

MS-LIVE-OHLCV-BACKTEST-1-IMPL — Real-Data Result

Datum: 2026-06-04 (~19:00 UTC) **Modus:** Read-only Backtest auf historische MS-Dry-Run Candidates · 0× Order · 0× Code-Touch im Bot **Datenquelle:** Binance Public Spot API · 5-min OHLCV · 4 h Window pro Candidate **Window:** 2026-05-28 07:37 UTC → 2026-06-04 15:37 UTC

1. Executive Summary

MS-Live darf in der aktuellen Form NICHT aktiviert werden. Aggregat-Level versagen alle 6 Strategien die Akzeptanz-Schwellen. Aber: **Symbol-Level zeigt klare Edge** für TON, WLD, MEME, HBAR — kompensiert durch katastrophale False-Positives auf XPL, POL.

Empfehlung: weder Activate noch komplett Verwerfen — **MS-CURATED-SYMBOL-ALLOWLIST-1** (P1, neuer Plan) prüfen, der MS-Live nur auf statistisch profitable Symbole zulässt.

2. Datenbestand

Metrik	Wert
TRADE_CANDIDATEs extrahiert	646
OHLCV-Fetches erfolgreich	646 (100 %)
Fetches fehlgeschlagen	0
Backtest-Window pro Candidate	4 h × 5 min = 48 Bars
Total Backtest-Laufzeit	207 s
Datenquelle	Binance Public Spot

3. Aggregat-Ergebnisse pro Strategy

Variant A — RAW (alle 646 Candidates, kein Dedup)

Strategy	N	WR	Exp %	PF	TP	SL	TO	Verdict
trend_follow	620	38.55 %	-0.241	0.856	13	47	560	NO_GO
breakout	23	0.00 %	-2.144	0.0	0	3	20	DEFER (N<50)
volatility_sweep	3	0.00 %	-0.494	0.0	0	0	3	DEFER (N<30)
mean_reversion	0	—	—	—	—	—	—	NO_DATA
vwap_mean_reversion	0	—	—	—	—	—	—	NO_DATA
oversold_bounce	0	—	—	—	—	—	—	NO_DATA

Variant B — POST-RECON (Dedup + Stablecoin-Block aktiv)

Dedup-Statistik: - 20 Candidates durch Stablecoin-Block weggefiltert - 413 durch SYM:120-min Cooldown weggefiltert - 144 durch SYMSTRAT:360-min Cooldown weggefiltert - **69 verbleibende Candidates** (10.7 % des Original-Samples)

Strategy	N	WR	Exp %	PF	Verdict
trend_follow	67	37.31 %	-0.272	0.840	DEFER (N<100, ähnliche Edge wie RAW)
breakout	1	0.00 %	-3.365	0.0	DEFER (N<<50)
volatility_sweep	1	0.00 %	-0.347	0.0	DEFER (N<<30)

Verdict pro Strategy (kombiniert)

Strategy	Akzeptanz-Schwelle	RAW erreicht?	POST_RECON erreicht?	Final
trend_follow	N≥100 / WR≥35% / Exp≥+0.3% / PF≥1.1	N ✓ WR ✓ Exp X PF X	N X	NO_GO
breakout	N≥50	X (23)	X (1)	DEFER — mehr Daten

volatility_sweep	N≥30	X (3)	X (1)	DEFER — mehr Daten
mean_reversion	N≥50	X (0)	X (0)	NO_DATA — Strategy hat nie ausgelöst
vwap_mean_reversion	N≥50	X (0)	X (0)	NO_DATA — OFF default
oversold_bounce	N≥30	X (0)	X (0)	NO_DATA — RSI<25-Gate zu strikt für aktuellen Markt

4. Regime-Slicing (RAW)

Regime	N	WR	Exp %	PF	Timeout
STRONG_TREND	357	38.7 %	-0.361	0.800	88.5 %
WEAK_TREND	286	35.3 %	-0.244	0.841	92.0 %
RANGE	3	0.0 %	-0.493	0.0	100.0 %
HIGH_VOL / CHOP	0	—	—	—	—

Befund: STRONG_TREND und WEAK_TREND verhalten sich sehr ähnlich (Edge gleichermaßen negativ). Regime ist **kein** primärer Edge-Treiber.

