

# MS-LIVE-OHLCV-BACKTEST-1 — Audit & Plan

**Datum:** 2026-06-04 (~18:30 UTC) **Modus:** READ-ONLY · Audit + Backtest-Methodologie · keine Code/Order-Aktion **Quelle:** MS-Dry-Run-Window 2026-05-26 06:32 UTC → 2026-06-04 ≈ **9 Tage 12 h Update:** ersetzt initialen Audit nach RECON-1-Cutover + 4 P0/P1-Commits

## 1. Ziel

Wahre Edge der 6 MS-Strategien quantifizieren, **bevor** `MULTI_STRATEGY_DRY_RUN=false` gesetzt wird. Der bisherige `final_score` -Proxy ist unzureichend (siehe `final_score=9.1` für 100 % stable U/USDT — Score sagt **nichts** über tatsächliche Profitabilität).

**Hartes Akzeptanzkriterium für MS-Live-Aktivierung:** je-Strategie real-data-validated Win-Rate  $\geq 50\%$  und Expectancy  $> 0$  USDT auf historischen Candidates.

## 2. Settings-Snapshot (aktuell live)

### MS-Pipeline

Setting	Wert	Bedeutung
<code>MULTI_STRATEGY_ENABLED</code>	True	Pipeline läuft
<code>MULTI_STRATEGY_DRY_RUN</code>	<b>True</b>	nur Logging, kein <code>execute_buy</code>
<code>MULTI_STRATEGY_TESTNET_RELAXED</code>	True	niedrigere Score-Schwellen für Testnet
<code>MULTI_STRATEGY_DRY_RUN_MAX_SYMBOLS</code>	10	Top-10 Universum pro Scan
<code>MULTI_STRATEGY_MAX_OPEN_POSITIONS</code>	10	Slot-Cap
<code>MULTI_STRATEGY_MAX_EXPOSURE_PCT</code>	0.40	40 % Portfolio-Exposure-Cap
<code>MULTI_STRATEGY_MAX_RISK_PCT</code>	0.02	2 % Risk-per-Trade
<code>MULTI_STRATEGY_TIER_HIGH/MID/LOW_PCT</code>	0.09 / 0.06 / 0.03	Position-Size-Tiers
<code>MULTI_STRATEGY_CASH_RESERVE_PCT</code>	0.05	5 % Cash-Floor
<code>MULTI_STRATEGY_FEE_BUFFER_PCT</code>	0.002	0.2 % Fee-Buffer
<code>MULTI_STRATEGY_QUOTE_ASSET</code>	USDT	Spot-Quote

### Regime + Stability

Setting	Wert
<code>REGIME_STABILITY_CANDLES</code>	1 (MS-RELAX-1: gesenkt von 3)
<code>REGIME_HIGH_VOL_ATR_RATIO</code>	1.5
<code>REGIME_RANGE_EMA_SPREAD_MAX</code>	0.03
<code>TESTNET_SCORE_THRESHOLD</code>	6.5 (STRAT-RECON-Phase-1: gesenkt von 7.5)

### NEU seit 2026-06-04 (Commit c4651dd)

Setting	Wert	Wirkung
<code>MS_CANDIDATE_COOLDOWN_ENABLED</code>	True	Repeat-Spam-Filter aktiv
<code>MS_CANDIDATE_COOLDOWN_SYMBOL_MIN</code>	120	max 1 Candidate je Symbol pro 2 h
<code>MS_CANDIDATE_COOLDOWN_SYMBOL_STRATEGY_MIN</code>	360	max 1 Candidate je Symbol+Strategy pro 6 h
Stablecoin-Block ( <code>STABLECOIN_BASE_BLOCKLIST</code> )	+ RLUSD, BFUSD, XUSD, GUSD, USDB	7+ Stables vor-Strategie-Filter

### Per-Strategy Toggles

Strategy	Aktiviert?	Bemerkung
<code>trend_follow</code>	✓	STRONG_TREND + WEAK_TREND
<code>breakout</code>	✓	WEAK_TREND

mean_reversion (RangeTrading-V1)	✓	RANGE (since 2026-05-31)
vwap_mean_reversion	✓ ( ENABLE_VWAP_STRATEGY=True )	RANGE
volatility_sweep	✓	RANGE + HIGH_VOL/RANGE-base
oversold_bounce	✓ (Shadow-V1)	WEAK_TREND + RANGE

