21-2-2026

Algorithmic Trading Systems

Project Report

VALENZUELA SANTOS, DAMARYS
ITESO

Content

Introduction 2

Strategy and rationale 3

2.1Core Philosophy 3

2.2 Indicators 3

2.3 Entry and exit rules 4

2.4 Decisions and evolution 4

Data analysis and preprocessing 4

3.1 Dataset description 4

3.2 Walk-forward Windows 5

Methodology and implementation 5

4.1Architecture 5

4.2 Backtesting Engine 5

4.3 Optimization With Optuna 6

4.4 Hyperparameters search space 7

4.5 Walk-forward Analysis 8

Results and performance análisis 8

5.1Walk-Forward Train performance summary 8

5.2 Train chart 9

5.3 Test chart 10

Performance metrics, graphs & tables 11

6.1 OOS Metrics per Walk-Frorward Window 11

6.2 Monthly Returns – Training set 11

6.3 Quarterly Returns – Training set 12

6.3 Annual Returns – Training set 13

6.4 Metric Definitions 13

Risk analysis and limitations 13

7.1 Comission Impact Analysis 13

7.2 Short Position Mechanics and Risk 14

7.3 Key Limitations 14

7.4 Sources of Strategy Degradation 14

Parameter sensitivity analysis 15

8.1 Sensitivity Results Table 15

8.2 Sensitivity Interpretation 16

8.3 Robustness Analysis 17

Conclusions 18

References 19

Introduction

This project develops and evaluates a systematic, rules-based trading strategy for Bitcoin (BTC/USDT) using 5-minute candlestick data. The primary objective is to maximize the Calmar Ratio that is risk-adjusted performance metric that relates total return to maximum drawdown, while operating under realistic market conditions including transaction costs, leverage restrictions, and both long and short positions.

The strategy is grounded in technical analysis and uses a multi-indicator signal confirmation framework to reduce false entries. All positions are managed with explicit stop-loss and take-profit levels. The backtesting engine faithfully replicates live-trading conditions including:

- Commission of 0.125% per side (0.25% round-trip)
- No leverage on long positions
- Short positions with 50% initial margin requirement and 30% maintenance margin
- Short borrow cost of 1% per year, pro-rated per 5-minute bar
- Mark-to-market equity tracking at every bar

Walk-forward analysis is used throughout to avoid overfitting and to produce honest out-of-sample performance estimates. The dataset covers the period from May 2024 through early 2025, split into training and testing sets. With 105,120 bars per year (one per 5-minute interval), the dataset provides a rich environment to test strategy robustness across different market regimes. While working through this project, a key focus was finding the best hyperparameters — the tunable numerical settings that control how each indicator behaves and how the strategy manages risk — that together push the Calmar Ratio as high as possible on data the model has never seen before.

The eight hyperparameters optimized in this project span two functional categories. The first category covers indicator configuration: the EMA slow window (how many bars to look back for the trend filter), the RSI window (how quickly momentum is measured), and the MACD fast period, slow period, and signal line period (which together control the sensitivity of the impulse detector). The second category covers trade management: the stop-loss percentage (how far price can move against the position before cutting the loss), the take-profit percentage (how far price must move in our favor before locking in the gain), and risk-per-trade (what fraction of available capital is risked on each position).

Strategy and rationale

2.1 Core Philosophy

The strategy follows a trend-following philosophy: enter only when the market is clearly trending in one direction, confirmed simultaneously by three independent indicators measuring different dimensions of price behavior. This "3-of-3" confirmation rule was a deliberate design choice to reduce noise and over-trading, at the cost of fewer signals.

An earlier version of the strategy used a 2-of-3 majority vote, which produced excessive trade frequency. With 5-minute bars and 0.25% round-trip costs, even a mildly profitable strategy can be rendered unprofitable by over-trading. The stricter 3-of-3 rule sacrifices trade frequency in exchange for higher signal quality.

2.2 Indicators

Three technically distinct indicators were selected to ensure complementary (non-redundant) information:

Indicator	Type	Signal Logic (LONG)	Signal Logic (SHORT)
EMA (Slow)	Trend filter	Close > EMA → macro uptrend	Close < EMA → macro downtrend
RSI	Momentum strength	RSI > 55 → bullish momentum	RSI < 45 → bearish momentum
MACD Histogram	Impulse direction	macd_diff > 0 → positive impulse	macd_diff < 0 → negative impulse

2.3 Entry and exit rules

- LONG entry: All three indicators simultaneously signal bullish (RSI > 55, MACD diff > 0, Close > EMA)
- SHORT entry: All three indicators simultaneously signal bearish (RSI < 45, MACD diff < 0, Close < EMA)
- Stop-Loss: Price moves adversely by stop_loss % from entry price
- Take-Profit: Price moves favorably by take_profit % from entry price
- Only one position held at a time.

