Flux Market Pressure Index: A Contrarian Indicator for Oil Swap Markets

Harinder Sandhu, Flux Insights

January 22, 2026

Abstract

Flux Insights presents V1 of the Market Pressure Index (MPI), a contrarian indicator backed by Onyx's market leading proprietary positioning and flow dataset, to predict price reversals in oil swap markets. We first discuss the approach implemented in constructing the index, present case studies on utilising the MPI as a discretionary trading tool, and finally develop a systematic MPI-based contrarian strategy across a basket of products, enhanced by technical indicators, showcasing 1.6 Sharpe in our training/validation set and 1.2 Sharpe out-of sample.

Contents

1 Introduction

Market sentiment is a critical driver of commodity prices, yet extreme sentiment often precedes reversals rather than trend continuations. The Market Pressure Index (MPI) is designed as a contrarian indicator that synthesizes multiple dimensions of trader behavior: Entry prices, positioning, and open interest, to identify when markets have reached unsustainable extremes.

The methodology draws on the principle that when prices deviate significantly from trader entry levels, while positioning becomes crowded, the market becomes vulnerable to reversals. This vulnerability is amplified when open interest is at historical extremes, indicating maximum trader commitment and reduced capacity for further trend extension.

The backbone of the index involves calculating price deviations from entry levels observed from Onyx's proprietary Commitment of Traders (COT) data in market-making oil derivatives. With market shares upwards of 30% across the oil derivatives complex, Onyx is capable of extracting highly meaningful counterparty entry levels, doing so in timely fashion with T-2 latency, which we prove remains a **measurable edge in providing trading insight**.

2 Theoretical Rationale

2.1 Why Contrarian Indicators Work

Contrarian indicators are predicated on the behavioral finance principle that markets tend to overshoot fundamental equilibrium due to herding behavior and momentum trading. When the majority of traders hold similar positions (e.g., everyone is long), several dynamics emerge.

First, the exhaustion of buying power, as more traders accumulate long positions, fewer new buyers remain to elevate prices, particularly as open interest in the contract begins to exceed historically high thresholds. The MPI seeks to implement the level of saturation as part of the signal and thereby quantify the extent to which prices are overextended.

Second, provided the majority of positions added are congruent with price direction, for example, long positions being added by speculative players as prices trend upwards, sellside pressure may emerge as these players look to lock in profits, resulting in mean reversion of prices, manifesting as the breakdown and reversal of a trend.

Third, with most trending contracts, technical moving averages or statistical methods involving deviations of prices away as a function of their standard deviation such as Bollinger Bands, provide an esimate of the ceiling/floor to a trend. In isolation, these techniques can prove problematic, since they largely hinge on psychological factors from market participants, without meaningful supportive data to suggest traders are seeking to add to or liquidate positions.

The MPI exploits these dynamics by identifying situations where current prices have deviated substantially from average trader entry levels, suggesting that many positions are either significantly profitable (vulnerable to profit-taking) or underwater (vulnerable to forced liquidation).

2.2 Utilisation of Entry Price Data

Traditional sentiment indicators that rely on sources of positioning data in the industry such as net long/short ratios, derived from highly lagged CFTC positioning data, struggle to separate signal from noise and fail to generate consistent, reliable indications of future market states. The MPI enhances this by not only integrating more recent positioning data into signals, but also providing quantitative metrics for trader entry levels, enabling another dimension of analysis beyond standard market data.

7-day entry prices: Captures short-term momentum and recent commitment

90-day entry prices: Reflects medium-term positioning and trends

365-day entry prices: Provides long-term context and structural position growth

By comparing current prices to these entry levels, the MPI quantifies the unrealized profit/loss embedded in the market, which directly links with reversal risk.

3 Methodology

The MPI is constructed through the following procedure, converting market data from the Flux COT API into a normalized indicator.

3.1 Price Deviations from Entry Levels

The foundation of the MPI is in measuring how far the current price has deviated from average trader entry prices for both long and short positions. For a given time window w (7-day, 90-day, or 365-day):

3.1.1 Long Entry Deviation

$$D_{\text{long}}^{(w)} = \frac{P_t - E_{\text{long}}^{(w)}}{E_{\text{long}}^{(w)}} \times 100 \tag{1}$$

where P_t is the current market price at time t, and $E_{\text{long}}^{(w)}$ is the average entry price for long positions over window w. A positive $D_{\text{long}}^{(w)}$ indicates that current prices are trading above levels where long traders entered, suggesting profitable long positions and therefore potential downward profit-taking pressure. This is a **bearish** signal in our contrarian strategy.

3.1.2 Short Entry Deviation

$$D_{\text{short}}^{(w)} = \frac{P_t - E_{\text{short}}^{(w)}}{E_{\text{short}}^{(w)}} \times 100 \tag{2}$$

where $E_{\text{short}}^{(w)}$ is the average entry price for short positions over window w. A positive $D_{\text{short}}^{(w)}$ suggests that current prices trade above short trader entries, meaning short positions are underwater, creating potential for short covering (buying to close shorts), which is a **bullish** signal.

3.1.3 Net Deviation

The net pressure is computed as the average of the two deviations:

$$D_{\text{net}}^{(w)} = \frac{D_{\text{long}}^{(w)} + D_{\text{short}}^{(w)}}{2} \tag{3}$$

This composite measure captures the overall price pressure: positive values indicate bullish deviation (prices above average entries), while negative values indicate bearish deviation. This approach is one of many that may be adopted to quantify the level of pressure present in the market arising from profit taking or stop-out potential.

3.2 Moving Average Smoothing

The raw daily deviations calculated with the above approach can be noisy due to intraday volatility and transient price movements. In order to reduce noise while maintaining responsiveness, we apply a 10-day moving average to the deviation series. This aims to filter out day-to-day randomness that doesn't reflect genuine sentiment shifts, whilst remaining short enough to capture meaningful sentiment changes, and so allowing time for a given trend to over-extend and be primed for reversal.

$$D_{\text{MA}}^{(w)}(t) = \frac{1}{10} \sum_{i=0}^{9} D_{\text{net}}^{(w)}(t-i) \tag{4}$$

3.3 Scaling Factors

Open interest (OI) represents the total number of outstanding contracts and serves as a key measure of market participation and commitment. When OI is at historical highs, the market is maximally risk-on and vulnerable to reversals. Conversely, low OI indicates limited participation and so dampens signal reliability, as price movements do not reflect players who are vulnerable or profitable with large market positions.

We compute the OI ratio relative to the 5-year maximum:

$$R_{\mathrm{OI}}(t) = \frac{\mathrm{OI}(t)}{\max_{s\in[t-1825,t]} \mathrm{OI}(s)} \tag{5}$$

where the denominator is the maximum OI observed over the preceding 5 years for the same contract.

3.3.1 OI Scaling Logic

The deviation is scaled based on the OI ratio:

$$D^{(w)}_{\text{OI-scaled}}(t) = \begin{cases} D^{(w)}_{\mathrm{MA}}(t) \times R_{\mathrm{OI}}(t) & \text{if } R_{\mathrm{OI}}(t) > 1.0 \quad \text{(AMPLIFY)} \\ D^{(w)}_{\mathrm{MA}}(t) \times [1 - (1 - R_{\mathrm{OI}}(t)) \times 0.5] & \text{if } R_{\mathrm{OI}}(t) \leq 1.0 \quad \text{(DAMPEN)} \end{cases} \tag{6}$$

Amplification: ($R_{\mathrm{OI}} > 1$): When OI exceeds its 5-year maximum, the deviation is multiplied by the OI ratio. For example, if OI is 120% of its historical max ($R_{\mathrm{OI}} = 1.2$), the deviation is amplified by 1.2×. This reflects heightened reversal risk when the market is overcrowded.

