

Sample report built from computer-generated synthetic data: it is nobody's account or strategy. This is what a full report looks like.

RIGOR · BACKTEST AUDIT



VERDICT

Class C: there is a material weakness; we would not rely on this backtest until it is resolved.

DIMENSIONS

Statistical significance

• Pass

Number of settings tried

• Weak

Costs

• Weak

Out of sample

• Weak

Data quality and trading pattern

• Pass

Benchmark

• Not applicable

EXECUTIVE SUMMARY

+173.7%

Total return

-8.2%

Maximum drawdown
(closed trades only)

-12.3%

Drawdown p95, 1 year
(closed trades only)

1.79

Annualised Sharpe

WHAT TO DO NOW

- 1 Ask the seller why your live account falls outside what the backtest led you to expect.
- 2 Compare your broker's spread and commission with the costs the result can bear: a cost a little above the reference erases the margin.
- 3 Ask how many configurations were tried before this one was chosen, and on which period it was chosen.

Every figure carries its tag: "Measured" when computed from your files; "Declared" when stated by you or the seller, not verified; "Not measured" when a piece was missing to compute it.

Sample report built from computer-generated synthetic data: it is nobody's account or strategy. This is what a full report looks like.

• RIGOR • BACKTEST AUDIT

Verdict C

Identifier sample

Generated 2026-09-24 00:00 UTC

engine version 0.1.0

simulation seed 12345

VERDICT

Class C: there is a material weakness; we would not rely on this backtest until it is resolved.

As a single test, the result is too consistent to be explained by chance alone (Sharpe ratio distinguishable from zero). With 120 trials counted in the files, the Sharpe adjusted for those trials misses the bar: if more configurations were tried, the result may come from picking the best one. The trades stay positive at the reference cost, but not at 3 times that cost. In the period held back for checking (out of sample) the result stays positive but worsens. No red flags in the data. Benchmark declared not applicable.

• Live account: Not consistent. Trading result: -305.20 on 1,500.00 deposited.

• Measured

[See the comparison with the backtest](#)

Every figure carries its tag: "Measured" when computed from your files; "Declared" when stated by you or the seller, not verified; "Not measured" when a piece was missing to compute it.

Money reconciliation

We compare starting capital, known flows and net closed-trade P&L with the closing balance. A match does not authenticate the history.

Matches within tolerance

Closed trades agree within tolerance; this does not authenticate the history.

Gross – itemised costs = net closed-trade P&L: $26,343.75 - 8,974.00 = 17,369.75$

Starting capital + known flows + net P&L = expected closing balance: $10,000.00 + 0.00 + 17,369.75 = 27,369.75$

USD

Starting capital	10,000.00	• Measured
Flows after the start	0.00	• Measured
Gross closed-trade P&L	26,343.75	• Measured
Itemised costs	8,974.00	• Measured
Net closed-trade P&L	17,369.75	• Measured
Open-position value	—	• Not measured
Expected closing balance	27,369.75	• Measured
Observed closing balance	27,369.75	• Measured
Difference (observed – expected)	0.00	• Measured
Tolerance	14.11	• Measured

Coverage and limits

- **Curve:** rebuilt from the platform deal rows
- **Closed trades:** 1282
- **Trades outside the curve period:** 0
- **Deposits and withdrawals:** listed by the platform
- **Open positions:** not valued separately
- **Currency:** USD

Executive summary

+173.7%

Total return

how much the account changed over the whole history

-8.2%

Maximum

drawdown (closed trades only)

the worst fall from a peak

-12.3%

Drawdown p95, 1 year (closed trades only)

a fall exceeded in 1 of every 20 simulated years

1.79

Annualised Sharpe

return against its ups and downs; higher is steadier

1.52

Profit factor

what was won for every 1 lost

1,282 · 55%

Trades · win rate

1.59

Extra cost that takes it to zero (bps per side)
how much more trading can cost before it reaches zero

+15,375.41

Without the best 5 trades
what is left of the net result without those 5; with all of them: +17,369.75

+140.0%

Without the best 5 periods

total return without those 5; with all of them: +173.7%

After subtracting what cash in dollars paid over the same dates (3-month US Treasury bills, 2.57% a year on average), the Sharpe is 1.58. The Sharpe above subtracts no rate. If the account is not in dollars, the fair rate to subtract is its own currency's. Source: [FRED](#). • Measured

What this means for you

Statistical significance

• Pass

With this much data, a result like this is hard to get by pure luck. That says nothing about what happens next: only that the history is not noise.

Number of settings tried

• Weak

Part of the result may come from picking the best of many configurations. Ask how many were tried and request the optimisation file.

Costs

• Weak

At normal costs the result stays positive, but at high costs it disappears. A wider spread or higher commission than assumed would erase it.

Out of sample

• Weak

Out of sample the result stays positive, but much worse than in sample. This is common in somewhat overfitted strategies.

Data quality and trading pattern

• Pass

We found no jumps, gaps or hidden-risk patterns in the files. That does not rule out errors the files do not show.

Benchmark

• Not applicable

No applicable reference was declared. A comparison with a passive alternative is outside this report.

What to do now

If you bought or are about to buy this robot or signal, this is what is worth clearing up first, from what the audit found.

- 1 Ask the seller why your live account falls outside what the backtest led you to expect. **Go to the section →**
- 2 Compare your broker's spread and commission with the costs the result can bear: a cost a little above the reference erases the margin. **Go to the section →**
- 3 Ask how many configurations were tried before this one was chosen, and on which period it was chosen. **Go to the section →**
- 4 Take this report's questions to the seller. **Go to the section →**
- 5 Keep this report and its identifier; if the robot changes, ask for a new audit.

What each class requires

The class does not measure how much was made, but how many questions your files answer. A better class does not mean the strategy will work.

A

Statistics and number of trials pass; costs, out-of-sample and benchmark pass or do not apply; the data has no serious or warning flags.

B

Statistics pass, the number of trials passes or was not declared and nothing fails, but costs, out-of-sample, benchmark, data quality or the number of trials still need measuring or strengthening.

C

One dimension fails, or statistics or number of trials are weak.

