINHO, F.; CAMARGO, M.; FIGUEIREDO, J. 2017. **Uma revisão de literatura sobre modelos de volatilidade em estudos brasileiros.** FACES Journal, vol. 16 nº 1, 10-28, 2017.

PONTES, T.; MAIA, S. 2017. **Precificação de Opções sobre contratos futuros de boi gordo na BM&FBovespa.** Economia Aplicada, v.21, n. 4, 2017, 737-760.

POON, S.; GRANGER, C. 2005. **Practical issues in forecasting volatilities.** Financial Analysts Journal, vol. 61, nº 1, 2005, p. 45-56.

SAITO, R.; ROCHMAN, R. 2008. **Avaliação de métodos numéricos na precificação de derivativos: revisão e aplicação à opção de compra da Telebrás PN.** Read Ed. 61 V. 14 nº 3, Set/ 2008.

SALIBY, E.; GOUVÊA, S.; MARINS, J. 2007. **Amostragem descritiva no apreçamento de opções europeias através de simulação de Monte Carlo: o efeito da dimensionalidade e da probabilidade de exercício no ganho de precisão.** Trabalhos para Discussão nº 134. Banco central do Brasil. Disponível em: <http://www.bc.gov.br>. Acesso em: 03 nov. 2021.

SILVA, M.; MAIA, S. 2014. **Previsões de volatilidade diária um passo à frente: Um estudo dos contratos futuros de soja com liquidação financeira da BM&FBOVESPA.** Disponível em https://www.anpec.org.br. Acesso em 03 nov. 2021.

SOUZA, W. MARTINES-Filho, J.; MARQUES, P. 2013. **Uso da estrutura a termo das volatilidades implícitas das opções de soja do CME Group para previsões em Mato Grosso.** Revista de Economia e Sociologia Rural, vol. 51, n° 2, 255-274, Piracicaba, SP, Brasil.

TONIN, J.; COELHO, A. 2012. **Testando modelos de precificação de opções: análise das opções de compra sobre contratos futuros de café arábica na BM&FBovespa.** Revista de Economia e Administração, julho de 2012, DOI: 10.11132/rea.2012.606.

UNITED STATES DEPARTMENT OF AGRICULTURE (USDA). 2024. **World Agricultural Supply and Demand Estimates (Wasde)**. Mar. 2024. Disponível em https://www.usda.gov/oce/commodity/wasde. Acesso em 09 mar. 2024.