Dampening ($R_{\mathrm{OI}} \leq 1$): When OI is below historical norms, the deviation is reduced. For instance, if OI is 50% of its max ($R_{\mathrm{OI}} = 0.5$), the dampening factor is $1-(1-0.5)\times 0.5 = 0.75$. This acknowledges that low participation reduces the significance of price deviations for purposes of capturing a reversal - furthermore, this enables the index to avoid the beginnings of trend formation when the open interest tends to remain around seasonal average levels.

3.3.2 Positioning Adjustment

The core positioning insight is that reversals are most likely when positioning and price deviations are aligned. For example, if the market is heavily long (net positive positioning) and prices are above entry levels (positive deviation), this represents an **overbought** condition ripe for a downward reversal. If the market is heavily short (net negative positioning) and prices are below entry levels (negative deviation), this represents an **oversold** condition ripe for an upward reversal.

Conversely, when positioning and deviation oppose each other (e.g., long positioning but prices below entries) players face downward stop-out risk, liquidating length by selling into the market, which would cause prices to accelerate lower and thereby **deepen the downward trend**. Given this, and that the MPI is contrarian in nature, the index is skewed in the **opposite direction**, towards neutrality, to reflect the divergence in positioning and deviation and thus the uncertainty in a contrarian move.

3.3.3 Position Ratio

We define the net position ratio as:

$$R_{\mathrm{pos}}(t) = \frac{\mathrm{NetPosition}(t)}{\mathrm{OI}(t)} \tag{7}$$

where $\mathrm{NetPosition}(t)$ is defined as the total net positioning in the market

This ratio ranges from -1 (fully short) to $+1$ (fully long), with 0 indicating balanced positioning.

3.3.4 Alignment Scaling

The alignment between positioning and deviation is captured by their product:

$$A(t) = R_{\text{pos}}(t) \times D^{(w)}_{\text{OI-scaled}}(t) \tag{8}$$

If $A(t) > 0$: Position and deviation have the same sign (aligned) $\rightarrow$ **Amplify toward extremes**

If $A(t) < 0$: Position and deviation have opposite signs (opposed) $\rightarrow$ **Dampen toward neutral**

3.3.5 Contrarian Scaling

$$D^{(w)}_{\text{contrarian}}(t) = \begin{cases} D^{(w)}_{\text{OI-scaled}}(t) \times [1 + |R_{\text{pos}}(t)| \times 0.8] & \text{if } A(t) > 0 \quad \text{(ALIGNED)} \\ D^{(w)}_{\text{OI-scaled}}(t) \times [1 - |R_{\text{pos}}(t)| \times 0.4] & \text{if } A(t) \leq 0 \quad \text{(OPPOSED)} \end{cases} \tag{9}$$

Example: Suppose $R_{\text{pos}} = 0.6$ (60% net long) and $D^{(w)}_{\text{OI-scaled}} = +15\%$ (prices 15% above entries). These are aligned ($A > 0$), so:

$$D^{(w)}_{\text{contrarian}} = 15\% \times (1 + 0.6 \times 0.8) = 15\% \times 1.48 = 22.2\%$$

The deviation is amplified by 48%, reflecting the heightened reversal risk from crowded long positioning at elevated prices.

3.4 Construct Pressure Index (0-100)

The final step transforms the contrarian-adjusted deviation into a percentile-based index. To avoid look-ahead bias, we use an **expanding window** rank that only considers data available up to time t:

$$\text{MPI}^{(w)}(t) = \text{Percentile}\left(D^{(w)}_{\text{contrarian}}(t), \{D^{(w)}_{\text{contrarian}}(s) : s \leq t\}\right) \times 100 \tag{10}$$

This yields an index from 0 to 100, where:

- **MPI > 80**: Overextended bullish sentiment (prices far above entries, positioning crowded long) $\rightarrow$. **Expect downward reversal**
- **MPI < 20**: Overextended bearish sentiment (prices far below entries, positioning crowded short) $\rightarrow$. **Expect upward reversal**
- **20 $\leq$ MPI $\leq$ 80**: Neutral zone with no strong directional conviction

4 Case Study: April 2025 US Tariff Period

4.1 Context

The 2 April 2025 "Liberation Day" package marked the formal pivot from Trump's earlier, more targeted trade actions into an explicitly system-wide reciprocal regime, with China singled out for the harshest treatment and pushed to an effective rate of about 54% on most imports once combined with prior levies. China's State Council Tariff Commission responded on 4 April with a blanket tariff on US goods layered on top of pre-existing energy duties, driving a structurally higher landed cost for US LPG and other hydrocarbons into China and forcing a rapid rethink of forward trade flows into East Asia. Within days, macro and sell-side forecasters were trimming China's growth outlook by around 1–2.5 percentage points, reinforcing expectations of weaker downstream petrochemical demand just as tariff barriers on US-origin feedstocks were hardening.

The LPG complex reacted quickly once it became clear that Beijing's retaliation explicitly hit US LPG, an export stream worth on the order of tens of billions of dollars annually and tightly linked to China's propane dehydrogenation (PDH) build-out. The Argus Far East Index (FEI) – a key benchmark for delivered propane and butane into East Asia and the settlement basis for a large share of physical and paper LPG trade – became the main transmission channel for this shock. As Chinese buyers stepped back and began to re-optimise toward Middle Eastern (Saudi CP) and other non-US supply, FEI-linked propane values came under pressure, with prices dropping sharply as traders priced in both a hole in Chinese pull and the prospect of US cargoes being re-routed toward Korea, Japan, India, and Southeast Asia, to discover new outlets.

We examine how the MPI performed during this period to validate its alerting capabilities. Since the MPI adopts an expanding window, we need to consider a wider timeframe prior to liberation day and the ensuing volatility,

4.2 Data and Timeframe

The analysis covers March 1, 2025 to April 30, 2025, a period encompassing:

- **Pre-tariff phase**: Early March, with stable market conditions
- **Tariff announcement**: Late March-Early April, triggering initial volatility
- **Peak tariff uncertainty**: Early to Mid April with extreme price movements
- **Post-tariff adjustment**: Mid to late April, as markets digested the new regime

We showcase the MPI using the three aforementioned periods to capture the market state in the C3 FEI contract prior to and during the tariff announcements, and the market response thereafter.

Figure 1: C3 Far East Index Propane May 2025 response to initial liberation day reciprocal tariff announcements from China. Source: Flux Terminal