YOUR REPORT

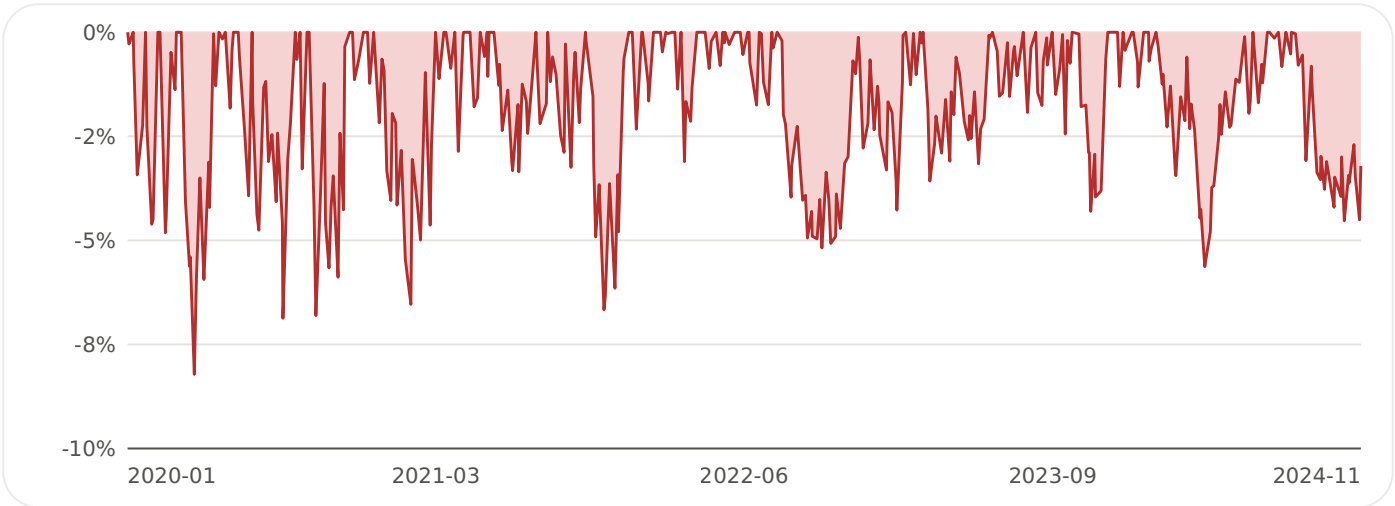
D

The data or the statistics fail, or two dimensions or more fail.

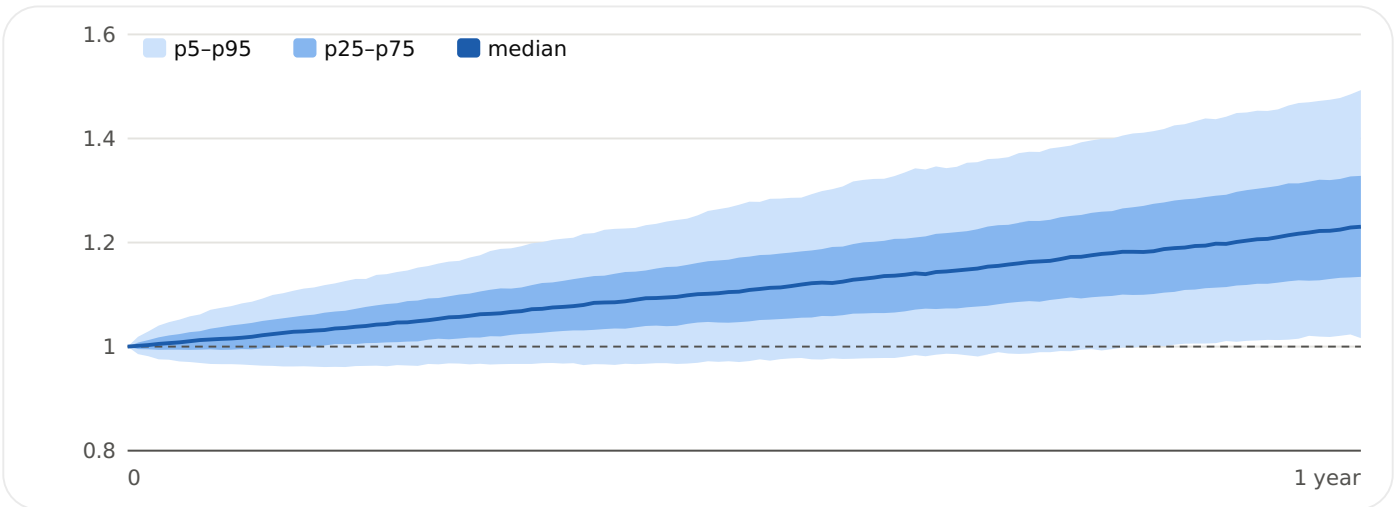
Charts



Measured Equity curve: 400 points, min 9,972, max 28k. balance rebuilt from closed trades; floating drawdown is not visible



Measured Drawdown (fall from the previous peak): max -8.2%. balance rebuilt from closed trades; floating drawdown is not visible



Measured Fan of resampled scenarios. Resampled from the supplied history; not a forecast and says nothing about future results.

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
2020	+0.9%	+3.8%	+3.8%	+3.6%	+3.0%	+4.2%	-3.3%	+4.1%	-0.5%	0.0%	+8.1%	+5.7%	+38.3%
2021	-2.7%	-0.7%	+4.4%	+8.7%	+6.3%	+0.1%	-0.3%	+1.3%	+0.3%	+2.8%	-5.5%	+8.7%	+24.9%
2022	-0.9%	+8.4%	+3.2%	+6.9%	+4.8%	+2.4%	+2.5%	-2.2%	-2.7%	+1.0%	+3.9%	-2.2%	+27.4%
2023	+0.9%	+3.8%	-0.6%	+1.1%	-0.6%	+1.7%	+0.3%	+1.0%	+1.5%	-1.8%	+5.0%	+4.1%	+17.4%
2024	+2.9%	-0.9%	+1.0%	-1.1%	+2.0%	+0.3%	+4.5%	-0.2%	-2.7%	-0.5%	+0.8%		+6.0%

Measured Monthly returns.

Red flags found

✓ No red flags in the audited files.

Internal file consistency

Heuristic checks of rows within this file; they do not compare two independent files or authenticate who created the history.

Method version: forensics-1 · File family: mt5_tester · Rows read: 2573 Measured

No calibrated signal appeared in these checks.

A signal does not prove forgery; no signal does not prove authenticity.