5. Symbol-Level Breakdown — der eigentliche Insight

Top-10-Symbole (RAW)

Symbol	N	WR %	Exp %	PF	Total Net %	Bewertung
TON/USDT	45	75.6	+2.136	4.639	+96.14	✓ Premium
WLD/USDT	42	59.5	+2.064	2.092	+86.69	✓ Premium
MEME/USDT	49	61.2	+0.846	1.753	+41.47	✓ Stabil
HBAR/USDT	39	59.0	+0.820	3.057	+31.99	✓ Stabil
ICP/USDT	41	56.1	+0.062	1.070	+2.53	- Grenzwertig
NEAR/USDT	36	52.8	-0.029	0.982	-1.05	- Neutral
XLM/USDT	36	27.8	-0.330	0.844	-11.88	X Schwach
POL/USDT	29	10.3	-0.420	0.063	-12.19	X Toxisch
FET/USDT	90	42.2	-0.209	0.876	-18.84	X Toxisch
XPL/USDT	37	5.4	-3.134	0.003	-115.96	X Katastrophal

Aggregat Top-10: -201 USDT net (von -200 total → Top-10 dominiert das Sample).

Kernerkenntnis: - 4 Symbole liefern **+256 USDT** kumuliert (TON+WLD+MEME+HBAR) - 4 Symbole liefern **-159 USDT** kumuliert (XPL+POL+FET+XLM) - XPL allein zerstört -116 USDT — das ist mehr als die Top-2 Gewinner kombiniert

Die MS-Strategy hat **Edge auf Symbol-Level**, aber das Sample-Level wird durch die Toxic-Symbols kannibalisiert.

Top-Symbole (POST-RECON, dedup-bereinigt)

Symbol	N	WR %	Exp %	PF
HBAR/USDT	2	100.0	+1.347	inf
ICP/USDT	4	75.0	+0.854	10.856
WLD/USDT	7	57.1	+2.821	3.24
TON/USDT	4	50.0	+1.719	2.47
NEAR/USDT	3	33.3	+0.412	1.642
XLM/USDT	4	25.0	+0.221	1.176
FET/USDT	9	33.3	-0.313	0.815
POL/USDT	3	33.3	-0.260	0.457
MEME/USDT	5	20.0	-1.442	0.202
XPL/USDT	4	25.0	-2.674	0.030

Nach Dedup sind die Sample-Größen je Symbol so klein, dass Aussagen schwach werden. Aber die **Rangfolge** bleibt erhalten:

6. Overall PnL-Simulation

Variante	N	TP	SL	TIMEOUT	WR	Total Net %
RAW	646	14	50	582	37.0 %	-200.35
POST_RECON	69	2	6	61	36.2 %	-21.90

Bei **200 USDT Position-Size pro Trade** ergäbe RAW ein theoretisches PnL von **~-400 USDT** in 7 Tagen. **POST-RECON** **~-44 USDT** in derselben Zeit — Dedup reduziert den Schaden 10x, aber nicht profitabel.

7. Timeout-Rate = strukturelles Problem

Metrik	RAW	POST_RECON
TIMEOUT-Rate	90.1 %	88.4 %

90 % aller Candidates erreichen weder TP noch SL binnen 4 Stunden. Die Strategy plant TP/SL zu eng (kleine ATR-Multiples) für die tatsächliche Marktvolatilität. Bei Timeout-Close zu Last-Price ist die MtM-Verteilung negativ-skew.

Implication: Strategy-Parameter (sl_atr_mult, tp_atr_mult) sind nicht real-data-tuned.

8. RAW vs POST-RECON Comparison

Aspekt	RAW	POST-RECON
Sample-Größe	646	69 (-89 %)
Total Net	-200 %	-22 %
Win-Rate	37.0 %	36.2 %
Expectancy	-0.31 %	-0.32 %
Trades pro Tag	≈ 93	≈ 10
Verdict	NO_GO	NO_GO (zu klein für GO)

Befund: Dedup ändert die Per-Trade-Edge **kaum** (-0.31 vs -0.32 %). Es reduziert nur **Frequency**. Die strukturelle Edge-Schwäche bleibt.