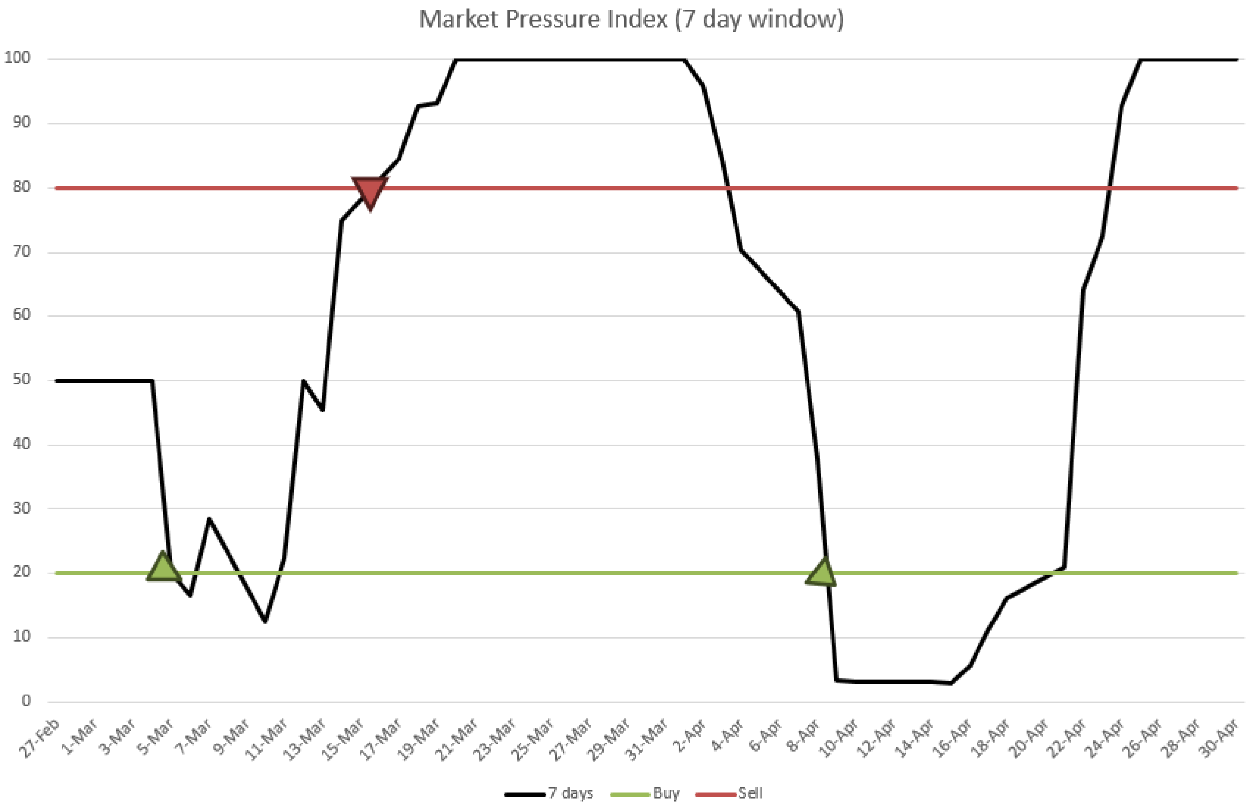


Figure 2: 7-Day Market Pressure Index during April 2025 US Tariff Period for May C3 FEI. Red dashed line at 80 indicates overextension, downward reversal inbound). Green dashed line at 20 indicates downward overextension. Gray dotted line at 50 represents neutral sentiment. 7D overextension noted on Mar 17th.

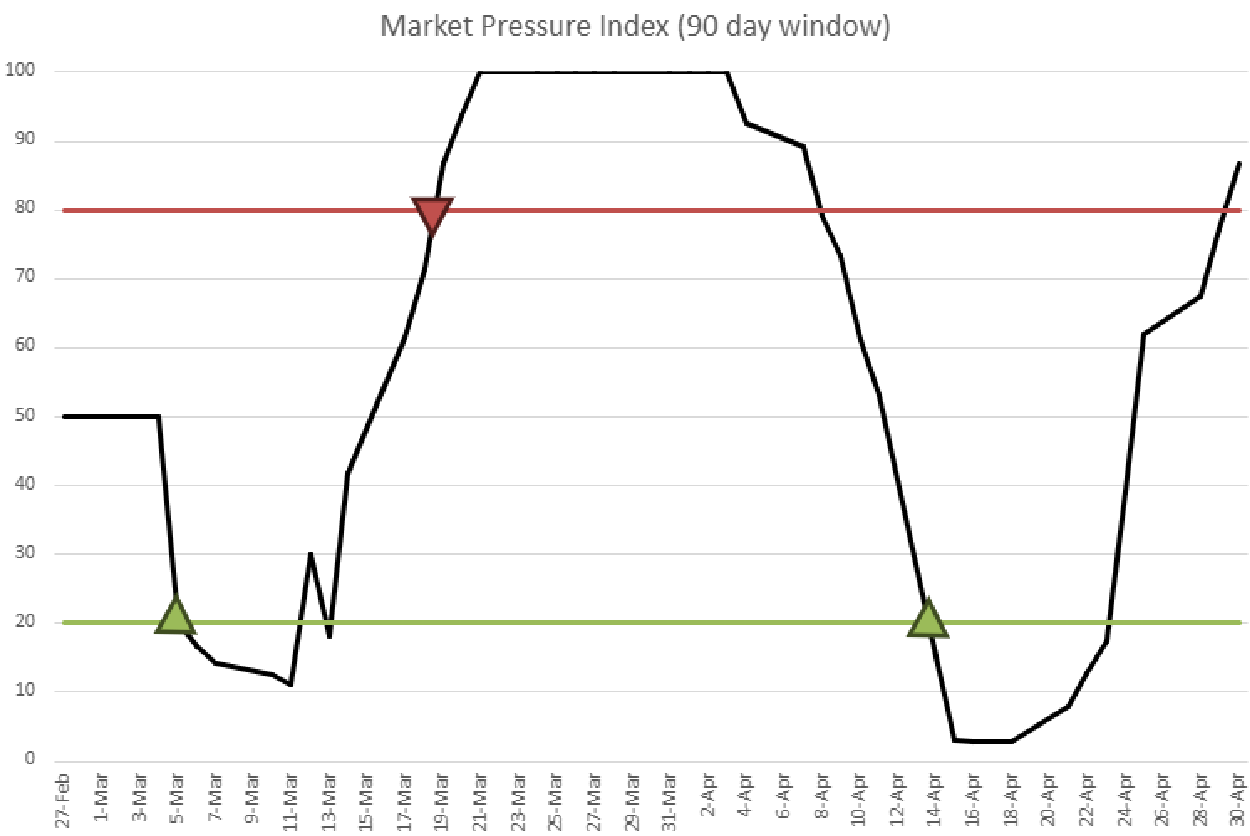


Figure 3: 90-Day Market Pressure Index during April 2025 US Tariff Period, showcasing similar behaviour, with a slightly later response of overextension (Mar 18th).