All checks and their status

CHECK	STATUS	CALIBRATION
FILE_TRACE	Data point	No applicable calibration
TOTALS_VS_ROWS	Data point	Calibrated for this family
SUMMARY_IDENTITIES	Data point	Calibrated for this family
BALANCE_CHAIN	No finding	Calibrated for this family
DEAL_SEQUENCE	No finding	Calibrated for this family
TESTER_NUMBERING	No finding	Calibrated for this family
TICKET_ORDER	Data point	No applicable calibration
DUPLICATE_TICKET	No finding	Calibrated for this family
TICKET_LINKS	Data point	No applicable calibration
CROSS_COPIES	Data point	No applicable calibration
SLTP_FILL	Data point	Calibrated for this family
PNL_SIGN	Data point	No applicable calibration
PRICE_IMPLIED_PNL	Data point	No applicable calibration
PRICE_PRECISION	No finding	Calibrated for this family
TIME_SANITY	No finding	Calibrated for this family

CHECK	STATUS	CALIBRATION
ROW_ORDER	No finding	Calibrated for this family
MARKET_HOURS	No finding	No applicable calibration
HIDDEN_CONTENT	No finding	Calibrated for this family
VOLUME_IN_OUT	No finding	No applicable calibration
STATEMENT_PERIOD	Data point	No applicable calibration
TV_INVARIANTS	Data point	No applicable calibration
NT_INVARIANTS	Data point	No applicable calibration
MONTHLY_DIGITS	Data point	No applicable calibration

Plan to reach a better class

What the audit's rules would need to see in each open dimension, most decisive first. A better class means the files answer more questions, not that the strategy will work.

01 Measure how many configurations were tried

• Weak

DSR 0.905 at 120 trials; it passes at 0.95 or more and fails below 0.5. With 16384 or more configurations tried it falls below 0.5.

- The trial count already comes from your files; the variants matrix (each configuration's results over time) would add the PBO.
- Fewer parameters and narrower ranges mean fewer trials.
- Validate the chosen configuration on a stretch not used while optimising.

If this dimension passed and the rest stayed the same, the class would be **B**.

02 Check the real costs

• Weak

The net reaches zero at 1.59 bps per side of extra cost. To pass this dimension it has to stay above zero at 3x the reference (3.00 bps per side).

- Compare that margin with your broker's real spread and slippage: on EURUSD at 1.10, 1 bp per side is about 1.1 pips.
- Declare the real cost per side when uploading: it is added to what the report already itemises.
- Fewer trades or a larger move per trade make costs weigh less.

03 Add an out-of-sample stretch

• Weak

Out-of-sample Sharpe 0.54; 0.5 or more is needed (met). In-sample Sharpe minus out-of-sample Sharpe: 1.35; up to 1.0 is accepted (not met).

- A sharp drop out of sample often appears when too many parameters were tuned: fewer parameters and a fresh validation on unseen data.

Backtest against the live account

If the live account's trades came from the same backtest, how unusual would its result be? We drew backtest trades at random, as many as the live account holds, 5,000 times, and placed the live account among those histories.

- **Not consistent** The live account does not behave like the backtest: its net result or its deepest fall is worse than in 99 % of the backtest's histories.

	LIVE, AT THE BACKTEST'S SIZE	EXPECTED RANGE (90 %)	BACKTEST
Trades	180		1,282
Period	2024-09-09 – 2025-06-20		2020-01-02 – 2024-11-29
Trades per month	19.3		21.8
Win rate	47%	49% ... 61%	55%
Net result	-1,526.00	-348.92 ... 5,177.01	17,369.75
Deepest fall	1,862.50	629.30 ... 2,186.22	1,528.76
Average win	102.62		103.36
Average loss	-105.69		-94.81

1%

Backtest histories with a net result as low or lower

• Measured

10%

Backtest histories with a fall as deep or deeper

• Measured

The live account trades 0.20 times the backtest's size: each live trade was scaled to the backtest's median size before comparing.

Part of the live account falls inside the backtest's period: those dates may have been used to fit the backtest, so the comparison is less demanding.

Backtest trades resampled with replacement, as many as the live statement holds; costs itemised per trade subtracted on both sides. Streaks are not preserved.

Same dates, trade by trade

From 2024-09-09 to 2024-11-29 both files cover the same days. Each live trade was looked up in the backtest: same side, same symbol and an entry less than 60 minutes apart.

Live trades found in the backtest	55 of 59 (93%)	Measured
Backtest trades the live account did not take	4 of 59 (7%)	Measured
Median price difference at entry	0.7 bp	Measured
Median price difference at exit	1.0 bp	Measured
Result difference on the paired trades	-82.60 (-1.50 per trade)	Measured

bp = basis points (0.01 % of the price); positive is worse for the account. The result difference is at the backtest's size; negative is what the live account made below the backtest on the same trades.

The account's real money

The percentage gain track-record sites show takes deposits and withdrawals out. Here it sits next to the money the account made or lost by trading, deposits made in a deep drawdown, and positions still open when the history was printed.

Review of the history you uploaded as the live account. Its flags are shown here and do not change the backtest's class.

Warning

Deposits in a deep drawdown

1 deposit arrived while the account was at least 20% below its peak

-30%

Percentage gain, as track-record sites show it.

Measured

-305.20

Trading result, in money, on 1,500.00 deposited.

Measured

4%

Open loss over the balance when the history was printed.

Declared

METRIC	VALUE	EVIDENCE	NOTE
Deposits	2	• Measured	
Money deposited	1,500.00	• Measured	
Withdrawals	1	• Measured	
Money withdrawn	300.00	• Measured	
Trading result, in money	-305.20	• Measured	Closed trades after commission and swap, in the account currency
Percentage gain	-30.03%	• Measured	Time-weighted: deposits and withdrawals are taken out, as track-record sites compute gain
Result on the money deposited	-20.35%	• Measured	Trading result / money deposited
Share of deposits withdrawn	20.00%	• Measured	Withdrawn / deposited
Deposits after trading began	1	• Measured	
Deposits in a deep drawdown	1	• Measured	
Floating result when printed	-35.00	• Declared	The platform's own summary at the time of the statement
Floating result / balance	-3.91%	• Declared	Floating result / balance

Deposits after trading began, largest first

DATE	AMOUNT	BALANCE BEFORE	DRAWDOWN THEN
2025-02-18	500.00	708.80	-32.03%

Read from the file as uploaded; nothing was checked with the broker.

What data the test ran on

The report header says how prices were simulated, how much of the history the tester had and which dates were tested. It is worth verifying, so here it is checked against its own trades.