## Hard-Blocks

- CHOP-Regime: alle Strategien aus
- HIGH\_VOL ohne sauberes base\_regime: alle aus
- `is_stable=False` : alle aus

## 3. Datenbestand zum Backtest

### Volumen

- **73 418 MS-Log-Records** (Decisions, Errors, Heartbeats)
- **646 TRADE\_CANDIDATES** mit voller Trade-Spec
- **100 % der 646** haben `entry`, `stop_loss`, `take_profit` gesetzt → direkter Backtest ohne Annahmen möglich
- Window: 2026-05-28 07:37 UTC → 2026-06-04 15:37 UTC (≈ 7 d 8 h aktive Candidates)

### Per-Strategy

Strategy	Candidates	Anteil	Zeitspanne
<b>trend_follow</b>	<b>620</b>	96.0 %	28.05 07:37 → 04.06 15:37
breakout	23	3.6 %	29.05 06:24 → 04.06 00:59
volatility_sweep	3	0.5 %	31.05 21:54 → 31.05 21:59 (alle ICP/USDT)
mean_reversion	0	0 %	— (alter N7.1 + RangeTrading-V1 hatten keine Hits)
vwap_mean_reversion	0	0 %	—
oversold_bounce	0	0 %	— (RSI<25 / RSI<30+reversal-Gate zu strikt für den Markt)

### Top-Symbole pro Strategy

**trend\_follow:** FET ×90 · MEME ×49 · TON ×45 · WLD ×42 · HBAR ×39 · ICP ×38 · POL ×36 · NEAR ×36 · MEME-cluster ×31 (TOP 9 = 70.6 % Volumen) **breakout:** INJ ×15 · WLFI ×8 **volatility\_sweep:** ICP ×3 (alle in 5-min-Window am 31.05 21:54-21:59)

### Sample-Size-Bewertung

Strategy	N	Statistical-Power-Bewertung
trend_follow	620	<b>sehr stark</b> — 99 %-CI ± ~5 % auf Win-Rate
breakout	23	<b>schwach</b> — 99 %-CI ± ~25 % auf Win-Rate; nur Trend-Indikation
volatility_sweep	3	<b>unzureichend</b> — keine statistische Aussage möglich
mean_reversion, vwap, oversold_bounce	0	<b>kein Backtest möglich</b> — Strategie hatte keine Hits

## 4. Backtest-Methodologie

### 4.1 Pro-Candidate-Simulation (Hauptmetrik)

Für jeden der 646 Candidates:

1. **OHLCV-Fetch** ab `candidate.ts` für Window N (Vorschlag: 4 h × 5 min = 48 Bars).
2. **Bar-Walk-Forward** (chronologisch):
3. Wenn `low ≤ stop_loss` UND `low ≤ take_profit` (gleicher Bar): konservativ als **SL-HIT** werten (Worst-Case-Modell).
4. Wenn `low ≤ stop_loss`: **SL-HIT** → `outcome = (stop_loss - entry) / entry`.
5. Wenn `high ≥ take_profit`: **TP-HIT** → `outcome = (take_profit - entry) / entry`.
6. Sonst: nächster Bar.
7. **Time-Out:** erreicht weder TP noch SL im Window → `outcome` bei letzter Bar-close-price `(close - entry) / entry` (Mark-to-Market).

### 4.2 Aggregat-Metriken pro Strategy

Metrik	Definition
win_rate	$\text{count}(\text{outcome} > 0) / \text{count}(*)$
avg_win_pct	$\text{mean}(\text{outcome} \mid \text{outcome} > 0)$
avg_loss_pct	$\text{mean}(\text{outcome} \mid \text{outcome} < 0)$
expectancy_pct	$\text{win\_rate} \times \text{avg\_win} - (1 - \text{win\_rate}) \times \text{avg\_loss}$
profit_factor	$\text{sum}(\text{outcome} > 0) / \text{abs}(\text{sum}(\text{outcome} < 0))$
MFE_avg	$\text{mean}(\text{max}(\text{high} - \text{entry}) / \text{entry})$ über Window
MAE_avg	$\text{mean}(\text{max}(\text{entry} - \text{low}) / \text{entry})$ über Window
time_to_tp_median	Median-Anzahl-Bars bis TP-Hit
time_to_sl_median	Median-Anzahl-Bars bis SL-Hit
timeout_rate	$\text{count}(\text{neither TP nor SL hit}) / \text{count}(*)$