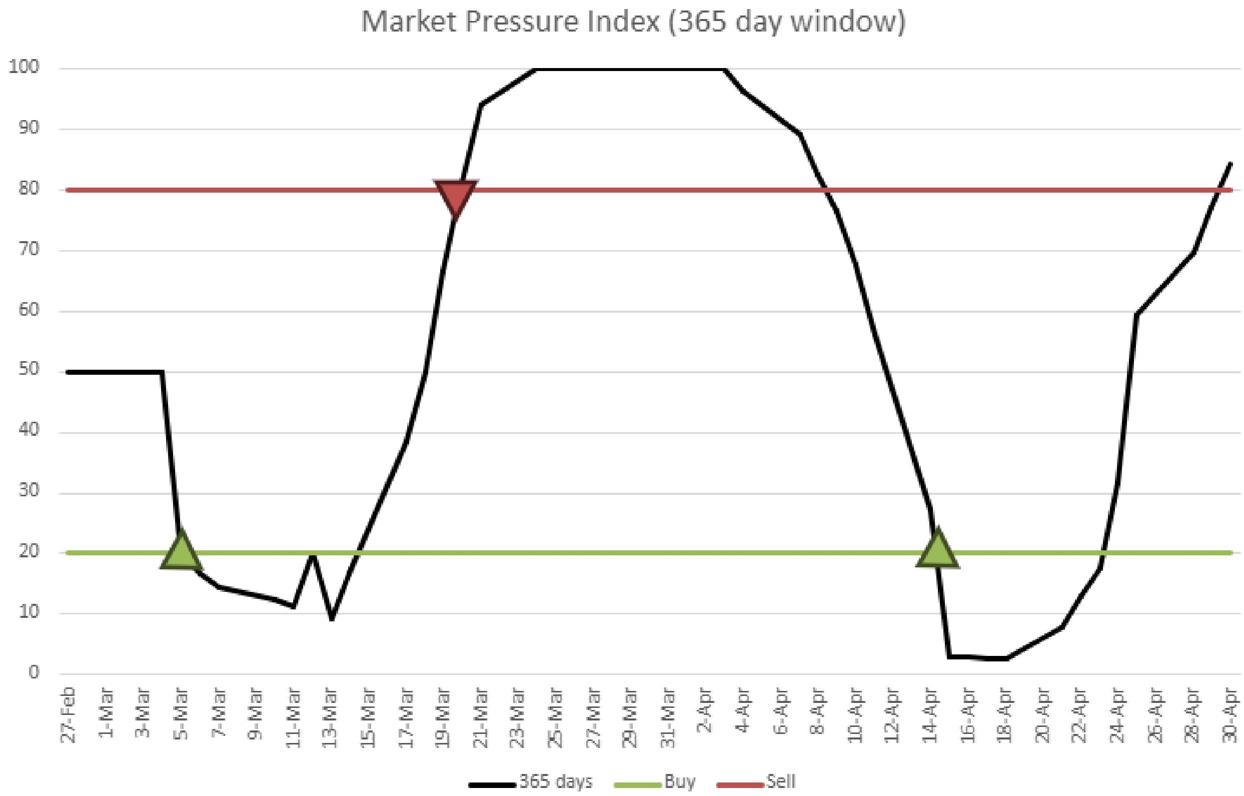


Figure 4: 365-Day Market Pressure Index during April 2025 US Tariff Period, with a further delay until Mar 19th.

4.3 Results

Figures 1-3 represent the realised Market Pressure Indexes during the tariff period and the month preceding. The charts reveals several critical inflection points:

4.3.1 Key Observations

1. **March 5-17: First Warning Signal (MPIs $\rightarrow$ 50 $\rightarrow$ 15)**: The MPIs dropped sharply to the overextended downward zone (MPI < 20), indicating a **buy signal**. This preceded a gradual price rise over the following 10 days, with prices trending from 560 to 590 following the bullish signal generated by our index. Note that we consider from 7th March to respect the T-2 latency.

2. **March 17-30: Sustained Overbought (MPIs > 80)**: The MPIs remained in the positively overextended zone for over 10 days, during which prices continued to gradually trend upwards. This sustained warning period, highlighted across the MPIs, suggested significant overextension, signalling not only to close long positions/tighten the exit initiated on Mar 7th as the index crossed neutrality, but also **trigger new shorts** for the speculative trader utilising the MPIs.

3. **April 1-7: Oversold Extreme (MPIs < 20)**: As prices collapsed, sentiment in the faster moving 7-day MPI flipped to neutrality on April 7th, followed by the 90 and 365d MPIs. This continued quickly to short overextension on April 9th followed by the 90d and 365d indices on April 14th and 15th, respectively. If trading the index systematically in isolation, the return to neutrality would have represented the exit of our short position initiated on March 23rd.

4. **April 10-20: Consolidation (MPIs $\approx$ 20-30)**: The index remained in the lower range, corresponding to a continued buyside signal.

5. **April 22-27: Second Overbought Warning (MPIs $\rightarrow$ 100)**: A surge in optimism drove the MPI back to 100, again signaling excessive bullishness. True to form, this was followed by a price correction in late April.

4.4 Signal Consistency Across Timeframes

A critical validation of the MPI is that the signal remains consistent across different time windows (7-day, 90-day, 365-day). During the April tariff period, we observed the following:

7-day MPI: Provided the most responsive signals, capturing short-term sentiment swings within days of reversals, suggesting strongest promise to shorter term traders seeking rapid signals.

90-day MPI: Exhibited similar patterns but with greater smoothing, filtering out the noisier 7-day signal while still identifying major turning points (e.g. the mid-March peak and early April trough).

365-day MPI: Showed longer-term structural positioning, confirming that even on an annual basis, the tariff period represented an extreme sentiment event.

5 Discretionary Evaluation

Traders can leverage the MPI as a timing tool in several ways, both defensively to protect existing positions and offensively to time reversal entries. On the former, the MPI functions as an early warning system by signaling when the contract is primed for reversal, providing traders and risk managers with an improved snapshot of the market regime, thereby allowing for risk-aware position sizing. The index also informs stop-loss placement beyond a technical support or resistance level, traders may tighten stops on long positions when the MPI exceeds 80 and on short positions when it falls below 20. Aligning MPI extremes with technically important levels would **further bolster its reliability**. In timing systematic entries, as we will demonstrate later in this paper, pairing the MPI with a technical confirmation of trend weakness provides **robust contrarian entry signals**.

For higher-conviction signals, traders should look for alignment across multiple MPI time windows. When the 7-day, 90-day, and 365-day MPIs all exceed 80 at the same time, the signal of an impending reversal is particularly strong, as it reflects extreme sentiment across multiple trading horizons. In contrast, when only the 7-day MPI reaches an extreme level while the longer-term measures remain neutral, the signal is more likely to be temporary and less reliable. That said, for speculative traders who already hold a short or mid-term contrarian view based on an alternative framework, an extreme 7-day MPI can be highly valuable, as it offers insight into the prevailing market psychology and the type of price action they are likely to face in the near term.

An important caveat when leveraging MPI for signal generation is that the expanding window yields slightly different index values depending on the timeframe used. This is partially mitigated by the open interest scaling factor, as widening the interval to significantly far from expiry will result in lower open interest compared to long term maximum levels, naturally dampening index scores toward neutrality. Nonetheless, this can skew later index scores and result in differences in buy/sellside signal timings. Furthermore, whilst the thresholds of 80/20 represent sensible bounds to signal reversals are impending based on forward returns, this is far from an optimised product specific approach.

To rectify this, we now construct a more rigid, rigorous, rule-based approach to utilise these signals with appropriate hyperparameter modifications, in full view of the MPI's strengths and limitations.

6 Systematic Trading

We aim to develop and validate a systematic trading strategy that combines MPI signals with technical reversal indicators to identify mean reversion opportunities in oil swap contracts.

6.1 Data Architecture and Preparation

In constructing this set our sources of data are solely **historical Flux OHLC and Flux COT APIs**.

Step 1: Data Loading and Quality Assessment. The process begins with loading 2 separate datasets, the first containing OHLC price data sources from Flux historical API, and the second, COT fields and metadata with product identifiers, tenors, and processing dates. In the COT dataset, for each date, we calculate the **time-to-settlement (TTS) as the number of months between the processing date and the contract's expiration date**, and include processing dates for TTS 5 to TTS 1. For example, take the Dec'24 MOPJ Crack contract, expiring 2024-12-31. At this stage of the data processing pipeline, the time series will begin from 2024-07-01 and end at 2024-11-31, reflecting TTS 5 through to TTS 1. We then left-join the COT data with OHLC prices, removing any null fields to ensure continuity for all subsequent calculations.

Step 2: Temporal Split Application. Before any feature engineering or technical indicator calculation, the cleaned dataset is segmented into training+validation and test sets based solely on the ProcessingDate field. This early application of temporal splits ensures strict separation between the three datasets and prevents any information leakage across the temporal boundaries. This is the safest possible approach, though it results in identical dataset features regardless of whether the split is performed before or after technical and MPI feature construction, since all indicators only reference historical data.

- **Training + Validation** Jul 31, 2022 – Dec 31, 2024 (29 months)
- **Test:** Jan 1, 2025 – Dec 12, 2025 (11 months)

Step 3: Technical Indicator Generation. Within each temporal split, a suite of technical indicators are calculated independently for each product-tenor combination. The calculation on TTS 5-1 data ensures that when we subsequently filter to TTS 2-1 for actual trading, all technical indicators have stabilised and are not still calibrating. The main indicators implemented are:

ADX (14-period): Measures trend strength via smoothed directional indicators, range 0–100. Declining ADX signals trend exhaustion, which acts as a prerequisite for reversal confirmation.

ATR (14-period): Average True Range for volatility-adjusted stop/target placement, ensuring position sizing scales with market conditions.