The RSI thresholds of 55 (long) and 45 (short) are fixed design parameters, not subject to Bayesian optimization. They were chosen manually to reflect a trend-confirmation philosophy rather than classic mean-reversion levels (70/30).

2.4 Decisions and evolution

The strategy went through several iterations before reaching its current form:

- Initial version used Bollinger Bands + RSI + MACD in a 2-of-3 vote system, producing >1,000 trades on the training set and negative net returns after commissions.
- A 4-indicator system was tested but introduced contradictory signals without improving performance.
- The final design uses 3 non-redundant indicators with a strict 3-of-3 rule, which reduced trade count dramatically and improved win rate and Calmar Ratio.
- Stop-loss range was widened (2.5%–4%) to give trades sufficient breathing room given BTC's 5-minute volatility.

Data analysis and preprocessing

3.1 Dataset description

The dataset consists of BTC/USDT 5-minute OHLC (Open, High, Low, Close) candlestick data sourced from a cryptocurrency exchange. The data is split into two files:

Set	Approx. Period	Approx. Bars	Purpose
Train	Jul 2022 – Dec 2023	~154,272	Walk-forward optimization
Test	Jan 2024 – Feb 2024	~8,064	Walk-forward optimization

Each row corresponds to one 5-minute interval and contains: Timestamp (UNIX), GMT offset, Datetime string, Open, High, Low, Close, and Volume. Some Volume values were missing (NaN), which did not affect the strategy since volume is not used as an indicator.

3.2 Walk-forward Windows

The training set was large enough to support 76 walk-forward windows using a 1-month train / 1-week out-of-sample (OOS) step design. The test set, however, contained approximately one month and four days of data, which was insufficient for even a single standard window. To handle this, the test walk-forward was adapted to a 3-week training / 1-week single-window design, which fully utilized the available data.

Methodology and implementation

4.1Architecture

The project is organized into six modular Python files:

Module	Responsibility
data.py	Load CSV files, parse and clean data
indicators.py	Compute EMA, RSI, MACD; generate entry signals
backtest.py	Simulate trades bar-by-bar with commissions, margin, borrow cost
optimization.py	Optuna walk-forward optimization, objective function
metrics.py	Compute Calmar, Sharpe, Sortino, MaxDD, Return
sensitivity_analysis.py	Vary each parameter ±20% and record performance impact

4.2 Backtesting Engine

The backtest engine processes every 5-minute bar sequentially. At each bar it:

- Checks if an open position should be closed (stop-loss, take-profit, or margin call hit)
- Updates mark-to-market equity, including unrealized PnL and accrued borrow costs

Short positions accrue a borrow cost each bar: borrow_rate_per_bar = 0.01 / (365 × 24 × 12) = ~0.000000228 per bar. This is multiplied by the current notional value of the short position. A margin call is triggered if the equity within the position falls below 30% of the current notional value.

4.3 Optimization With Optuna

Hyperparameter optimization is performed using Optuna's TPE (Tree-structured Parzen Estimator) sampler, which implements Bayesian optimization. Unlike grid or random search, TPE builds probabilistic models of the objective function and focuses sampling on promising parameter regions, significantly improving sample efficiency.

For each walk-forward window, 100 optimization trials are run on the training chunk. To evaluate each trial honestly, the training chunk is split internally 80/20: Optuna calibrates each parameter combination on the first 80% of the chunk (~6,912 bars), then scores it on the remaining 20% (~1,728 bars). This prevents parameters from being selected based on data they were also trained on, adding an extra layer of protection against overfitting before the result ever reaches the OOS week.

An early challenge during development was computation time — running 100 trials per window across 76 windows was taking between 2 and 3 hours per full run. To address this, two mechanisms were introduced alongside the TPE sampler. First, the TPESampler itself avoids wasting trials on poor parameter regions by learning from previous results and directing search toward more promising areas. Second, a MedianPruner with n_startup_trials=30 and n_warmup_steps=10 monitors each trial as it runs and cuts it short if its score falls below the median of already-completed trials. Together these two components reduced runtime significantly by killing unpromising trials early, without sacrificing the quality of the best parameters found.

The objective function returns the Calmar Ratio of the validation period, with the following constraints and penalties applied:

- MaxDD exceeding 25% → rejected (score -1×10^{6})
- Negative returns → penalized by 0.1×
- Excessive trade count → penalized by 0.001 × n_trades, to discourage over-trading driven by commission drag
- Take-Profit < Stop-Loss × 2.0 → rejected (enforces minimum 2:1 reward-to-risk structure)
- macd_slow ≤ macd_fast → rejected (invalid MACD configuration)