Metric	Value	Evidence	Note
Price modelling	—	• Not measured	The report does not state a modelling mode we recognise
Data quality	—	• Not measured	The report prints no data quality (n/a)
Tested from	2020-01-02	• Declared	As printed in the report header
Tested to	2024-11-29	• Declared	As printed in the report header
Trades outside those dates	0	• Measured	Trades that open or close outside the dates the header says were tested

The modelling mode, data quality and stated dates raise no flag.

Read from the header as uploaded: an edit is caught only when the header does not fit itself or the trades.

Stress tests: without the best outcomes

We remove the best periods and trades from what you uploaded and measure what is left. If the total falls to zero or below, it rests on a few events that may not repeat. This is not a forecast.

1 of 8 scenarios ends at zero or below.

On the curve (compounded total return)

Scenario	Left	Change	Still above zero?
Original • Measured	173.7%		
Without the best 1 % of periods (13)	100.4%	-73.3%	• Yes
Without the best 5 periods	140.0%	-33.7%	• Yes
Without the best 10 periods	113.8%	-59.9%	• Yes

SCENARIO	LEFT	CHANGE	STILL ABOVE ZERO?
Without the best month (2021-12)	151.7%	-22.0%	• Yes

On the closed trades (net result after commission and swap)

SCENARIO	LEFT	CHANGE	STILL ABOVE ZERO?
Original • Measured	17,369.75		
Without the best trade	16,897.65	-472.10	• Yes
Without the best 5 trades	15,375.41	-1,994.34	• Yes
Without the best 10 % of trades (129)	-13,338.41	-30,708.16	• No
Without the best month (2022-02)	15,784.89	-1,584.86	• Yes

The best 5 trades add up to this multiple of the net result: 0.11x • Measured

What living through this history was like

A total and a maximum drawdown do not say what the history was like to live through: how long it went without a new high, how long the worst fall took to come back, and what the worst day and month were. These are the numbers that make people switch a system off.

184 days

Longest time without a new high
(3 Aug 2022 to 3 Feb 2023)

• Measured

19 days

Worst fall: days from the high (19 Mar 2020) to the low (7 Apr 2020)

• Measured

36 days

Days from that low back to the high

• Measured

-3.7%

Worst day (24 Feb 2020)

• Measured

-5.5%

Worst month (Nov 2021)

• Measured

68%

Months that ended up (40 of 59);
longest run of losing months: 3

• Measured

2.77

Calmar ratio over 4.9 years:
compound annual return divided
by the deepest fall

• Measured

-1.5%

Average return per day in the
worst 5 % of days (65 of 1282)

• Measured

-3.8%

Average return per month in the
worst 5 % of months (3 of 59)

• Measured

The deepest falls

• Measured

FALL	HIGH TO LOW	BACK AT THE HIGH	TOTAL LENGTH
-8.2%	19 Mar 2020 → 7 Apr 2020 (19 days)	13 May 2020 (36 days)	55 days
-6.9%	30 Jun 2020 → 14 Aug 2020 (45 days)	27 Aug 2020 (13 days)	58 days
-6.8%	21 Sep 2020 → 1 Oct 2020 (10 days)	13 Nov 2020 (43 days)	53 days
-6.7%	28 Oct 2021 → 24 Nov 2021 (27 days)	24 Dec 2021 (30 days)	57 days
-6.5%	24 Dec 2020 → 16 Feb 2021 (54 days)	23 Mar 2021 (35 days)	89 days

Each fall runs from the last point at a high to its lowest point and ends on the first date back at that high. One not back by the file's last date is still open.

The curve is rebuilt from closed trades: open losses do not show, so the real falls lasted and measured at least this much.

Calendar days from the uploaded equity curve; months from each month's last point.

When it wins and when it loses

Your trades grouped by entry day and time. If nearly all the result comes from one day or one session, a change of server time, holidays or news can erase it.

36%

36% of the net result comes
from Thursdays.

• Measured

33%

33% of the net result comes
from the 16:00–19:59 session.

• Measured

ENTRY DAY	TRADES		NET RESULT	WIN RATE
Monday	256		+3,097.50	58%
Tuesday	256		+2,908.32	51%
Wednesday	256		+4,239.72	58%
Thursday	257		+6,331.45	56%
Friday	257		+792.76	50%

ENTRY TIME	TRADES		NET RESULT	WIN RATE
00:00–03:59	206		+4,137.98	56%
04:00–07:59	169		+2,372.34	52%
08:00–11:59	351		+3,055.70	56%
12:00–15:59	165		+2,045.94	53%
16:00–19:59	391		+5,757.79	54%

Entry times as the file states them (platform or server time); net result after the fees the file itemises per trade.

Does it still work in the recent period?

A long history can look good in total while its last stretch no longer adds up. We cut the history's time in three equal stretches and compare the last one with the two before it, trade by trade.

• Holds

The last third of the history shows no drop into losses beyond what chance explains.

+14.94

Average per trade before 11 Apr 2023

• Measured

+10.77

Average per trade since 11 Apr 2023

• Measured

+4,609.62

Net result since 11 Apr 2023 (428 trades)

• Measured

-0.6

Distance between the two averages, in standard errors (-2 or lower: a drop chance hardly explains)

• Measured

EXIT YEAR	TRADES		NET RESULT	WIN RATE
2020	261		+3,827.60	56%
2021	261		+3,438.56	54%
2022	260		+4,734.67	57%
2023	260		+3,828.60	56%
2024	240		+1,540.32	50%

Closed trades by exit date; net result after the fees the file itemises.

Did its average return change at some point?

We look for the moment the curve's average return changed most and measure whether that change is larger than the normal swing of its returns (a CUSUM test that allows for one return influencing the next). It does not change the class.

• No clear change

There is no clear change in the average return across the history ($p = 0.154$): the differences between stretches fit the normal swing of its returns. It does not prove there was none: a small change can go unseen.

• Measured

CUSUM of the returns in time order (Ploberger and Kramer); cautious long-run variance; p-value from the Brownian bridge.

How did it do in the known crises?

The curve's return, from its month-end balances, through each market fall on the public record that it covers in full (the market's peak to its trough). A month with no trades counts as flat. The dates are fixed: they are not fitted to the file.