Step 4: TTS 2-1 Filtering for Trading Signals. After technical indicators are fully calculated on TTS 5-1 data, we filter each dataset to include only records where TTS is either 1 or 2. This narrower window corresponds to M1 and M2 contracts, where MPI signals are most reliable and actionable, due to increased liquidity, tighter spreads, and more representative data. The MPI signals themselves are pre-calculated and exist only within this TTS 2-1 range in the source data, as this represents the period of maximum signal quality for mean reversion detection.

Step 5: Signal Generation and Backtesting. With fully initialized technical indicators and MPI signals available within the TTS 2-1 window, the system proceeds to generate trading signals according to the specified entry logic (MPI extremes confirmed by ADX reversals within a window) as we will highlight. Backtesting then simulates trade execution with trading at T+2 open prices, product-specific costs to simulate bid/offer, and position sizing based on entry-time ATR values.

6.2 Signal Specification

Signal generation is characterized by a combination of an MPI alert and technical price reversal confirmation. The exact parameters required for a signal are subject to product specific tuning during cross validation, thus we present an example case below.

6.2.1 MPI Extreme Detection

Oversold (long candidate):

$$S_{MPI}^{buy}(t) = \bigwedge_{w \in \{7d, 90d, 365d\}} \left(\mathrm{MPI}_t^{(w)} < \theta_{buy} \right) \tag{11}$$

Overbought (short candidate):

$$S_{MPI}^{sell}(t) = \bigwedge_{w \in \{7d, 90d, 365d\}} \left(\mathrm{MPI}_t^{(w)} > \theta_{sell} \right) \tag{12}$$

where $\theta_{buy} = 20$, $\theta_{sell} = 80$ (validated thresholds). Logical conjunction $\bigwedge$ requires that ALL timeframes are extreme simultaneously.

6.2.2 ADX Decline with Minimum Strength

Upon MPI extreme at t, we search a confirmation lookback window for example $W_{conf}(t) = \{t-4, t-3, t-2, t-1, t, t+1\}$ for ADX decline. This does not present a leakage problem because execution is always done at T+2, given COT data availability, whereas technical indicators may be generated T+1:

$$\text{ADX}_{max} = \max_{i \in W_{conf}} \text{ADX}_i \tag{13}$$

$$\text{ADX}_{min} = \min_{i \in W_{conf}} \text{ADX}_i \tag{14}$$

$$\text{ADX}_{drop} = \text{ADX}_{max} - \text{ADX}_{min} \tag{15}$$

Confirmation requires:

$$S_{ADX}(t) = (\text{ADX}_{drop} \geq 5) \wedge (\text{ADX}_{max} > 40) \tag{16}$$

The second condition (peak > 40) emerged from regime analysis: where drops from weak trends (ADX < 40) tend to produce more spurious signals. This ensures reversals arise from meaningful trends, and not choppy ranges.

6.2.3 Final Signal Logic

Complete confirmation:

$$\text{LONG} = S_{MPI}^{buy}(t) \wedge S_{ADX}(t) \tag{17}$$

$$\text{SHORT} = S_{MPI}^{sell}(t) \wedge S_{ADX}(t) \tag{18}$$

This three-stage hierarchy trades **signal frequency for quality** as each stage filters distinct failure points and sources of spurious signals. MPI excludes neutrally positioned contracts or contracts trading close to current prices, ADX eliminates weak trends, and the lookback window ensure only recent, relevant, actionable signals are retained.

6.3 Position Sizing and Risk Management

We utilise fixed-dollar risk per trade with ATR-based stops ensuring volatility-adjusted sizing, ATR multipliers are subject to optimisation during cross validation.

$$d_{stop} = k_{stop} \cdot \text{ATR}_{entry} \tag{19}$$

$$N = \frac{R}{d_{stop}} = \frac{R}{k_{stop} \cdot \text{ATR}_{entry}} \tag{20}$$

where $R = \$1000$ (risk per trade), N = contracts. The initial stop/target placement is as follows:

Long: $P_{stop} = P_{entry} - k_{stop} \cdot \text{ATR}_{entry}, \quad P_{target} = P_{entry} + k_{target} \cdot \text{ATR}_{entry}$

Short: $P_{stop} = P_{entry} + k_{stop} \cdot \text{ATR}_{entry}, \quad P_{target} = P_{entry} - k_{target} \cdot \text{ATR}_{entry}$

The trailing stop uses ATR_{entry} (not current ATR) to preserve fixed risk. For longs:

$$P_{stop}^{t} = \max\left(P_{stop}^{t-1}, P_{close}^{t} - k_{stop} \cdot \text{ATR}_{entry}\right) \tag{21}$$

For shorts, we use min(). This approach ensures that the stop only tightens in the profit direction and never widens and enables the trading system to protect its downside whilst allowing the contrarian momentum to run, thereby improving performance.

6.4 Trade Execution

The signal generated at t triggers entry at the 9am bar at $t+2$, at which point we implement realistic bid-offer crossing costs. These costs are implemented as product-specific fees deducted at trade entry and exit. Each product has a predefined bid-offer crossing fee f_{txn} in dollars/cents per contract, defined in the `FLAT_PRODUCT_FEES` dictionary, presented in the appendix. When a position is opened at T+2 OPEN price, the entry fee is immediately deducted from available funds, similarly, as the position is closed via stop loss, take profit, or end-of-data, the exit fee is deducted from the realized PnL. The complete transaction cost for a full trade cycle with is tracked as:

$$\text{Total Fees} = 2 \cdot N \cdot f_{txn} \tag{22}$$

6.5 Train-Validation-Test Protocol

6.5.1 Hyperparameter Set

The hyperparameter set during training and validation includes eight key hyperparameters that control signal generation, risk management, and trade execution. The grid is set up coarse, with wide increments between , with parameters selected based on contrarian trading principles.

Signal generation parameters include MPI buy threshold (range: 10-30), MPI sell threshold (range: 70-90), ADX drop threshold - the fall in the technical ADX indicator to suggest the beginning of trend weakness (range: 3-10), maximum ADX level - the threshold for ADX to surpass to suggest we are in a strong trend (range: 35-48), and lookback window (range: 1-5 days). Risk management parameters comprise stop loss as ATR multiple (range: 2.0x-4.0x), take profit as ATR multiple (range: 2.0x-5.0x), and risk per trade in dollars (fixed at $1,000 for comparability). With this range of values, our hyperparameter set encompasses approximately 7000 possible combinations.

6.5.2 Initial Hypothesis Test

We first run a broad hypothesis test utilising our train/validation data across the universe of available products to gauge which present a structural edge. We select 500 parameter sets at random from the 7000 possible sets and backtest across the entire training and validation space, selecting top contracts subject to median, mean Sharpe and proportion of hyperparameter sets with "valid" status, those with at least 10 active trades.

The use of median and mean Sharpe ratios mitigates sensitivity to outliers, while the proportion of valid parameter sets serves as a proxy for structural consistency and tradability across regimes. Importantly, we draw **no statistical inference** regarding true alpha; instead, the procedure functions as a filtering mechanism to reduce the product universe to candidates exhibiting stable, repeatable behaviour prior to further evaluation.