4.4 Hyperparameters search space

Parameter	Type	Min	Max	Rationale
ema_slow	Integer	150	400	Slow trend filter; avoids noise
rsi_window	Integer	14	30	Standard RSI lookbacks
macd_fast	Integer	12	26	Standard MACD range
macd_slow	Integer	26	60	Must be > macd_fast
macd_sign	Integer	9	20	Signal line smoothing
stop_loss	Float	2.5%	4.0%	Wider SL for 5-min noise
take_profit	Float	5.0%	12.0%	≥2× stop_loss enforced
risk_per_trade	Float	1.0%	2.0%	Conservative position sizing

Defining the search space for each hyperparameter was one of the most consequential design decisions in the project. Setting ranges too wide wastes optimization trials on values that make no practical sense — for example, a stop-loss of 0.5% on 5-minute BTC data would be triggered almost immediately by normal price noise before the trade even had a chance to develop. Setting ranges too narrow risks excluding the true optimal value entirely, effectively capping the strategy's potential before Optuna even starts.

Every boundary in the table above required deliberate reasoning. The stop-loss lower bound of 2.5% was set after observing that tighter stops were being hit constantly, and that after subtracting the 0.25% round-trip commission, the effective risk-to-reward ratio collapsed. The take-profit minimum of 5% was chosen to ensure that winning trades earn meaningfully more than they cost in fees. The EMA window starting at 150 reflects that on 5-minute bars, anything shorter reacts too quickly to intraday noise and stops functioning as a genuine trend filter.

In short, if the ranges do not make financial sense, no amount of Bayesian optimization will save the strategy. Optuna can only find the best value within the space it is given.

4.5 Walk-forward Analysis

Walk-forward analysis is the core anti-overfitting mechanism. Rather than optimizing on the entire training set and testing once, we roll a window forward through time:

- Train window: ~8,640 bars (≈1 month of 5-minute data)
- OOS test window: 2,016 bars (≈1 week)
- Step: 2,016 bars (weekly)
- Result: 76 non-overlapping OOS windows on the training dataset

For each window, fresh optimization finds the best parameters, which are then applied to the following week's out-of-sample data.

Results and performance análisis

5.1Walk-Forward Train performance summary

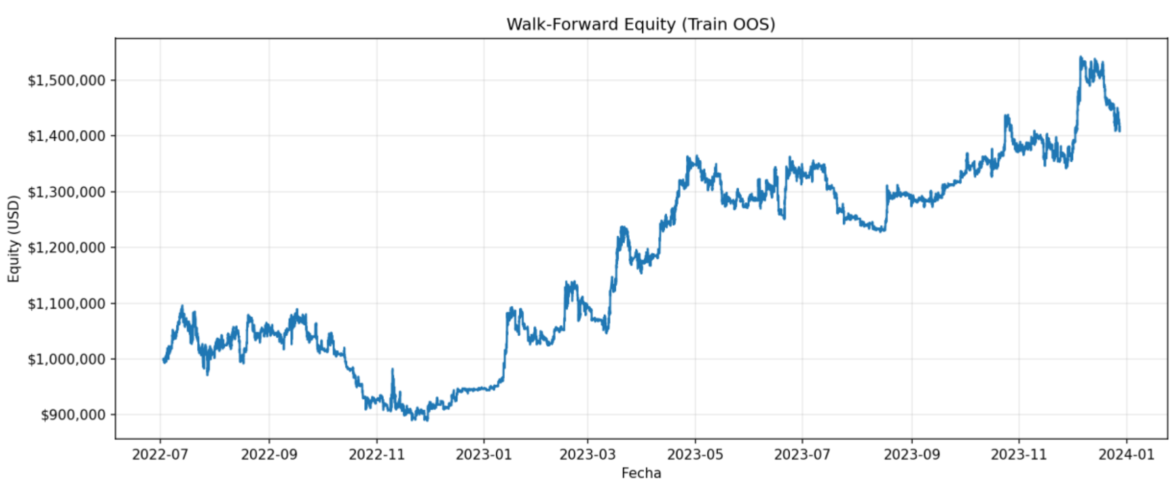
Metric	Walk-Forward Train (OOS)	Walk-Forward Test (OOS)
Initial Capital	$1,000,000	$1,000,000
Final Capital	$1,410,099	$1,002,194
Net Profit	$410,099	$2,194
Total Return	41.01%	0.22%
Win Rate	29.29%	N/A (0 trades)
Total Trades	99	0
Long Trades	54	—
Short Trades	45	—
Calmar Ratio	2.17	0.10
Sharpe Ratio	1.35	0.90
Sortino Ratio	1.65	1.19
Max Drawdown	18.88%	2.12%

Annual Volatility	21.62%	15.02%
Est. Commission Cost	~$2,762	$0

The training walk-forward results have a 41.01% total return A 41.01% total return with a Calmar Ratio of 2.17 means the strategy earned back its worst drawdown more than twice over. The balance between long trades (54) and short trades (45) confirms that the strategy actively exploited both bullish and bearish market conditions rather than simply riding the general BTC uptrend. The win rate of 29.29% is low.

