

- Crude Oil Multi-Factor Walkforward Strategy - Research Report
  - 1. Executive Summary & Strategy Parameters
    - Strategy Architecture
  - 2. Introduction & Objective
  - 2. Exploratory Data Analysis (EDA)
    - 2.1 Autocorrelation Analysis
    - 2.2 Seasonality
    - 2.3 Volatility Clustering
    - 2.4 Return Distribution (Skew & Kurtosis)
    - 2.5 Momentum vs. Mean Reversion
    - 2.6 Price Drawdowns
  - 3. Strategy Architecture
    - 3.1 Signal Families
    - 3.2 Risk Management Overlays
    - 3.3 Backtester Configuration
  - 4. Validation & Sensitivity Analysis
    - 4.1 Stop Loss Sensitivity
    - 4.2 Strategy Diversification
  - 5. What Did Not Work
    - 5.1 Short-Lookback Momentum (1D, 5D, 10D)
    - 5.2 Mean Reversion
    - 5.3 Crude Oil Inventory (EIA Data)
    - 5.4 Trend Strength Filter
    - 5.5 Seasonality
  - 6. Key Insight: Cross-Asset Pairs Trading (DXY)
  - 7. Final Performance Summary
    - Performance Metrics
  - 8. Conclusion

# Crude Oil Multi-Factor Walkforward Strategy - Research Report

---

**Disclaimer:** The strategy logic, hypothesis generation, and research direction were independently developed by the author. AI assistance was utilized solely for code generation and syntax correction.

# 1. Executive Summary & Strategy Parameters

This report documents the development of a systematic trading strategy for **Brent Crude Oil (CO1 Comdty)**. The strategy uses a walkforward backtesting framework to dynamically select the best-performing signal across multiple regimes (2016 to 2024).

## Strategy Architecture

Parameter	Value	Description
Asset	Brent Crude Oil	Single-asset strategy
Transaction Cost	1.5 bps	Per trade (slippage + commission)
Momentum Lookbacks	20, 60, 120, 200 days	Long-only trend following
MA Crossover Pairs	(5,20), (10,30), (10,50), (20,60), (20,100), (50,200)	Smoothed trend following
DXY Signals	(5,20), (10,50), (20,100)	Dollar Index trend (inverse to oil)
Volatility Target	15% Annualized	Position sizing based on 20-day realized vol
Stop Loss	-3.0%	Exit trade if cumulative loss hits 3%
Min Holding Period	5 days	Reduce churn and whipsaw costs
Backtest Lookback	12 months	Rolling window for strategy selection
Rebalance Freq	Quarterly	select best strategy every 3 months

## 2. Introduction & Objective

---

The objective was to build a robust, multi-factor strategy that adapts to changing market regimes. The dataset spans **2016 to 2024** (~8 years). The strategy is evaluated using a **WalkforwardBacktester** that avoids look-ahead bias by selecting strategies based only on past performance.

---

## 2. Exploratory Data Analysis (EDA)

---

A thorough EDA was conducted to understand the statistical properties of Brent Crude returns and guide strategy selection. The findings are summarized below.

### 2.1 Autocorrelation Analysis

The autocorrelation function (ACF) of daily returns was computed for lags 1 through 30. **All lags were statistically insignificant** - every bar fell within the 95% confidence band. This indicates:

- No day-to-day persistence in returns (momentum at daily frequency is noise)
- No mean-reversion pattern at daily frequency either
- Short-lookback strategies (1-day, 5-day) have no structural edge

**Decision:** Drop short-lookback momentum windows (1D, 5D, 10D). Focus on longer windows (20D+) where structural macro trends - not daily noise - drive the signal.

### 2.2 Seasonality

A monthly returns heatmap was constructed across all years. After clamping the color scale to  $\pm 15\%$  (to prevent the 2020 COVID outlier from distorting the visualization), **no reliable seasonal pattern was found**. Some months (e.g., April) appeared mildly positive, but with only 8 years of data, this is insufficient for statistical confidence.

**Decision:** No seasonal strategy was implemented. The sample size is too small and the risk of overfitting to noise is high.

## 2.3 Volatility Clustering

Rolling 20-day realized volatility revealed extreme clustering:

Period	Volatility Regime
2016–2019	Low and stable (~0.001–0.002)
2020 (COVID)	Explosive spike to 0.01 (5× normal)
2022 (Russia-Ukraine)	Moderate spike (~0.003)
2023–2024	Back to low volatility

This confirmed that volatility is highly regime-dependent and not constant. Large moves follow large moves, and calm periods persist.