Table 1 showcases the contracts selected from the hypothesis testing procedure

Table 1: Hypothesis Testing Statistics for Top Candidates

No.	Instrument	Median Sharpe	Mean Sharpe	Valid (%)	Earliest Full Datapoint
1	EBOB	0.2832	0.2665	87.6	2022-04-05
2	C3 Conway	0.1650	0.1384	98.6	2022-02-07
3	92	0.1372	0.1343	88.0	2022-05-16
4	EW Gasoline	0.1328	0.0929	94.2	2022-06-10
5	MOPJ	0.1267	0.1247	83.8	2022-02-07
6	0.5 Bgs	0.1123	0.1104	87.4	2022-03-14
7	380 Crack	0.1092	0.0985	99.8	2022-03-18

6.5.3 Cross Validation Framework

To reduce the risk of fitting the MPI-based model to any single dataset, we use a cross-validation (CV) framework designed specifically for time-series data. The combined training and validation dataset is evaluated using TimeSeriesSplit with three chronological folds. Each fold uses an expanding window where the model is trained on all data available up to a given point in time and validated on the immediately following segment. This approach preserves the continuity of the dataset whilst concurrently mitigating look ahead bias. The model performance is athen veraged across the three validation folds, providing a more reliable estimate of generalization than a single split. We intentionally choose relatively large validation windows of at least 150 data points (approximately 5 months) to ensure sufficient trading activity for robust evaluation. After selecting hyperparameters based on cross-validation performance, the model undergoes one stage of refinement, followed by portfolio optimisation across the selected products. Final results are then evaluated on a fully out-of-sample test set that is never used during training or validation.

Rather than exhaustive grid search which would be computationally expensive, we employ random sampling to explore the hyperparameter space efficiently. For initial optimization, we pre-sample 1000 candidate combinations and evaluate each across all products in the training period, ensuring each fold sees the same combination of hyperparameter set. We then select the top 100 performers on the training portion of each fold for validation.

6.5.4 Hyperparameter Performance Evaluation

Consistently strong performance across multiple folds is key to find parameters that generalize well across different market regimes, rather than those only responsive to temporary market states. Our main metric during this procedure is the annualised validation Sharpe ratio, with a secondary focus on percentage drawdown and trading activity. We implement a series of strict rules ex-ante to ensure balanced, representative evaluation of each hyperparameter set across folds.

First, we require at least 2 active tenors (a tenor within which a signal is generated) and 4 trades executed in each 100-day validation period, to ensure that the annualised fold-averaged validation Sharpe is an accurate measure of the strategy. Second, when selecting a hyperparameter set, we insist that the set appear as a top performer in at least 2 of the 3 validation folds, otherwise we reject the product on grounds of parameter instability. Once a hyperparameter set is selected, our final condition is a requirement that the fold averaged annualised validation Sharpe ratio is non-negative.

6.5.5 Local Grid Refinement Around Best Parameters

After identifying promising parameter combinations through random sampling on the training set, we perform targeted refinement on the full train+validation dataset. This refinement uses a local random grid search exploring the parameter space immediately adjacent to the best training parameters.

The refinement process generates a compact grid centered on the candidate parameters for each product. For example, if the training phase identified an MPI buy threshold of 20, the refinement tests values of 15, 20 and 25. This local exploration balances finding better parameter values if they exist nearby, while limiting overfitting risk that comes from exhaustive searching.

The refined candidates are evaluated on the combined train+validation dataset. We select the parameter combination achieving the highest Sharpe ratio on this combined period, representing the best balance between optimization and generalization. If refinement yields no improvement over the original training-selected parameters, we retain the original optimal CV parameters.

6.6 Trading Across Tenors

We implement a daily mark-to-market (MTM), recalculating portfolio equity on every calendar day as the sum of three components: initial capital, cumulative realized profit/loss from all closed trades, and current unrealized profit/loss from all open positions.

$$\text{Equity}_t = \text{InitialCapital} + \sum_{\text{closed}} \text{P\&L}_i + \sum_{\text{open}} \text{UnrealizedP\&L}_{i,t}$$

For each product-tenor combination, the trading simulation calculates unrealized PnL daily based on current position (if any), entry price, current market price, and position size in contracts. These unrealized PnL values are aggregated across all active tenors to produce total portfolio unrealized PnL for each date.

This daily MTM calculation enables accurate Sharpe ratio computation using actual daily return series rather than artifacts of trade timing. Returns are calculated as percentage changes in daily equity, producing a time series that properly reflects volatility and risk experienced by the portfolio. The resulting Sharpe ratios provide reliable risk-adjusted performance measures suitable for comparing strategies and making allocation decisions.

Thus for each of our products, we calculate the final total equity across tenors over the entirety of the training and validation set with optimal refined CV hyperparameters.

6.7 Portfolio Construction and Product Selection

Once the individual product backtests and performance metrics have been determined, we construct a diversified portfolio of products that collectively aim to deliver attractive risk-adjusted returns.

Portfolio weights are computed using a mean–variance framework (MVO) based on the maximum Sharpe ratio (tangency portfolio). Expected returns are estimated as the sample mean of historical daily returns obtained from the training and validation backtest periods. Portfolio risk is captured by the sample covariance matrix of these returns.

Let $\boldsymbol{\mu} \in \mathbb{R}^N$ denote the vector of expected returns for N products, and let $\boldsymbol{\Sigma} \in \mathbb{R}^{N\times N}$ denote the corresponding return covariance matrix. The initial portfolio weights are obtained in standard fashion, combining expected returns with the inverse of the covariance matrix, corresponding to the classical tangency portfolio. In the code, the Moore–Penrose pseudo-inverse $\boldsymbol{\Sigma}^+$ is used in place of the standard matrix inverse in case of matrix singularity.

$$\mathbf{w} = \frac{\max\big(\boldsymbol{\Sigma}^+\boldsymbol{\mu},\, 0\big)}{\mathbf{1}^\top \max\big(\boldsymbol{\Sigma}^+\boldsymbol{\mu},\, 0\big)} \tag{23}$$

Here, $\mathbf{w} \in \mathbb{R}^N$ denotes the vector of portfolio weights, $\mathbf{1}$ is a vector of ones, and the $\max(\cdot, 0)$ operator is applied element-wise to enforce a long-only constraint by eliminating negative weights. The denominator ensures that the portfolio is fully invested, such that the weights sum to one.

To limit portfolio concentration, an upper bound $w_{\max}$ of 30% is applied to each separate product. If any product allocation exceeds this limit, the excess weight is truncated and redistributed among the remaining uncapped products.

The final portfolio satisfies the following constraints:

$$w_i \geq 0, \quad w_i \leq w_{\max}, \quad \sum_{i=1}^{N} w_i = 1,$$

which results in a fully invested, portfolio with bounded position sizes. In our implementation, no product achieved a weighting exceeding 30%, resulting in no enforcement of the $w_{\max}$ threshold.

6.8 Out-of-Sample Testing and Final Validation

After completing all optimization and refinement on train and validation data, final parameters are evaluated on the completely held-out test period. This provides an unbiased estimate of strategy performance in genuine forward conditions where no parameter tuning has occurred.

Test period evaluation uses identical methodology as training/validation: daily MTM equity calculation, realistic bid/offer costs, and consistent signal generation rules. The test period represents approximately 12 months of trading data, as the entirety of 2025, providing sufficient duration and potential for trade execution to assess performance while remaining recent enough to reflect current market conditions.

6.9 Summary of Complete Protocol

Our complete methodology follows this sequential workflow:

1. **Data Preparation**: Load OHLC price data for all products and tenors, calculate technical indicators (ATR, ADX, MPI), and enforce strict temporal data splits (trainval/test) to prevent leakage.
2. **Initial Hyperparameter Search**: Perform random sampling of hyperparameter space, evaluate candidates on training data across all products, and identify top-performing parameter combinations based on portfolio Sharpe ratios.
3. **Validation Assessment**: Evaluate top training candidates on validation data using fixed parameters (no reoptimization), select best-performing combination based on validation Sharpe ratio, and store results for subsequent refinement.
4. **Local Refinement**: Generate ± 1 step parameter variants around best training parameters, test refinement candidates on combined train+validation data using daily MTM equity tracking, and select final parameters achieving highest Sharpe on full train+validation period.
5. **Full Train+Validation Backtest**: Run comprehensive backtest using final parameters on entire train+validation period, compute performance metrics including Sharpe ratio from daily returns, and generate equity curve showing portfolio growth and drawdowns.
6. **Product Selection and Portfolio Construction**: Evaluate individual product performance across validation folds, select products showing consistent favourable Sharpe ratios, calculate mean-variance optimal weights with position limits, and construct final portfolio allocations.
7. **Out-of-Sample Test**: Apply final parameters and portfolio weights to completely held-out test period, calculate all performance metrics using identical methodology, compare test vs validation performance to assess robustness, and document any regime-dependent behavior observed.