• Measured

CRISIS	STRATEGY	MARKET OVER THOSE MONTHS
Covid crash 2020-02 – 2020-03	+7.8%	S&P 500 -19.9% Nasdaq Composite -15.8%
Inflation and rates, 2022 2022-01 – 2022-09	+24.2%	S&P 500 -24.8% Nasdaq Composite -32.4%
Crypto winter 2022 2021-11 – 2022-12	+30.9%	Bitcoin (Coinbase) -73.0%

Market: the change from the close of the month before the window to the close of its last month, a fixed historical figure checked on 2026-09-25 against the index levels on FRED ([S&P 500](#), [Nasdaq Composite](#), [Bitcoin \(Coinbase\)](#)); no other data of these indices is read or shown. These are US equities and bitcoin: if the strategy trades another market (currencies, commodities, another country), take them only as context for what the market went through, not as its yardstick.

Worst 12 months in a row: +4.2%; best: +41.7%. 100% of the 12-month periods ended positive. Measured

How did it do in calm and in turbulent markets?

Each return in the file is placed by the VIX (how much the options market expects the S&P 500 to move over the next month) at the close of the market day before it starts: a calm market below 20, a turbulent one from 20. Since 1990 the VIX has closed at 20 or more on about one day in three. Period: 2020-01-01 to 2024-11-29. Measured

	CALM MARKET (VIX < 20)	TURBULENT MARKET (VIX ≥ 20)
Share of the time	51%	49%
Returns counted	661	621
Return per month (compounded)	1.66%	1.79%
Sharpe (return per unit of risk)	1.85	1.74

The gap in mean return between the two columns (0.19 standard errors) is not enough to say it behaves differently depending on the market. Measured

VIX: public data from [FRED](#) (series VIXCLS, from CBOE) read when the report was made. It measures US equities: if the strategy trades another market, read it as a general gauge of fear in markets. It does not change the class.

What was the account worth in your currency and after inflation?

The curve's levels, in dollars, converted at each day's exchange rate (the Federal Reserve's New York noon buying rate), from 2020-01-01 to 2024-11-29. If you live in another currency, this is what the account was worth in it. The difference from the dollar row comes from the exchange rate, not the strategy: when the dollar rises against your currency the result in it rises, and when it falls, it falls. ● Measured

CURRENCY	TOTAL RETURN	A YEAR	WORST FALL
Dollars (the account)	+173.7%	+22.7%	-8.2%
Dollars after US inflation	+123.8%	+17.8%	-7.6%
Mexican pesos (MXN)	+194.8%	+24.6%	-15.5%
Mexican pesos (MXN) after its own inflation	+128.3%	+18.3%	-19.1%
Brazilian reais (BRL)	+308.3%	+33.1%	-15.8%
Brazilian reais (BRL) after its own inflation	+208.2%	+25.7%	-16.1%
Euros (EUR)	+191.3%	+24.3%	-12.3%
Euros (EUR) after its own inflation	+140.2%	+19.5%	-12.6%
Pounds sterling (GBP)	+186.0%	+23.8%	-13.3%
Pounds sterling (GBP) after its own inflation	+129.0%	+18.4%	-15.7%
Japanese yen (JPY)	+278.8%	+31.1%	-13.1%
Japanese yen (JPY) after its own inflation	+246.1%	+28.7%	-13.9%
Canadian dollars (CAD)	+195.9%	+24.7%	-11.3%
Canadian dollars (CAD) after its own inflation	+150.2%	+20.5%	-10.8%
Swiss francs (CHF)	+149.5%	+20.4%	-11.0%
Swiss francs (CHF) after its own inflation	+135.2%	+19.0%	-11.0%

US inflation over those dates was +22.3% in total (+4.2% a year). ● Measured

The rows "after its own inflation" divide by each country's official consumer price index of each month, or that of the latest month published; a currency without a current official index shows only its row before inflation. The return a year is shown from one year of history. Exchange rates and US prices from [FRED](#), read when the report was made. It does not change the class. Consumer prices: [Mexican peso](#), Source: INEGI, Índice Nacional de Precios al Consumidor (INPC), used here to take inflation out of the balances; [real](#), Banco Central do Brasil (IBGE's IPCA); [euro](#), Eurostat (through FRED); [pound](#), Office for National Statistics, licensed under the Open Government Licence v3.0; [yen](#), created by editing Japan's Consumer Price Index (Statistics Bureau, Ministry of Internal Affairs and Communications), through e-Stat; [Canadian dollar](#), Bank of Canada (Statistics Canada's CPI, available free of charge at [bankofcanada.ca](#)); [Swiss franc](#), Eurostat's harmonised index.

How it behaves after losing

What a trading journal would tell you: whether losses are held longer than gains, whether a new trade follows a loss quickly, and how trades do after a losing streak. It does not change the class: these are questions to ask.

• No pattern

Nothing stands out in how it trades after losing.

1.1×

How long a losing trade lasts against a winning one
(median: 4.0 h against 3.7 h) • Measured

53%

Win rate after 2 losses in a row (261 trades; whole history: 55%) • Measured

Closed trades by entry and exit time; net result after the fees the file itemises.

Does it work on each instrument?

When a robot or a signal trades several markets, the total can come from one of them while the others lose. It does not change the class: these are questions to ask.

• Spread out

No single instrument carries the result on its own.

58%

Share of the net result that comes from AUDUSD

• Measured

INSTRUMENT	TRADES		NET RESULT	WIN RATE
AUDUSD	648		+10,057.16	55%
EURUSD	634		+7,312.59	54%

Closed trades by the instrument the file names; net result after the fees the file itemises.

Trade statistics

At the same share of losing trades and in random order, the longest losing run is typically 8 in a row, and 1 history in 20 reaches 11. This history had 7. • Measured

How much of this could be chance?

With 1,282 trades, every figure has a margin. 95 % range: the underlying values consistent with these trades, if each is independent of the others and the system did not change. It is not a prediction.