### 4.3 Fee/Slippage-Modell

Komponente	Wert
Entry-Fee	0.1 %
Exit-Fee	0.1 %
Slippage (konservativ)	0.05 % je Seite
<b>Total Roundtrip-Cost</b>	<b>0.3 %</b>

Netto-Outcome:  $\text{gross\_outcome} - 0.003$  für TP-Hit / SL-Hit (Time-Out: 0.002 wegen einseitiger Tighter-Modell).

### 4.4 Dedup-Counterfactual (zwei Läufe)

**Variante A — Raw:** alle 646 Candidates ohne Dedup (entspricht Pre-RECON-1-Verhalten). **Variante B — Post-RECON-1:** Candidates nach simuliertem  $\text{SYM}:\{\text{symbol}\}$  120-min +  $\text{SYMSTRAT}:\{\text{symbol}\}:\{\text{strategy}\}$  360-min Cooldown filtern (entspricht aktuellem Bot-Verhalten ab 15:28 UTC). Schätzung: **~110 verbleibende Candidates** (aus PDF-Eval).

Vergleich beider Varianten gibt Antwort: würde Dedup die Edge-Quote verbessern oder den Sample-Pool zu stark schrumpfen lassen?

### 4.5 Regime-Slicing

Per-`regime` Outcome-Breakdown: - STRONG\_TREND: erwartete Edge-Konzentration - WEAK\_TREND: gemischt - RANGE: nur volatility\_sweep (3 Candidates → unbrauchbar) - HIGH\_VOLATILITY: separate Bewertung

## 5. Datenquelle OHLCV

### Anonyme Binance-Spot-API

```
GET https://api.binance.com/api/v3/klines
?symbol=XLMUSDT
&interval=5m
&startTime={ms}
&endTime={ms+4h}
```

Item	Wert
Rate-Limit (anonymous)	1200 req/min · 6000 req/5min
646 Candidates × 1 Fetch	≈ 33 s @ 50% Auslastung
Datenkost	0 € (Public API)
Auth nötig	Nein
Universe	Binance-Mainnet-Public (= reale historische Preise) — kein Testnet-OHLCV (Testnet hat synthetisches Spread-Modell, unbrauchbar)
Test-Symbol	XLMUSDT live-verified

### Backup-Plan

Falls Rate-Limit-Issue: lokales `data-fetcher`-Module mit `requests_cache` für 24 h Cache.

## 6. Implementations-Skizze (out-of-scope für Audit)

```
# scripts/ms_live_ohlcv_backtest.py
def backtest_candidate(c, window_hours=4, bar_minutes=5):
    bars = binance_fetch(c['symbol'], c['ts'], window_hours, bar_minutes)
    for b in bars:
        if b.low <= c['stop_loss']:
            return ('SL', (c['stop_loss'] - c['entry']) / c['entry'])
        if b.high >= c['take_profit']:
            return ('TP', (c['take_profit'] - c['entry']) / c['entry'])
    return ('TIMEOUT', (bars[-1].close - c['entry']) / c['entry'])
```

Output: Pandas-DataFrame mit allen 646 Outcomes, Aggregat-Report pro Strategy + Regime.

**Implementations-Phase wäre separat als P3-Commit `ms-live-ohlcv-backtest-impl-1` mit Plan-vor-Code.**

## 7. Erwartete Ergebnisse (qualitative Prognose)

### trend\_follow (620 Candidates)

- Erwartung: Win-Rate 35-50 %, Profit-Factor 1.0-1.4
- Begründung: Trend-Following hat strukturell niedrige Hitrate aber höhere RR
- Falls Win-Rate < 35 % UND Profit-Factor < 1.0 → **MS-Live für trend\_follow abgelehnt**
- Risk-Distribution wichtig: TON-Cluster (45 Hits in kurzer Zeit) sollte 99 %-CI nicht dominieren — Cluster-Bootstrap-CI empfohlen