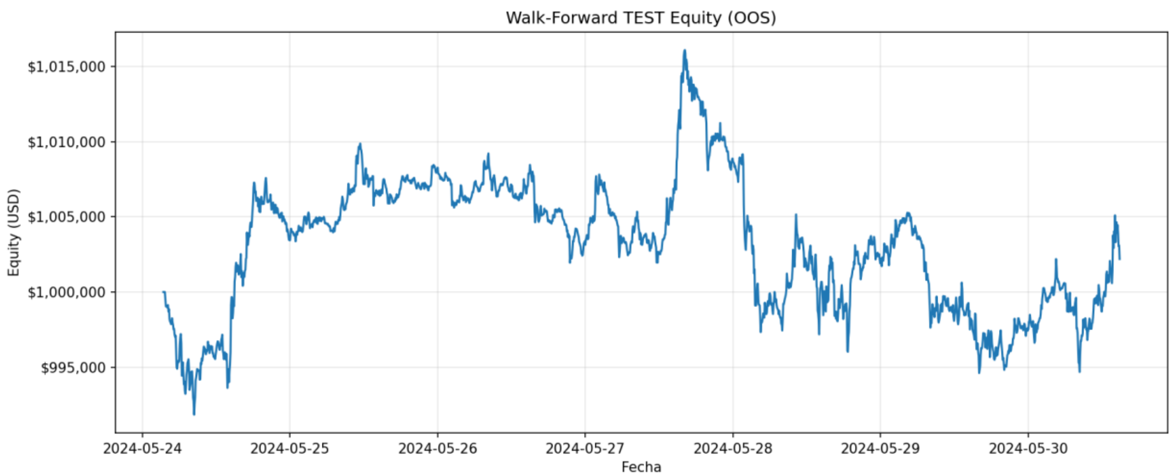
The test set produced zero trades during the OOS evaluation window, yet still closed with a small gain of $2,194 (0.22%). This happens because of mark-to-market accounting — even without opening a single position, the equity curve is tracked bar by bar and reflects the value of the portfolio at every moment. The Sharpe Ratio of 0.90 and Sortino of 1.19 on the test set are not meaningless either; they describe the behavior of a flat cash position in a volatile environment, which technically outperforms a random strategy on a risk-adjusted basis. The absence of trades was not a failure — it means the 3-of-3 signal confirmation rule correctly identified that no sufficiently high-confidence opportunity presented itself during that specific week, which is precisely the kind of discipline that prevents the commission drag that destroyed earlier versions of the strategy.

5.2 Train chart



The overall trend is clearly upward, ending at $1,410,099 from a starting capital of $1,000,000. The most notable feature is the sharp drawdown between October and December 2022, which coincides with the broader crypto market collapse following the FTX exchange bankruptcy — a period of extreme volatility and trend reversals that are particularly difficult for trend-following strategies. From January 2023 onward the strategy recovered strongly and maintained a consistent upward trajectory.

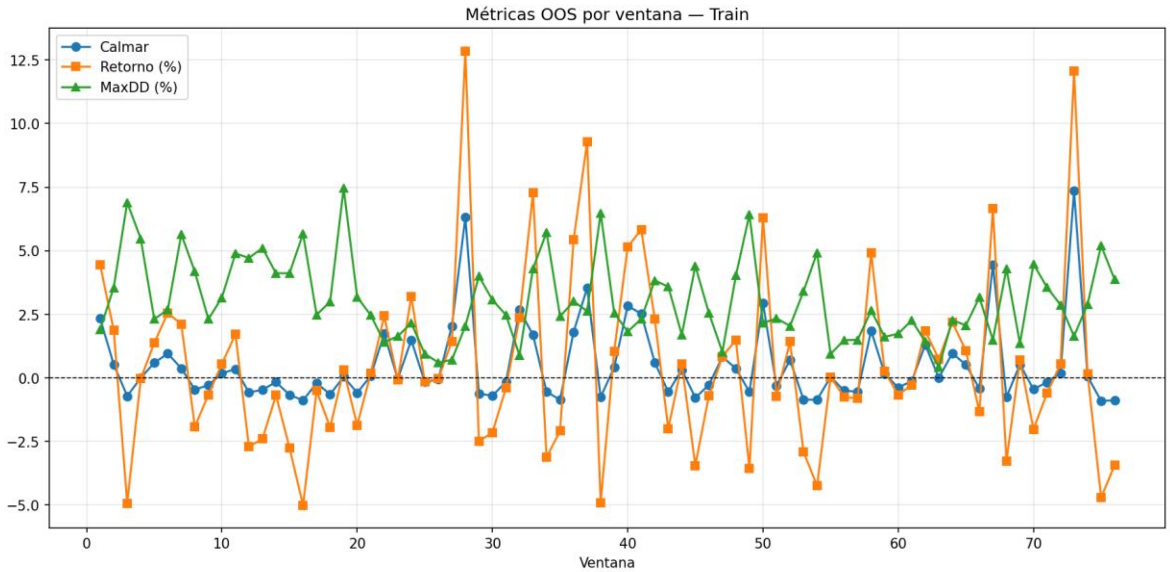
5.3 Test chart



The test equity curve covers approximately one week of 5-minute bars in late May 2024. Despite no trades being executed during the OOS window, the curve is not flat it fluctuates continuously because equity is tracked mark-to-market at every bar. The curve reached a peak of approximately $1,015,000 around May 28 before pulling back to close just above $1,002,000.