7 Results

We now discuss the results formulated from the approach outlined above, with separate treatment for each of the training/validation and out of sample datasets. The portfolio allocation below showcases the relative contributions from each of the instruments passing hypothesis testing and cross validation, determined from MVO.

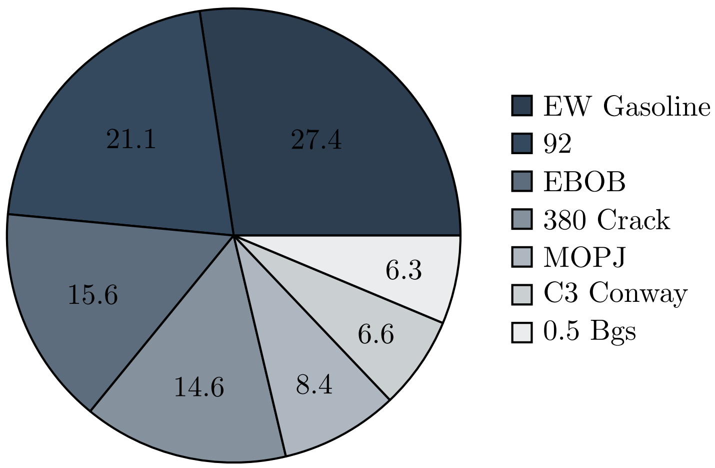


Figure 5: Contribution to portfolio from each instrument

7.1 Training Performance

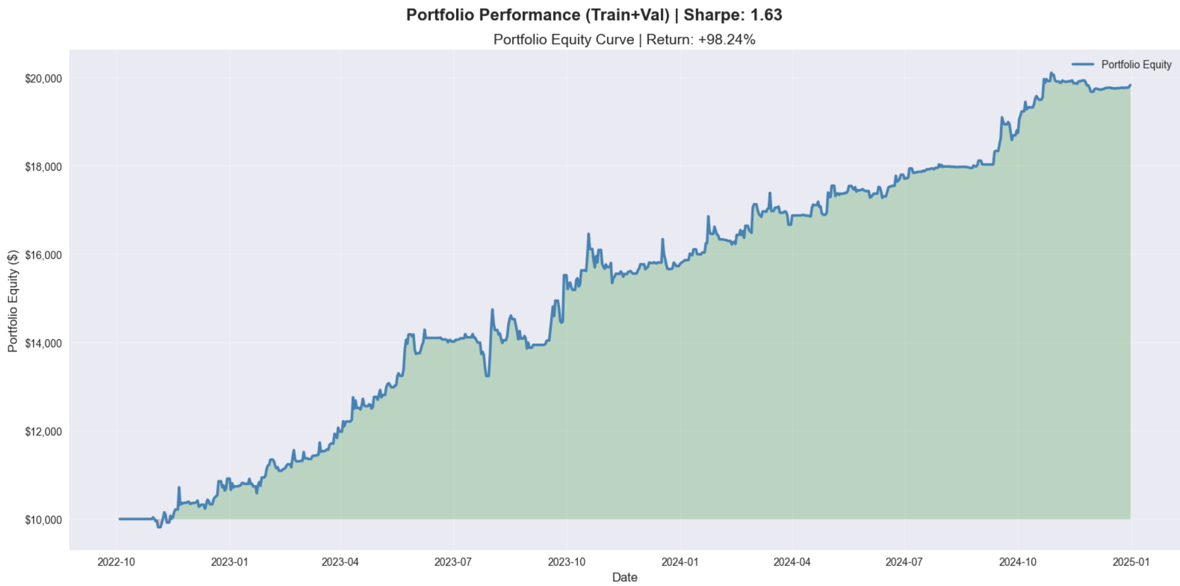


Figure 6: Training+Validation Portfolio performance following cross validation hyperparameter tuning, refinement and portfolio optimisation.

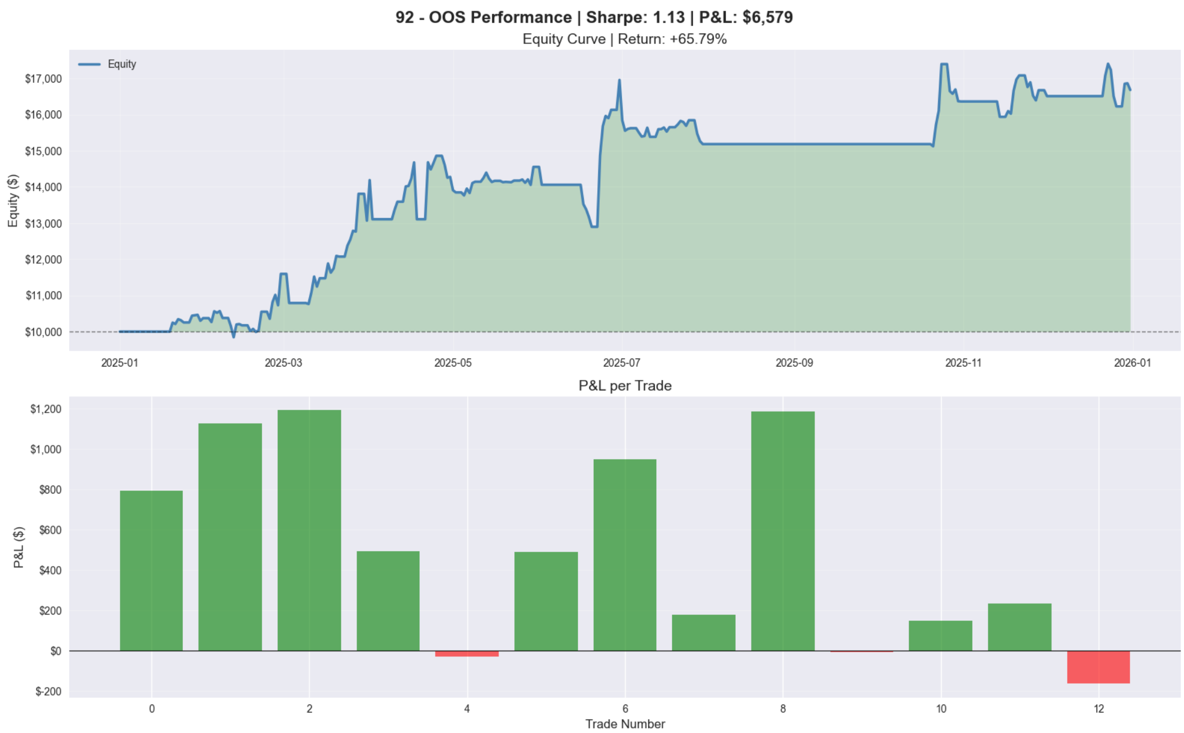
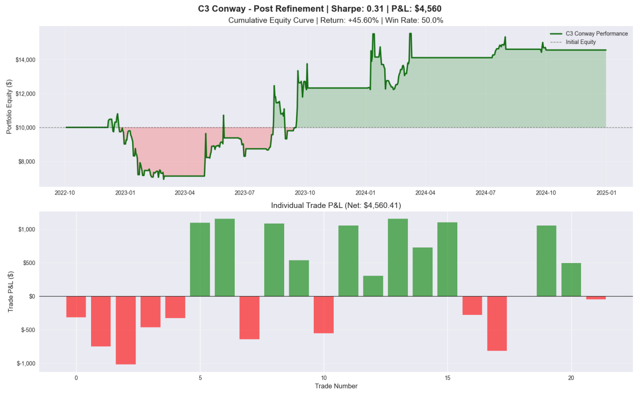
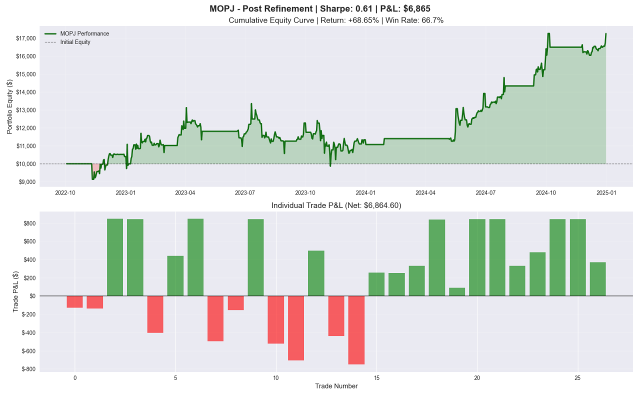