### breakout (23 Candidates)

- Sample zu klein für harte Aussage
- Direktional-Indikator: wenn alle 23 verloren hätten → klares "Nein"; wenn ≥ 15 gewonnen → "Lokales Setup gut, mehr Daten nötig"

### volatility\_sweep (3 Candidates)

- Nicht backtestbar — alle 3 in einem 5-min-Fenster, gleicher Symbol (ICP)
- Empfehlung: **OFF in MS-Live** bis ≥ 30 Candidates akkumuliert

### mean\_reversion, vwap, oversold\_bounce (0 Candidates)

- Strategien haben in Live-Markt nie ausgelöst
- Können nicht backtestet werden
- Empfehlung: **OFF in MS-Live** + Audit der Strategy-Logic (möglicherweise zu strikte Filter)

## 8. Akzeptanz-Schwellen für MS-Live

Strategy	Min-N	Min-Win-Rate	Min-Expectancy	Min-Profit-Factor
trend_follow	100	35 %	+0.3 % netto	1.1
breakout	50	40 %	+0.4 % netto	1.2
volatility_sweep	30	45 %	+0.5 % netto	1.3
mean_reversion	50	50 %	+0.3 % netto	1.1
vwap_mean_reversion	50	50 %	+0.3 % netto	1.1
oversold_bounce	30	45 %	+0.4 % netto	1.2

**Ein Strategy passiert MS-Live nur wenn ALLE 4 Spalten erfüllt sind.**

## 9. Limitierungen + Bias-Vermeidung

#	Limitierung	Mitigation
1	Survivorship-Bias (Binance delisted Pairs nicht im Backtest)	Filter: nur Pairs verwenden, die heute noch handelbar sind (übereinstimmend mit Live-Universum)
2	Look-Ahead-Bias via <code>take_profit</code> aus	Candidate hat SL/TP zum Zeitpunkt <code>ts</code> berechnet — kein Lookahead. ✓ sauber

	Strategy-Code	
3	Slippage zu optimistisch	Konservativ 0.05 % je Seite eingerechnet; bei dünnen Pairs (BFUSD/XPL) erhöhen
4	Worst-Case-SL-vor-TP bei gleichem Bar	Modell-Default; alternativ Tick-Data-Modus (out-of-scope)
5	Cluster-Korrelation (TON 45 Hits)	Cluster-Bootstrap-CI je Symbol+1h-Window
6	Backtest-Window 4 h fix	Sensitivity-Analyse mit 2 h / 8 h / 24 h
7	Pre-Cutover-Period nicht repräsentativ	Post-Cutover (15:28 UTC am 2026-06-04) Vergleichs-Run wenn $\geq 50$ neue Candidates
8	Testnet-Sandbox-Quirks	Backtest auf Mainnet-OHLCV (Public), Live-Execution dann auf Testnet — Konsistenz-Annahme dokumentieren

## 10. Deliverables

#	Output	Format
1	Per-Candidate-CSV (646 rows $\times$ ~15 Spalten)	CSV
2	Per-Strategy-Report-Markdown	MD
3	Aggregat-PDF mit Grafiken (Win-Rate-Histogram, MFE/MAE-Scatter, Profit-Curve)	PDF
4	Roadmap-Update MS-LIVE-OHLCV-BACKTEST-1 $\rightarrow$ status=done	ROADMAP.php Edit
5	Operator-Empfehlung pro Strategy: GO/NO-GO/DEFER	Section in PDF

## 11. Operator-Entscheidungen vor Implementation

Q	Frage	Default-Empfehlung
Q1	Window-Größe?	<b>B</b> 4 h (50/50 zwischen TP-time und SL-time)
Q2	Bar-Größe?	<b>A</b> 5 min (genug Auflösung, Binance liefert $\geq 1000$ Bars je Request)
Q3	SL-vor-TP bei gleichem Bar?	<b>A</b> SL-HIT (konservativ)
Q4	Fee/Slippage?	<b>A</b> 0.3 % roundtrip (siehe §4.3)
Q5	Dedup-Variante A (Raw) oder B (Post-RECON-1)?	<b>C</b> beide für Vergleich
Q6	Akzeptanz-Schwellen wie in §8?	<b>A</b> ja, sonst Modifikation pinnen
Q7	Implementation-Phase als nachgelagerter GO oder nun gebündelt?	<b>A</b> nachgelagert — Audit hier ist Read-Only-Plan
Q8	PDF-Output auf files.gewerbespeicher-rechner.de?	<b>A</b> ja