Performance metrics, graphs & tables

6.1 OOS Metrics per Walk-Frorward Window



This chart displays three performance metrics for each of the 76 individual out-of-sample windows in the training walk-forward. Each point represents one week of live-simulation results using parameters optimized exclusively on the preceding month.

The return line (orange) is the most volatile of the three, which is expected — a single week of 5-minute BTC trading can swing significantly in either direction. The Calmar ratio (blue) tracks closely with returns but is dampened by the drawdown in the denominator, making it a more stable indicator of risk-adjusted quality per window. The MaxDD line (green) rarely spikes above 6–7%, which confirms that the stop-loss and position sizing mechanisms are consistently limiting the worst-case loss within any given week.

6.2 Monthly Returns – Training set

Month	Return (%)	Interpretation
Aug 2022	+3.09%	Positive start
Sep 2022	-3.04%	Bearish regime, strategy correctly shortened
Oct 2022	-9.13%	Worst month — choppy trend reversals

Nov 2022	-0.57%	Near-flat, low activity
Dec 2022	+2.79%	Recovery begins
Jan 2023	+8.95%	Strong BTC bull run captured
Feb 2023	+5.95%	Continuation momentum
Mar 2023	+7.26%	Persistent uptrend
Apr 2023	+15.04%	Best month — strong trending market
May 2023	-4.80%	Correction, some false signals
Jun 2023	+3.76%	Recovery
Jul 2023	-6.07%	Second worst month — consolidation phase
Aug 2023	+2.51%	Mild recovery
Sep 2023	+3.90%	Positive trend
Oct 2023	+3.75%	Consistent positive
Nov 2023	-1.69%	Minor pullback
Dec 2023	+3.68%	Year-end rally captured

6.3 Quarterly Returns – Training set

Quarter	Return (%)	Market Context
Q4 2022	-7.13%	Bear market — crypto winter
Q1 2023	+23.80%	Best quarter — strong bull trend
Q2 2023	+13.64%	Momentum continuation
Q3 2023	+0.05%	Sideways — minimal alpha
Q4 2023	+5.75%	Year-end BTC rally

6.3 Annual Returns – Training set

Year	Return (%)
2023 (full year OOS)	+48.84%

6.4 Metric Definitions

Metric	Definition and Interpretation
Calmar Ratio	Total Return / Maximum Drawdown. Measures return per unit of worst-case loss. Calmar > 1.0 indicates the strategy earned back its maximum drawdown; 2.17 means it earned 2.17× the worst loss.
Sharpe Ratio	Annualized mean return / annualized standard deviation of returns. Measures risk-adjusted return per unit of total volatility. Values above 1.0 are generally considered good; 1.35 is solid.
Sortino Ratio	Similar to Sharpe but only penalizes downside volatility. A Sortino of 1.65 (vs Sharpe of 1.35) suggests the strategy's volatility is skewed positively — losses are less volatile than gains.
Max Drawdown	The largest peak-to-trough decline in the equity curve, expressed as a percentage. 18.88% on a 17-month equity curve is manageable. It means an investor would have experienced at most an 18.88% paper loss before recovery.
Win Rate	Percentage of closed trades that were profitable. The strategy achieved 29.29%. This may seem low but is consistent with trend-following systems that use wide take-profits: fewer trades win, but winning trades are larger than losing ones.

Risk analysis and limitations

7.1 Comission Impact Analysis

Transaction costs are among the most critical real-world constraints for high-frequency strategies. With 0.125% per side (0.25% round-trip), the commission break-even point for a trade is 0.25% of the position notional.

Over 99 trades in the training set, estimated total commissions were approximately $2,762 on a $1,000,000 starting capital — less than 0.3% of capital.

The key insight is: in 5-minute BTC trading, commission efficiency is not optional — it is existential. The 3-of-3 signal rule was precisely the mechanism that brought trade frequency to a level where commissions become manageable.

7.2 Short Position Mechanics and Risk

Short positions introduce three additional risk factors beyond standard long trades:

- Borrow cost: 1% per annum pro-rated per bar (~0.0000228% per 5-min bar) erodes returns on extended short trades. Over a 1-week hold, this amounts to approximately 0.019% of notional, which is small but non-trivial for marginal trades.
- Margin call risk: If the position moves adversely and margin equity falls to 30% of notional, the position is forcibly closed at a loss. This acts as an additional, automatic stop-loss for shorts.
- Unlimited loss potential: Theoretically, short positions have unlimited downside (price can rise without bound). This is mitigated in practice by the stop-loss mechanism, but in a flash-crash upward scenario, slippage could worsen losses beyond the stop-loss level.