51.95% – 57.39%Win rate • Measured**6.75 – 20.35**

Expectancy per trade

• Measured**1.32 – 1.76**Profit factor • Measured

METRIC	VALUE	EVIDENCE	NOTE
Trades	1,282	• Measured	
Win rate	54.68%	• Measured	Share of trades with a net profit after the fees the file itemises
Gross profit of winners	77,425.10	• Measured	
Gross loss of losers	-51,081.35	• Measured	
Commission and swap	8,974.00	• Measured	Commission and swap as reported, a positive cost
Net result	17,369.75	• Measured	Gross pnl minus reported fees
Expectancy per trade	13.55	• Measured	Average net result per trade, account currency
Win rate before fees	56.16%	• Measured	Share of trades with pnl > 0
Profit factor	1.52	• Measured	Gross profit / gross loss, before commission and swap; a platform that counts them inside each trade can show a slightly lower figure
Average win	107.53	• Measured	
Average loss	-90.89	• Measured	
Average win / average loss	1.18	• Measured	Average win / average loss
Share of the largest win	0.61%	• Measured	Largest single win / gross profit

METRIC	VALUE	EVIDENCE	NOTE
Most consecutive wins	12	• Measured	
Most consecutive losses	7	• Measured	
Typical longest losing run by chance	8	• Measured	Median longest losing run when trades lose as often as these, in random order
Longest losing run by chance, 1 in 20	11	• Measured	Longest losing run chance reaches once in twenty, at the same loss rate
Chance of a run this long	90.27%	• Measured	Chance of a losing run at least this long, at the same loss rate
Mean hours per trade	3.89	• Measured	
Median hours per trade	3.85	• Measured	
SQN	1.65	• Measured	$\text{Sqrt}(\min(N, 100)) \times \text{mean} / \text{std of per-trade gross pnl}$
Trades per month	21.76	• Measured	First entry to last exit

Long

METRIC	VALUE	EVIDENCE	NOTE
Trades	652	• Measured	
Win rate	53.07%	• Measured	
Net result	5,783.59	• Measured	After the fees the file itemises per trade

Short

Metric	Value	Evidence	Note
Trades	630	Measured	
Win rate	56.35%	Measured	
Net result	11,586.16	Measured	After the fees the file itemises per trade

Resampled one-year risk

6.86%

Maximum drawdown over one year · p50

Measured

12.26%

Maximum drawdown over one year · p95

Measured

16.57%

Maximum drawdown over one year · p99

Measured

Probability of a fall of at least	In the simulations of the history
10%	14.75% Measured
20%	0.20% Measured
30%	0.00% Measured
50%	0.00% Measured

Consecutive periods below the peak, in the simulations: median 72
186.05

Is the file's worst fall normal for these returns?

Worst fall in the file: 8.2%. With the same returns in 1,000 random orders, the worst fall runs from 8.4% to 16.5% in 9 of 10 orders (median 11.4%).

It is milder than in almost every random order: only 4.0% of them fall this little. Losses followed losses less than chance would give. Smoothed curves, curves that average down losing positions, or a favourable order that need not repeat look like this. The file's fall may understate the risk.

Changing the order does not change the Sharpe, the volatility or the final result: it only shows the fall these returns usually bring over the whole file. It is not the one-year fall in the table above.

Assumptions:

- Resampled estimate from the supplied history: it is not a prediction.

- It assumes the future resembles the history; if the market changes, it no longer holds.
- A curve of daily closes does not show floating drawdown within the day.

Lone peak or plateau?

We compare the chosen settings with those one step away on each parameter in your optimisation file. If moving one parameter by one step sinks the result, the settings were fitted to the history's noise.

Chosen: the pass with the highest profit, because the report has no inputs matching a pass.

FastMA 24

SlowMA 55

91%

Of the chosen profit the neighbours keep (median).

• Measured

100%

Of the neighbours end with a profit.

• Measured

• Plateau

The neighbours keep much of the result: it looks like a plateau.

Neighbours one step away

PARAMETER	VALUE	PROFIT
FastMA	23	118.00
SlowMA	54	99.00

METRIC	VALUE	EVIDENCE	NOTE
Profit of the chosen settings	119.00	• Measured	From the rows of the optimisation export
Optimisation passes	120	• Measured	From the rows of the optimisation export
Passes with a profit	99.17%	• Measured	From the rows of the optimisation export
Chosen pass position (top share)	0.83%	• Measured	Rank of the chosen pass / passes

METRIC	VALUE	EVIDENCE	NOTE
Neighbours found	2	• Measured	From the rows of the optimisation export
Neighbours with a profit	100.00%	• Measured	From the rows of the optimisation export
Profit the neighbours keep	91.18%	• Measured	Median neighbour profit / chosen profit

If you turn on the forward period in the MT5 tester and upload that optimisation, the report adds “Does it hold in the forward period?": it compares every pass on the optimised period and on a later one the optimiser did not use to choose.

What is left once luck is discounted?

The more configurations are tried, the higher the best one comes out even when none has an edge. Here the file's Sharpe sits next to what pure luck would give with the configurations counted, using the published math of Bailey and López de Prado and of Harvey and Liu. It is the same calculation that decides the "Number of settings tried" dimension, in numbers.

• Beats luck, without margin

The Sharpe of 1.79 beats the 1.19 that 120 settings with no skill would show, but not by the margin we ask: the confidence that it is not luck (DSR) is 90%, and passing this dimension needs 95%.

1.19

Sharpe 120 settings with no skill would show (the file's: 1.79)

• Measured

2.2 years

Years of history at which that luck falls below this Sharpe (the file has 4.9 years)

• Measured

1.16

Sharpe left after discounting 120 settings (Harvey and Liu)

• Measured

CONFIGURATIONS TRIED	SHARPE LUCK WOULD SHOW	HISTORY NEEDED	IS THIS HISTORY ENOUGH?
10	0.72	10 months	yes
100	1.16	2.1 years	yes
1,000	1.49	3.4 years	yes

E[max Sharpe] of unskilled trials (Bailey & Lopez de Prado); minimum backtest length (Bailey, Borwein, Lopez de Prado & Zhu); Bonferroni haircut (Harvey & Liu).

How much capital it needs, at what size

How much money it takes so that a bad year does not take more than a given share of the account, with this file's trades. We drew 2,000 years of trades at random and took the fall only 5 % of them exceed, or the history's own if larger.

2,442.59

Reference fall in money, at the backtest's size. • Measured

1,528.76

Deepest fall of the history in its own order. • Measured

261

Closed trades per year in the history • Measured

IF YOU ACCEPT
LOSING UP TO 10%

24,425.89

Capital needed at
the backtest's size

0.41x

Size on the file's
starting balance
(10,000)

IF YOU ACCEPT
LOSING UP TO 20%

12,212.94

Capital needed at
the backtest's size

0.82x

Size on the file's
starting balance
(10,000)

IF YOU ACCEPT
LOSING UP TO 30%

8,141.96

Capital needed at
the backtest's size

1.2x

Size on the file's
starting balance
(10,000)

IF YOU ACCEPT
LOSING UP TO 50%

4,885.18

Capital needed at
the backtest's size

2.0x

Size on the file's
starting balance
(10,000)

1x is the backtest's lot size; 0.50x is half of it. Above 1x the fall in money grows in the same proportion.