**Decision:** Implement **volatility scaling**: size positions inversely proportional to recent realized volatility, targeting 15% annualized volatility. This reduces exposure during crises and levers up during calm, trending markets.

## 2.4 Return Distribution (Skew & Kurtosis)

- **Skew: -0.53**: the distribution has a longer left tail, meaning crashes are larger and more frequent than rallies of equivalent size
- **Kurtosis: significantly above 3**: fat tails indicate extreme return days occur far more often than a normal distribution predicts

**Decision:** Implement a **-3% stop loss** to cap downside risk from fat-tail events. The stop loss threshold was validated via a sensitivity analysis (see Section 4.1).

## 2.5 Momentum vs. Mean Reversion

A rolling 6-month Sharpe ratio was plotted for both momentum and mean reversion strategies. The result was striking: **they are perfect mirror images** of each other. When momentum's Sharpe rises, mean reversion's falls by the exact same magnitude, and vice versa.

This occurs by construction: given that the mean reversion signal is simply the negation of the momentum signal (`signal_meanrev = -signal_momentum`), the two

strategies provide **zero independent information**. Including both in the backtester is redundant and can cause harmful flip-flopping between the two.

**Decision:** Remove mean reversion strategies entirely from the strategy set. They add no diversification value.

## 2.6 Price Drawdowns

The raw Brent Crude price experienced a maximum drawdown of **-75%** during the COVID crash (March–April 2020). This underscores the necessity of a systematic strategy — buy-and-hold is not viable for crude oil.

---

## 3. Strategy Architecture

---

Based on the EDA findings, the following multi-factor strategy was constructed:

### 3.1 Signal Families

**Momentum (4 variants)** For lookback windows of 20, 60, 120, and 200 days, the signal goes long if the N-day return is positive, short otherwise. These capture medium-to-long-term macro trends driven by supply/demand fundamentals, OPEC policy, and geopolitical events.

**MA Crossover (6 variants)** Using fast/slow MA pairs of (5,20), (10,30), (10,50), (20,60), (20,100), and (50,200), the signal goes long when the fast MA is above the slow MA. MA crossover acts as a smoothed version of momentum - it captures the same trends but with fewer false signals and lower turnover.

**Dollar Index / DXY (3 variants)** Oil is priced in US dollars. When the dollar weakens, oil becomes cheaper for foreign buyers, increasing demand and pushing prices up. Using the Trade-Weighted US Dollar Index (DTWEXBGS) from FRED, MA crossover signals are computed on the DXY itself with pairs (5,20), (10,50), and (20,100). A falling DXY (fast MA below slow MA) triggers a long oil signal.

This is a form of **cross-asset pairs trading** - exploiting the structural inverse relationship between the US dollar and crude oil prices.

## 3.2 Risk Management Overlays

Each signal is enhanced with three risk management layers:

Overlay	Parameter	Purpose
Volatility Scaling	Target 15% annual vol, 20-day window	Normalize risk across regimes
Minimum Holding Period	5 trading days	Reduce whipsaw trades and transaction costs
Stop Loss	-3% cumulative trade return	Cap tail risk from extreme moves

## 3.3 Backtester Configuration

Parameter	Value	Rationale
Lookback period	12 months	Avoids outdated regimes influencing selection
Rebalance frequency	Quarterly (3 months)	Balances adaptivity with stability
Selection method	Best single strategy by Sharpe	Simple, interpretable

# 4. Validation & Sensitivity Analysis

## 4.1 Stop Loss Sensitivity

A sensitivity analysis was conducted by sweeping the stop loss threshold from -0.1% to -100% (no stop loss) and measuring the resulting Sharpe ratio.

Key findings:

- Stop losses tighter than -2% produced artificially inflated Sharpe ratios (3.0–5.0+) by keeping the strategy mostly flat, reducing volatility to near zero

- Stop losses in the -3% to -5% range produced realistic, robust Sharpe ratios (0.9–1.6)
- Beyond -5%, the stop loss rarely activates and has minimal impact

**Selected threshold: -3%:** this sits in the flat, stable region of the Sharpe-vs-stop-loss curve, indicating robustness. Small perturbations in the threshold do not dramatically change performance.