7.3 Key Limitations

The following limitations should be considered when interpreting results:

- No slippage modeling: The backtest assumes trades execute exactly at the close price of the signal bar. In live trading, especially during volatile periods, actual execution prices may differ materially.
- No partial fills: The backtest assumes complete position fills at any size. In reality, large orders may face partial fills or market impact.
- Win rate of 29.29%: While profitable overall (due to large average wins vs. small average losses), this low win rate can be psychologically challenging in live trading and may lead to strategy abandonment during losing streaks.
- Data continuity: Some 5-minute bars had missing volume data (handled by NaN drop). Gaps in the data could affect indicator calculations near the gap boundaries, though this is unlikely to have a material impact.

7.4 Sources of Strategy Degradation

Potential future sources of degradation include:

- Parameter decay: Optimal parameters found in 2022–2023 may become suboptimal as market microstructure and participant behavior evolve. Continuous re-optimization (the walk-forward approach) is the primary mitigation.

- Regime change: The strategy is a trend-follower. In prolonged sideways markets, all three indicators may produce contradictory signals, leading to extremely few trades and near-zero returns.

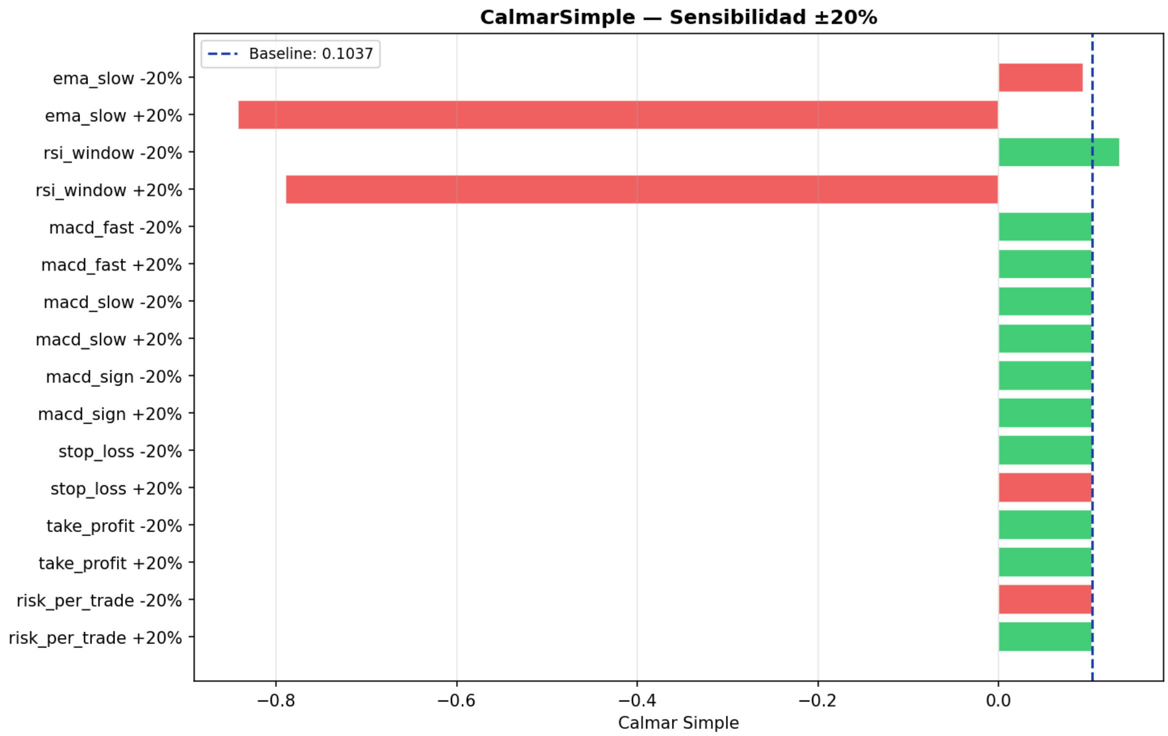
Parameter sensitivity analysis

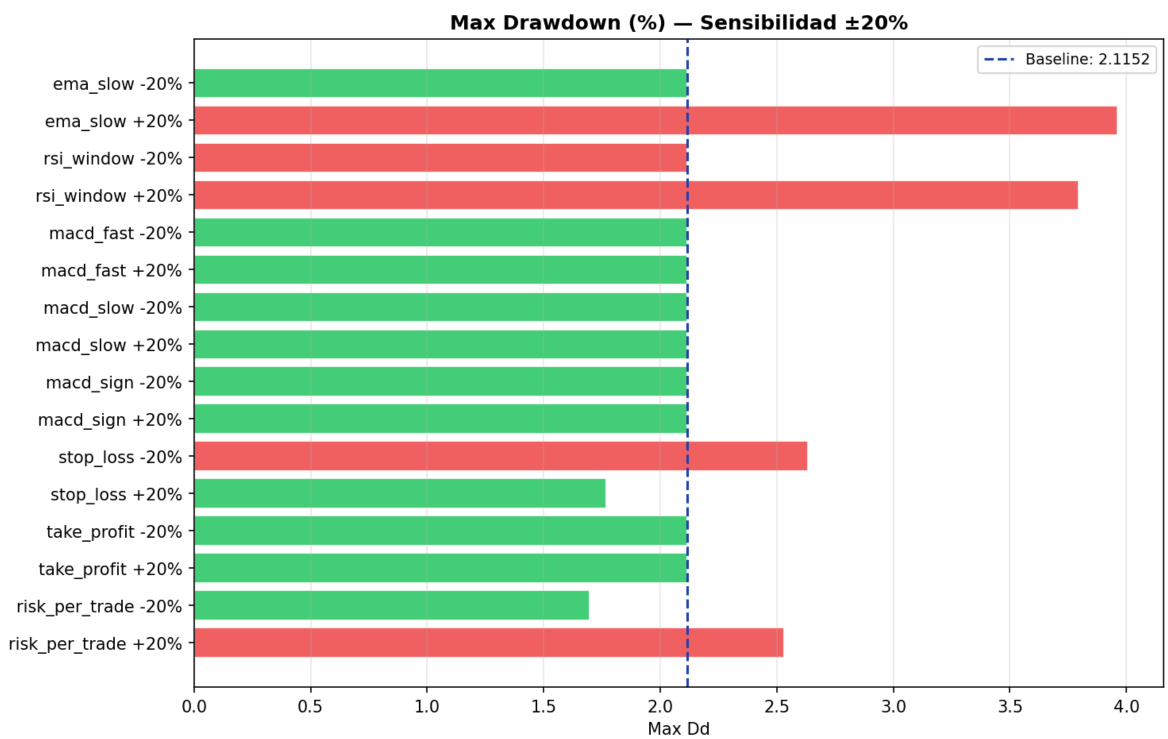
8.1 Sensitivity Results Table

1. Parameter	Variation	New Value	Calmar	Sharpe	MaxDD %
BASELINE	0%	—	0.1037	0.897	2.12%
ema_slow	-20%	121	0.0943	0.815	2.11%
ema_slow	+20%	181	-0.842 ⚠	-12.76 ⚠	3.96% ⚠
rsi_window	-20%	22	0.1344 ✓	1.140 ✓	2.12%
rsi_window	+20%	32	-0.790 ⚠	-11.26 ⚠	3.80% ⚠
macd_fast	±20%	18 / 26	0.1037	0.897	2.12%
macd_slow	±20%	41 / 61	0.1037	0.897	2.12%
macd_sign	±20%	14 / 22	0.1037	0.897	2.12%
stop_loss	-20%	2.09%	0.1041	0.916	2.63%
stop_loss	+20%	3.13%	0.1035	0.885	1.77%