→ **REPEAT-CANDIDATE-DEDUP-1 schützt vor Frequenz-Bleeding, behebt aber nicht die TP/SL-Tuning-Lücke.**

9. Empfehlung pro Strategy

Strategy	Verdict	Begründung	Nächster Schritt
trend_follow	NO_GO	N=620 statistisch stark, aber Exp=-0.24% und PF=0.86	TP/SL-Re-Tuning + Symbol-Allowlist
breakout	DEFER	N=23 zu klein; 0 % Wins	Mehr Markt-Daten (4 Wochen wait)
volatility_sweep	DEFER	N=3 unbrauchbar	Re-eval bei N≥30
mean_reversion	NO_DATA	0 Candidates in 9 d	Strategy-Logic-Audit (RangeTrading-V1 zu strikt?)
vwap_mean_reversion	NO_DATA	0 Candidates	OFF — pre-existing-Erkenntnis
oversold_bounce	NO_DATA	RSI<25-Gate filtert allen Markt aus	Strategy-Threshold-Lockerung erwägen

10. Empfohlene Folge-Phasen

ID	Prio	Zweck
MS-CURATED-SYMBOL-ALLOWLIST-1	P1	Allowlist {TON, WLD, MEME, HBAR, ICP} für MS-Live; alle anderen Symbole bleiben Dry-Run-only. Plan + 2-w-Monitor.

MS-TP-SL-RETUNING-1	P2	Per-Strategy ATR-Multiplier re-calibrate gegen tatsächliche MFE/MAE-Distribution aus diesem Backtest. Hauptziel: TIMEOUT-Rate < 50 %.
MS-WEEKLY-BACKTEST-RERUN-1	P3	Backtest weekly automatisch laufen lassen, in <code>bot_statuses.metadata.json.ms_backtest_weekly</code> exposen.
OVERSOLD-BOUNCE-RSI-RELAX-1	P3	Investigieren ob RSI<25-Gate ZU strikt ist (0 Hits in 9 d).
MS-LIVE-OHLCV-BACKTEST-1	done	Audit + Implementation abgeschlossen, weitere Re-Runs als Sub-Phasen.

11. Limitierungen + Bias-Hinweise

#	Limitierung	Auswirkung
1	4-h Window — viele Candidates brauchen länger	Timeout-Rate evtl. überzeichnet; 8-h-Run als Sensitivity nötig
2	Worst-Case-Modell SL-vor-TP bei gleichem Bar	Echte Outcomes evtl. minimal besser (TP-First selten)
3	Mainnet-OHLCV vs Testnet-Execution	Annahme: identische Spread-Strukturen; in Praxis kann Testnet-Slippage abweichen
4	Strategy-Defaults haben sich im Window geändert (RangeTrading-V1 ersetzte mean_reversion am 31.05)	Phase-D-Daten für mean_reversion sind frühe Lifecycle
5	0.3 % Fee+Slippage konservativ	Reale Costs könnten niedriger sein, aber Testnet-Slippage unklar
6	Cluster-Bias (TON 45 Hits in 2-h-Window)	TON-Edge möglicherweise zeitlich-konzentriert nicht repräsentativ
7	TIMEOUT-MtM kann positiv/negativ je nach Bar-Close-Drift	90 % Timeout-Rate macht das zur dominierenden Komponente

12. Boundaries dieser Implementation

0x Bot-Code-Touch · 0x Trading-State-Touch · 0x Orders · 0x MS-Live-Aktivierung · 0x Mainnet · 0x Env-Änderung · 0x DB-Write · 0x Bot/Worker-Recreate · 0x ConfigProfile-Apply · 0x Push.

Erstellte Dateien (alle lokal in /tmp/ms_backtest): - `candidates.json` — 646 extrahierte Candidates - `cache/` — 646 OHLCV-Snapshots - `outcomes.csv` — Per-Candidate Backtest-Outcomes - `report.json` — strukturiertes Aggregat - `report.md` (diese Datei) - PDF — separat publiziert

13. Operator-Entscheidung

Option	Aktion
A	GO MS-CURATED-SYMBOL-ALLOWLIST-1 PLAN (P1 neu) — restriktiv MS-Live auf TON/WLD/MEME/HBAR/ICP zulassen
B	GO MS-TP-SL-RETUNING-1 PLAN (P2 neu) — Strategy-Parameter re-tunen
C	Bundle A+B: Symbol-Allowlist + Re-Tuning kombiniert
D	Idle — Dry-Run weiter, 2 w neue Daten sammeln, Re-Run
E	MS komplett deaktivieren (<code>MULTI_STRATEGY_ENABLED=false</code>) — Legacy-Only-Mode

Default-Empfehlung: A (P1-Allowlist mit nachweislich profitablen Symbolen, geringes Risiko).

STOP

Backtest abgeschlossen, Roadmap-Update wird vorbereitet, **kein Push, keine MS-Live-Aktivierung**, kein Roadmap-Commit ohne separates Operator-GO.