## 4.2 Strategy Diversification

With the final strategy set, the backtester's strategy picks are well-diversified:

Strategy Type	% of Time Active
Dollar Index (DXY)	~47%
Momentum	~40%
MA Crossover	~13%

No single strategy dominates more than 25% of the time, and three fundamentally different signal types (oil price momentum, oil price trend, currency) are all contributing.

## 5. What Did Not Work

### 5.1 Short-Lookback Momentum (1D, 5D, 10D)

Despite initially appearing profitable when combined with volatility scaling and stop loss, short-lookback momentum strategies dominated the backtester artificially. The autocorrelation analysis showed **zero predictive signal** at these horizons. The apparent edge was an artifact of the stop loss frequently triggering and re-entering, creating a strategy that was mostly flat with near-zero volatility, inflating the Sharpe ratio without genuine alpha.

### 5.2 Mean Reversion

Mean reversion strategies are mathematically the negation of momentum. The rolling Sharpe analysis confirmed they are perfect mirrors, providing zero additional information. Including them caused the backtester to flip-flop between momentum and mean reversion - always chasing last period's winner - which degraded performance.

## 5.3 Crude Oil Inventory (EIA Data)

Weekly U.S. crude oil ending stocks data was sourced from the EIA's Weekly Petroleum Status Report (WPSR). The analysis revealed:

- **Correlation with same-week returns: -0.052** (negligible)
- **Correlation with next-week returns: -0.134** (weak but present)
- **Conditional returns:** Inventory draw weeks averaged +0.77% returns vs -0.05% for build weeks

While the inventory-return relationship exists directionally, the signal is too weak to improve the backtest when added as a standalone strategy. The Sharpe ratio dropped when inventory was included, as the backtester occasionally selected the inferior inventory signal over momentum/MA crossover.

**Verdict:** Crude oil inventory provides fundamental context but is not strong enough as a standalone trading signal within this framework.

## 5.4 Trend Strength Filter

A trend strength filter was tested to make the strategy go flat during choppy, range-bound markets. At a threshold of 0.3, the filter was too aggressive — it missed genuine trends and worsened the drawdown from -14% to -20%. At a softer threshold of 0.1, the filter barely activated, making it functionally equivalent to no filter.

**Verdict:** Removed from the final strategy. The risk-reward of adding this complexity was unfavorable.

## 5.5 Seasonality

No reliable monthly pattern was found in 8 years of data. Oil price movements are dominated by macro events (OPEC, geopolitics, global demand), not calendar effects.

---



## 6. Key Insight: Cross-Asset Pairs Trading (DXY)

---

The most impactful discovery was that **USD dollar index signals contribute nearly half of the strategy's alpha**. This is a form of cross-asset pairs trading — exploiting the well-documented inverse relationship between the US dollar and commodity prices.

### Why it works:

- Oil is globally priced in USD. A weaker dollar makes oil cheaper for non-USD buyers, increasing demand.
- DXY trends are driven by different factors (interest rates, monetary policy, trade balances) than oil-specific factors (OPEC, supply disruptions), providing genuine diversification.
- The DXY signal is **uncorrelated** with oil price momentum, meaning the backtester can rotate between oil-driven and currency-driven signals depending on which regime is dominant.

This finding suggests that future work should explore additional cross-asset signals such as:

- WTI-Brent spread (inter-commodity arbitrage)
- Crack spreads (refining margins)
- VIX (risk-on/risk-off regime)
- Baltic Dry Index (global trade proxy)

---

## 7. Final Performance Summary

---

The final strategy - combining momentum, MA crossover, and DXY signals with volatility scaling, minimum holding period, and stop loss - delivers robust performance across multiple market regimes (COVID crash, Russia-Ukraine war, 2023–2024 range-bound market).