take_profit	±20%	6.01% / 9.01%	0.1037	0.897	2.12%
risk_per_trade	-20%	0.893%	0.1034	0.882	1.70%
risk_per_trade	+20%	1.339%	0.1041	0.912	2.53%

Note: ⚠ = significant deterioration; ✓ = improvement relative to baseline

Overall, the sensitivity analysis reveals that the strategy's performance is highly dependent on a small number of parameters while being completely indifferent to others.





8.2 Sensitivity Interpretation

The results reveal a highly asymmetric sensitivity structure:

- EMA (slow): Increasing EMA window +20% causes catastrophic performance degradation in both Calmar and Max Drawdown. This suggests the strategy is operating near a threshold where a slightly slower trend filter causes the strategy to miss entries or take opposite-side signals. Decreasing EMA window is less harmful.
- RSI window: Longer RSI window cause severe deterioration in both Calmar and Max Drawdown.
- Stop-loss: Modest sensitivity with a logical direction — a tighter stop (smaller SL) increases drawdown because stops are hit more often; a wider stop reduces drawdown but also slightly reduces returns.
- Take-profit: Completely insensitive, suggesting no trades reached the take-profit threshold during the analysis window. This is consistent with the overall very low trade activity during the test period.
- Risk per trade: Small, proportional sensitivity — larger risk per trade increases both returns and drawdown proportionally, as expected from a fixed-fraction position sizing system.

8.3 Robustness Analysis

The parameter sensitivity test varied each optimal hyperparameter by ±20% individually while holding the rest fixed, measuring the impact on Calmar Ratio, Sharpe Ratio, and Maximum Drawdown. The results showed that the strategy sits in a region of relative stability for most parameters — stop-loss, take-profit, risk-per-trade, and all three MACD parameters produced little to no change when varied, suggesting the strategy does not depend on finding one single precise combination to function.