It counts closed trades only: losses of positions while they were still open are not included, so the capital needed may be larger.

Assumptions:

- Fixed sizes: no compounding and no size change after wins or losses.
- Trades are drawn independently of one another; the history's own fall covers streaks.
- Costs are those the uploaded file itemises.
- Money figures are at the sizes the file used; the relative size is computed on its starting balance, without later deposits.
- It measures the history's losses; it is not a forecast.

Prop-firm challenge simulator

Rules simulated: Generic · Two-step evaluation, phase 1. Generic reference rules, not any one firm's terms.

OUTCOME	IN THE SIMULATIONS OF THE HISTORY
Reaches the target	89.56% • Measured
Breaks the daily loss limit	0.00% • Measured
Breaks the total loss limit	2.28% • Measured

Does not finish in time

8.16% • Measured

88.7% – 90.4%

95 % interval of reaching the target • Measured

58 / 92 / 138

Business days to the target
(p25 / p50 / p75) • Measured

- Reference rules typical of two-step evaluations; not any one firm's terms.

Assumptions:

- Resampled estimate from the supplied history: it is not a prediction.
- Daily data cannot see intraday floating drawdown, so the estimate is optimistic against the daily and total limits.
- It assumes the future resembles the history and that every day with a non-zero return counts as a trading day.

Which firm's rules does your history fit?

The same history, resampled the same way, under each firm's published rules, from most to least likely to pass every phase of the program within the best-day rule, where the firm has one; ties go by name. It compares rules; it does not recommend buying any challenge. • Measured

CHALLENGE	PASSES	PASSES WITHIN THE BEST-DAY RULE	WHAT STOPS IT MOST
FundedNext · Stellar 2-Step 2 phases	90%	no rule	Does not finish in time
FundedNext · Stellar Lite 2 phases	88%	no rule	Breaks the total loss limit
FTMO · FTMO Challenge 2-Step 2 phases	87%	no rule	Does not finish in time
The5ers · High Stakes 2 phases	87%	no rule	Does not finish in time
The5ers · Hyper Growth 1 phase	86%	no rule	Breaks the total loss limit

CHALLENGE	PASSES	PASSES WITHIN THE BEST-DAY RULE	WHAT STOPS IT MOST
FTMO · FTMO Challenge 1 - Step 1 phase	82%	82%	Breaks the daily loss limit
FundedNext · Stellar 1 - Step 1 phase	80%	no rule	Breaks the total loss limit
The5ers · Bootcamp 3 phases	64%	no rule	Breaks the total loss limit
Topstep · Trading Combine 50K 1 phase	60%	60%	Breaks the total loss limit
Topstep · Trading Combine 100K 1 phase	50%	50%	Breaks the total loss limit
Topstep · Trading Combine 150K 1 phase	50%	50%	Breaks the total loss limit

Questions to ask the vendor

1. Is there a live or demo account with at least 11 months of auditable history, with the same robot and settings?
2. Which modelling mode and history quality was the backtest run with (real ticks, 1-minute OHLC, open prices only)?
3. Which spread, commission and swap were used? Are they your broker's?
4. Ask for the (floating) equity curve, not only the balance: the balance hides open losses.

Technical detail by dimension

DIMENSION	STATUS	REASONS
Statistical significance	● Pass	PSR 1.000 >= 0.95; bootstrap p5 Sharpe > 0
Number of settings tried	● Weak	DSR 0.905 between 0.5 and 0.95 with 120 trials counted in the files

Costs	• Weak	Net pnl at 1x is 6,441.05 > 0 but at 3x is -15,404.86 <= 0
Out of sample	• Weak	Out-of-sample Sharpe 0.54 > 0; gap 1.35
Data quality and trading pattern	• Pass	No red flags
Benchmark	• Not applicable	Client declared no applicable benchmark

Thresholds applied: PSR to pass 0.95 · minimum PSR 0.8 · DSR to pass 0.95 · minimum DSR 0.5 · maximum PBO 0.5 · cost multiple it must withstand 3 · minimum out-of-sample Sharpe 0.5 · maximum out-of-sample Sharpe drop 1 · maximum drawdown versus the benchmark (times) 1

Annualised performance

Metric	Value	Evidence	Note
Total return	173.70%	• Measured	
Compound annual return	22.75%	• Measured	
Annual volatility	11.86%	• Measured	
Sharpe	1.79	• Measured	
Sortino	2.80	• Measured	
Maximum drawdown	-8.22%	• Measured	

Statistical significance

Metric	Value	Evidence	Note
Observations	1,282	• Measured	
Sharpe per period	0.1107	• Measured	
Skewness	0.0974	• Measured	
Kurtosis	4.33	• Measured	

METRIC	VALUE	EVIDENCE	NOTE
Probabilistic Sharpe (PSR)	100.00%	• Measured	P[true Sharpe > 0] given length, skew and kurtosis
Minimum track record needed	222	• Measured	Observations needed for PSR to reach 0.95
Observations missing	0	• Measured	

Sharpe corrected for autocorrelation (Lo, 2002): it is not appreciably below 1.79, so the order of the returns does not inflate the plain Sharpe appreciably. When the correction raises it, the report does not use it, so as not to favour the file. • Measured

The returns do not depend on each other appreciably: taking that into account does not change the probability that the true Sharpe is above zero. • Measured

Multiplicity (number of trials)

Trials used in the deflated Sharpe: 120 • Measured passes in the MT5 optimisation export

METRIC	VALUE	EVIDENCE	NOTE
Trials used	120	• Measured	Passes in the MT5 optimisation export
Sharpe variance used	0.0008	• Measured	
Sampling - error floor	0.0008	• Measured	Sampling variance of the Sharpe estimator
Variance increase from dependence	1.03	• Measured	How many times the Sharpe's variance grows when the returns are not taken as independent (1 means no change)
Effective observations after dependence	1,244	• Measured	
Observed across variants	—	• Not measured	No variants uploaded
DSR at the declared trials	90.49%	• Measured	PSR against E[max Sharpe] of 120 trials

METRIC	VALUE	EVIDENCE	NOTE
DSR at the trials used	90.49%	• Measured	PSR against E[max Sharpe] of 120 trials, passes in the MT5 optimisation export
Trials that bring DSR to 0.5	16,384	• Measured	Smallest power-of-two trial count with DSR < 0.5

Variance used: the larger of the one observed across the variants you uploaded and the one sampling error produces.