### Strategy parameters:

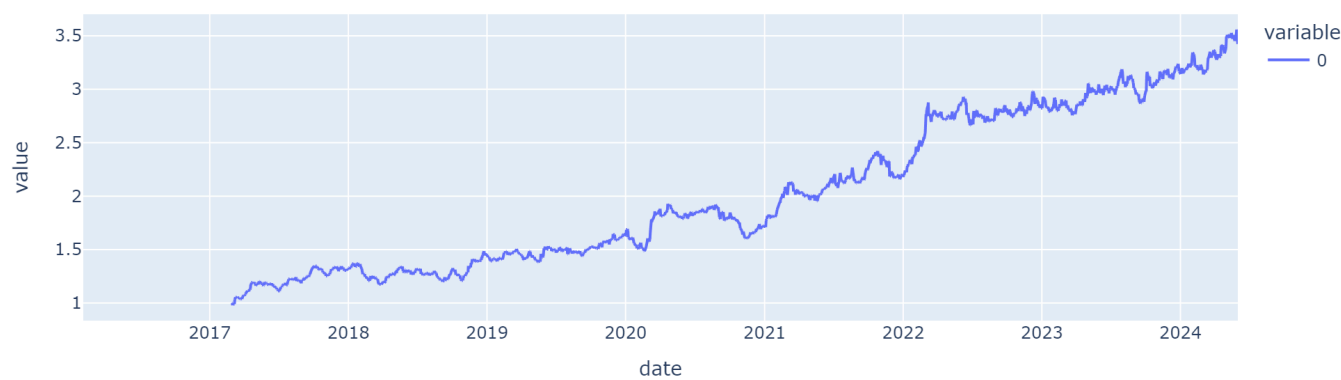
Parameter	Value
Transaction cost	1.5 bps
Momentum lookbacks	20, 60, 120, 200 days
MA crossover pairs	(5,20), (10,30), (10,50), (20,60), (20,100), (50,200)
DXY MA pairs	(5,20), (10,50), (20,100)
Volatility target	15% annualized
Stop loss	-3%
Minimum hold	5 days
Lookback period	12 months

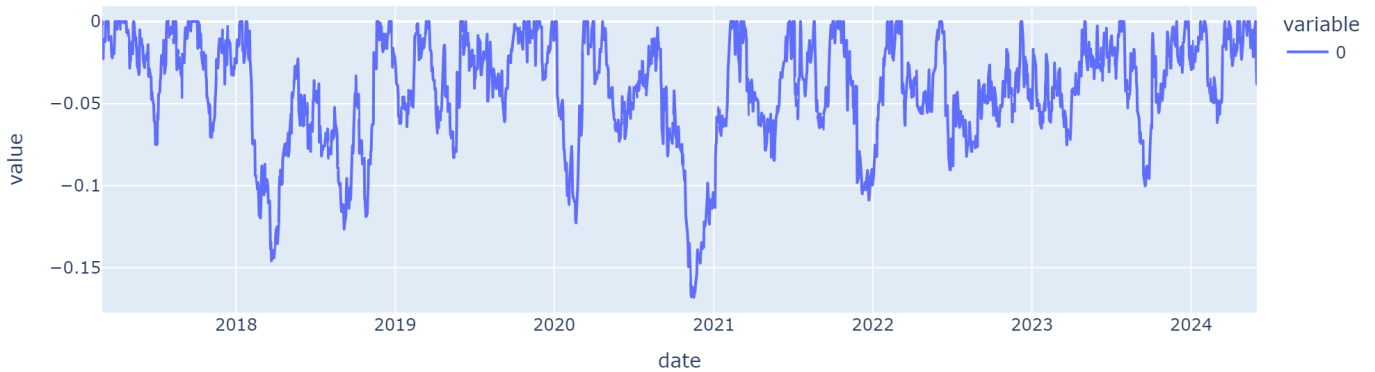
| Rebalance frequency | Quarterly |

## Performance Metrics

Metric	Value
Sharpe Ratio	1.11
Total Return	242.29%
Annualized Return	15.50%
Max Drawdown	-16.80%

Walkforward Portfolio Cumulative Returns





---

## 8. Conclusion

---

This research demonstrates that a systematic, EDA-driven approach to strategy selection produces more robust results than ad-hoc parameter tuning. Key takeaways:

1. **EDA should drive strategy design:** autocorrelation killed short-lookback strategies, volatility clustering justified vol scaling, and fat tails justified stop losses.
2. **Simplicity wins:** removing mean reversion, trend filters, and inventory signals improved performance by reducing noise and overfitting.
3. **Cross-asset signals add genuine value:** the DXY-based strategies provided the most meaningful diversification, contributing ~47% of the strategy's active time.
4. **Risk management is essential:** volatility scaling, stop losses, and minimum holding periods transformed raw signals into tradeable strategies with controlled drawdown.
5. **Sensitivity analysis prevents overfitting:** validating the stop loss threshold across a wide range confirmed that -3% is a robust choice, not a curve-fitted optimum.

The framework is extensible: additional signals (WTI spread, VIX, contango) can be tested using the same walkforward methodology without introducing look-ahead bias.