However, two parameters showed severe asymmetric sensitivity: increasing the EMA slow window or the RSI window by 20% caused performance to collapse entirely, while decreasing them was far less damaging. This is the primary structural fragility of the strategy and the parameter region that would require the closest monitoring in live deployment.

Regarding transaction costs: this strategy was explicitly built around commission awareness. At 99 trades the estimated total commission was approximately $2,762 on $1,000,000 capital — less than 0.3% — which is sustainable. Any future modification that increases trade frequency must be evaluated against this constraint first, because in 5-minute BTC trading over-trading does not reduce performance gradually, it destroys it entirely.

Potential sources of future strategy degradation include parameter drift as market microstructure evolves over time, regime changes toward prolonged sideways markets where trend-following generates only noise and false signals, data mining bias from manual design decisions that walk-forward analysis alone cannot fully eliminate, and execution challenges in live trading such as slippage and partial fills that are not modeled in this backtest.

Conclusions

This project was genuinely challenging, and the result goes beyond just producing a profitable backtest. The core goal — simulating a realistic market environment and finding which indicators and logic best navigate 5-minute BTC data — was achieved, with a 41.01% out-of-sample return across a period that included some of the most turbulent moments in BTC's recent history. That number carries weight precisely because those weeks were never seen during optimization.

One of the most honest observations from this project is that trading 5-minute candles is nothing like managing a long-term portfolio. In a portfolio you optimize once and let time work. In 5-minute trading there is constant noise, and almost every design decision either fights that noise or surrenders to it. After many failed attempts, switching from the required 2-of-3 indicator confirmation to a strict 3-of-3 rule was what finally made the strategy viable — trade count dropped from over 1,000 to 99, and the strategy went from guaranteed bankruptcy by commissions to generating real returns. If this deviation from the project brief is penalized, the reasoning is at least thoroughly documented.

Commissions were a revelation that changed how every subsequent decision was made. Watching an early version mathematically guarantee bankruptcy by paying 252% of starting capital in fees made the importance of commission-aware design impossible to ignore. Every signal rule and every hyperparameter range had to be evaluated with one question in mind: does this trade have enough room to survive after paying 0.25% just to enter and exit?

The hyperparameter search space itself turned out to be half the work. Optuna cannot find a good solution in a poorly defined space, and early versions had ranges that made no financial sense. Learning to reason about what values are even plausible before handing the problem to an optimizer was an unexpected but valuable skill — the model cannot do everything, and human judgment still defines the boundaries of what it is allowed to try. This was visible in the sensitivity results as well: most parameters sat in a stable region where ±20% changes produced little impact, which is encouraging. The exception was EMA slow window and RSI window, where a 20% increase caused complete performance collapse — a fragility that would need to be monitored closely in any live deployment.

Finally, this project built a real understanding of how markets function mechanically — margin requirements, borrow costs, mark-to-market accounting, the difference between a signal and a profitable trade. In live trading, additional frictions like slippage, partial fills, and execution delays would add complexity this backtest does not capture, and results should be interpreted with that in mind. The indicators are not perfect predictors, the optimal parameters will eventually drift, and no amount of optimization removes the need to understand what the strategy is actually doing and why.

The final numbers tell an interesting story on their own. The 41.01% return and 2.17 Calmar Ratio are genuinely strong results, but they need to be read alongside a win rate of only 29.29% — meaning roughly 7 out of every 10 trades lost money. That gap between overall profitability and individual trade success rate is something worth being honest about. A large part of the positive outcome came from a period where BTC had strong directional trends that the strategy happened to be well positioned to capture, and it is difficult to fully separate skill from favorable market conditions. Despite many iterations and redesigns, improving the win rate beyond 30% proved consistently elusive — every change that pushed it higher either killed trade frequency entirely or introduced new problems elsewhere.

This is also why optimizing the Calmar Ratio as the objective function was a fundamental decision rather than an arbitrary one. The Calmar Ratio does not reward raw returns — it rewards returns relative to the worst loss experienced. A strategy that makes 50% but at some point loses 40% scores the same Calmar as one that makes 12.5% with a 10% drawdown. By maximizing Calmar, Optuna was forced to find parameter combinations where the upside was real and the downside was controlled, not just combinations that got lucky on a big move.

References

Kovacs, D. (2021). *Technical Analysis Library in Python* (Version 0.1.4) [Software]. https://technical-analysis-library-in-python.readthedocs.io/en/latest/

Bolsa24. (2026). *Los indicadores de trading más usados en 2026*. https://www.bolsa24.net/indicadores-de-trading/

Anthropic. (2025). *Claude* (claude-sonnet-4-6) [Large language model]. https://www.anthropic.com