TRIALS	EXPECTED MAX SHARPE WITHOUT SKILL	DEFLATED SHARPE (DSR)
1	0	100.00%
5	0.0338	99.67%
20	0.0539	97.74%
100	0.0717	91.52%
120	0.0735	90.49%

Stationary bootstrap (per period)

Stationary block bootstrap, resamples: 500 · block 20

	ESTIMATE	P5	P50	P95
Sharpe per period	0.1107 • Measured	0.0688 • Measured	0.1109 • Measured	0.1513 • Measured
Total return	173.70% • Measured	84.39% • Measured	171.78% • Measured	302.55% • Measured

Declared out-of-sample

METRIC	VALUE	EVIDENCE	NOTE
Out-of-sample start	2024-06-03	• Declared	Declared by the client; not verifiable
Sharpe gap (in minus out of sample)	1.35	• Measured	In-sample minus out-of-sample annualised Sharpe

In sample

METRIC	VALUE	EVIDENCE	NOTE
Observations	1,152	• Measured	
Annualised Sharpe	1.89	• Measured	
Probabilistic Sharpe (PSR)	100.00%	• Measured	
Total return	168.44%	• Measured	

Out of sample

METRIC	VALUE	EVIDENCE	NOTE
Observations	130	• Measured	
Annualised Sharpe	0.5448	• Measured	
Probabilistic Sharpe (PSR)	64.96%	• Measured	
Total return	1.96%	• Measured	

Trading costs

METRIC	VALUE	EVIDENCE	NOTE
Reference cost (bps per side)	1	• Declared	Declared by the client; charged on top of the fees the report itemises
Break-even cost (bps per side)	1.59	• Measured	Extra cost per side, on top of the report's fees, at which the ledger nets to zero
Break-even cost multiple	1.59	• Measured	

MULTIPLIER	BPS PER SIDE	GROSS	COST	NET	WIN RATE	TRADES
0x	0.00	26,338.00	8,974.00	17,364.00	54.68%	1,282
1x	1.00	26,338.00	19,896.95	6,441.05	51.48%	1,282
2x	2.00	26,338.00	30,819.91	-4,481.91	48.52%	1,282
3x	3.00	26,338.00	41,742.86	-15,404.86	45.40%	1,282

This table recomputes each trade from its prices and size: with no extra cost it gives 17,364.00, 5.75 away from the trades' net result (17,369.75), from price rounding or currency conversion.

Costs the report itemises

METRIC	VALUE	EVIDENCE	NOTE
Commission	-8,974.00	• Measured	Signed total the report itemises; negative is a cost
Swap	0.00	• Measured	Signed total the report itemises; negative is a cost

Supplied benchmark

• Not measured no benchmark uploaded

Combinatorially symmetric cross-validation (CSCV) overfitting

• Not measured no variants uploaded

Sub-periods (calendar years)

YEAR	RETURN	MAX DRAWDOWN
2020	38.28%	-8.22%

2021	24.87%	-6.66%
2022	27.42%	-5.18%
2023	17.40%	-4.30%
2024	5.96%	-5.63%

Rolling windows

WINDOW	MIN RETURN	MIN DRAWDOWN	SHARE NEGATIVE
65	-4.68%	-8.22%	14.37%
131	-3.29%	-6.87%	2.78%
261	3.66%	-6.66%	0.00%

Red flags

✓ No red flags in the audited files.

Audited files (sha256)

report.html	f07658c4049d4cc344aece0211cf60babe77cc06eaeb6a9f554bf4b1a804f75b
optimization.xml	49227a5fca40ead53d2020bad7b4f19f8723bc7ef6d4079e7ff791e2fce72626
live.csv	0648d915126fac0966c5ca86aae7b6a1f10fc8b03b5f7233e1032d070c1b9a4b
Dataset digest	ced13c6f8a82d6d27bb49448ed86dec90c021c7d46baa31f69d86566e35eaf55

2020-01-01 → 2024-11-29 · daily (trading days) · 1,282 observations · equity curve

Parse warnings

- report: contract size inferred from reported profit: AUDUSD x100,000, EURUSD x100,000
- report: the file's times carry no timezone (platform or server time); they were read as UTC
- report: the balance curve is built from closed trades only; it does not show floating (open-trade) drawdown, so the real drawdown was at least as deep

- optimization: the pass count is the number of configurations the optimiser tried; a genetic optimisation lists only the passes it evaluated
- live: contract size inferred from reported profit: AUDUSD x100,000, EURUSD x100,000
- live: deposits or withdrawals were removed: the curve is a flow-adjusted index that starts at the initial balance
- live: the file's times carry no timezone (platform or server time); they were read as UTC
- live: the balance curve is built from closed trades only; it does not show floating (open-trade) drawdown, so the real drawdown was at least as deep

File format: MetaTrader 5 Strategy Tester (HTML)

Optimisation export: 120 configurations tried Measured

Figures the platform states Declared

Balance cells that do not match	0
End	2024-11-29
Largest balance difference	0.00
Period	H1 (2020.01.02 - 2024.11.29)
Reconstructed final balance	27,369.75
Reported final balance	27,369.75
Start	2020-01-02
Strategy	SyntheticSampleEA
Symbol	EURUSD

Declared by the client

Metric	Value	Evidence	Note
Trials	120	Declared	
Cost per side (bps)	1.00	Declared	
Out-of-sample start	2024-06-03	Declared	

Metric	Value	Evidence	Note
Benchmark applies	no	• Declared	
Initial balance	—	• Not measured	Not declared

No strategy description was written.

Not measured

Supplied benchmark
No benchmark uploaded.
Combinatorially symmetric cross-validation (CSCV) overfitting
No variants uploaded.

Declared holdout seal

Seal id	sample
Selection start	2020-01-01
Selection end	2024-06-02T23:59:59Z
Holdout start	2024-06-03
Holdout end	2024-11-29
Sealed at (UTC)	2026-09-24
Seal (sha256)	6c8325dd86a2f85c1c008ce937dc7426a1984a5d2f7ccc27a3a322c9e60a8fcc

Notice. This audit is a statistical research tool applied to client-supplied data. It is not investment advice, executes no trades, holds no funds or keys, and does not predict future results. Every value carries its evidence tag: Measured was computed from the file, Declared was asserted by the client and could not be verified, Not measured could not be computed from what was supplied.

sha256 of the audit JSON bc3f6d2923074aa2aff786714f98f62a8cb11ef5600d4b859b86a304b66a6acd