## 12. Boundaries dieses Audits

0 $\times$  Code-Touch · 0 $\times$  State-Edit · 0 $\times$  Bot/Worker-Recreate · 0 $\times$  Env-Änderung · 0 $\times$  DB-Migration · 0 $\times$  MS-Live-Aktivierung · 0 $\times$  Mainnet · 0 $\times$  Order · 0 $\times$  OHLCV-Fetch (das passiert in der separaten Implementation-Phase).

## 13. Bisherige Erkenntnisse aus MS-LIVE-READINESS-FULL-EVAL-1 PDF (2026-06-04)

Aus dem 222-h-Audit (Vorgänger-Phase, PDF veröffentlicht 13:00 UTC):

Befund	Konsequenz für Backtest
95.8 % aller Candidates aus trend_follow	Backtest fokussiert primär auf trend_follow
3 P0-Blocker identifiziert	2 davon gefixt (DEDUP + STABLECOIN) — Variante-B-Run zeigt Wirkung
Stable-Leakage 20 Candidates (USDE/RLUSD/BFUSD)	bereits geblockt — Variante B würde diese 20 ausschließen
Sample-Konzentration (TON 45, FET 30, HBAR 32)	Cluster-Bootstrap zwingend
MS-MTF-1 Wirkung gering (17 % statt 50-70 %)	Eigener Sub-Audit "MS-MTF wirklich nötig?" möglich

## 14. Empfehlung

Item	Empfehlung
MS-Live-Aktivierung heute	NEIN — kein Backtest-Beleg vorliegend
Backtest-Implementation	als separater Operator-GO-Punkt MS-LIVE-OHLCV-BACKTEST-1-IMPL (P2, $\leq 2$ h Aufwand)
Backtest-Run-Cadence	erstes Run heute, weekly Re-Run mit akkumulierten Daten
MS-Live-Re-Evaluation	Nach Backtest-Ergebnis: pro-Strategy GO/NO-GO/DEFER-Entscheidung
Roadmap-Status	bleibt planned bis Implementation-GO; Audit-Notes erweitert (siehe §16)

## 15. Beobachtbarer Status post-RECON-1 (Side-Channel-Beleg)

Seit RECON-1-Cutover (17:01 UTC) bis 17:42 UTC (40 min): - **174** MS-Log-Records - **0** neue TRADE\_CANDIDATEs - **18** Stablecoin-Pre-Filter-REJECTs (RLUSD 12, USDE 4, EUR 2) - **0** repeat\_candidate\_cooldown REJECTs (kein Repeat-Trigger weil keine Candidates)

→ Markt war im Eval-Fenster aktuell trade-arm. Backtest auf vollem 9-d-Window bleibt der primäre Hebel.

## 16. Roadmap-Update

ROADMAP.php Task MS-LIVE-OHLCV-BACKTEST-1 wird aktualisiert: - **status:** weiterhin planned (Audit ist Read-Only, Implementation ist separater Schritt) - **progress:** 0 → 30 (Audit + Methodologie fertig, Plan-Phase abgeschlossen) - **notes:** ergänzt um Datenbestand-Snapshot (646 Candidates), Methodologie-Outline (§4), Akzeptanz-Schwellen (§8), Sample-Size-Bewertung (§3) - **next\_action:** „Implementation-Phase ms-live-ohlcv-backtest-impl-1 als separater Operator-GO mit Plan-vor-Code“

PDF veröffentlicht: [files.gewerbespeicher-rechner.de/MS\\_LIVE\\_OHLCV\\_BACKTEST\\_AUDIT\\_{TS}.pdf](files.gewerbespeicher-rechner.de/MS_LIVE_OHLCV_BACKTEST_AUDIT_{TS}.pdf)

## STOP

Audit abgeschlossen. **Keine** Implementation, **kein** Code-Touch, **kein** OHLCV-Fetch. Operator-GO für MS-LIVE-OHLCV-BACKTEST-1-IMPL erforderlich, um zur tatsächlichen Backtest-Phase überzugehen.