Figure 7: Strongest in-sample product, Singapore 92 performance in training+validation set, with trading frequency, winrate and Sharpe.

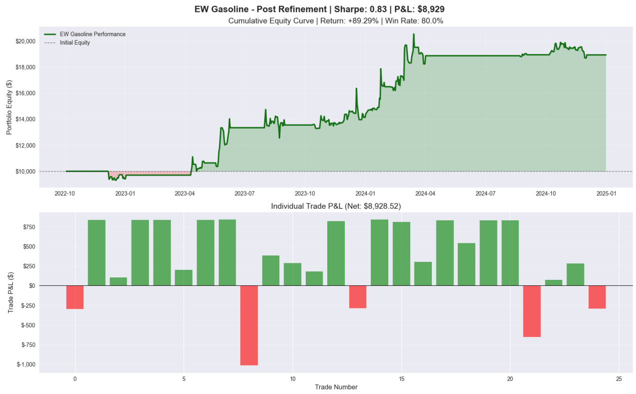
As figures 2 and 3 highlight, backtest results utilising refined CV parameters demonstrate strong performance - the slight drawdown earlier in the year is mitigated by strong positive returns thereafter, with promising results especially in gasoline contracts, 92 and EBOB, with Sharpe ratios at 1.13 and 0.91 respectively, while EBOB showcases the strongest raw performance over the 2.5. Diversification boosts the portfolio Sharpe to 1.63 in-sample.



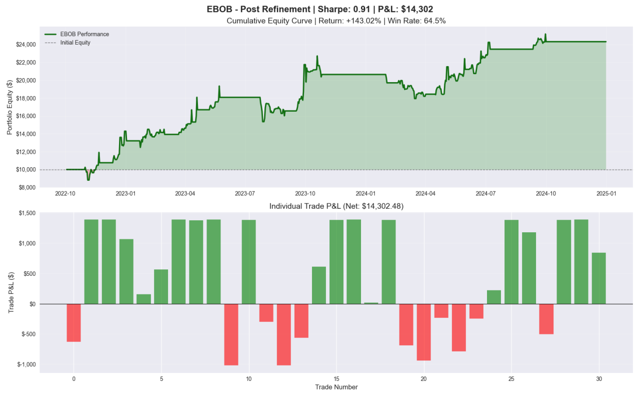
(a) C3 Conway Propane full training set performance



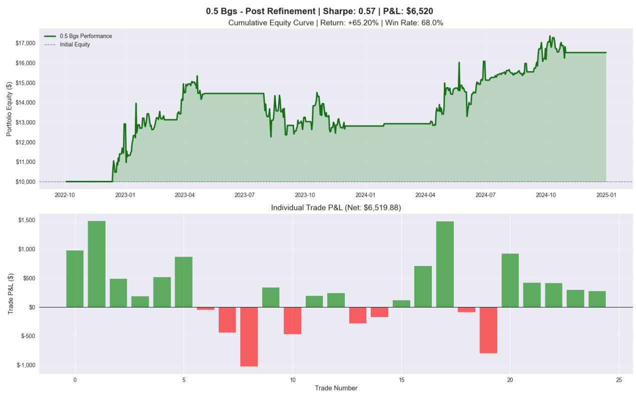
(b) MOPJ full training set performance



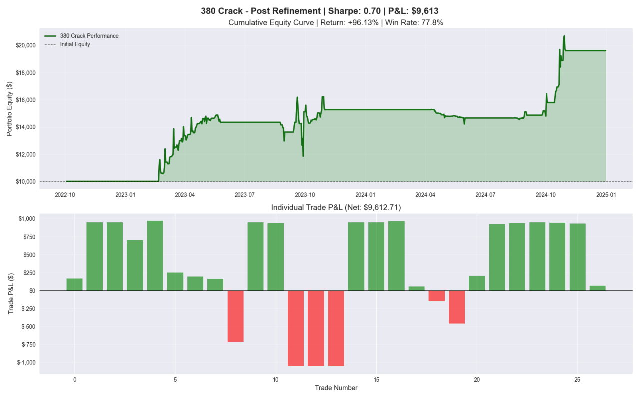
(c) East/West Gasoline full training set performance



(d) EBOB full training set performance



(e) 0.5% Barges full training set performance



(f) Singapore 380 Crack full training set performance

Figure 8: Refined parameter set following CV across training+validation periods

7.2 Out of Sample Performance

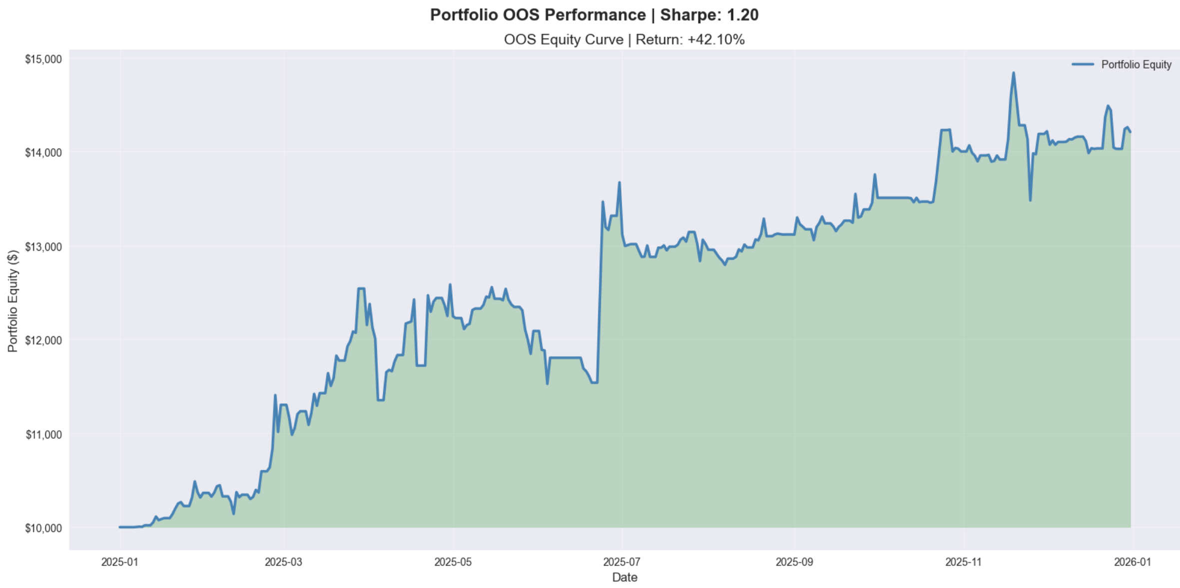


Figure 9: Out of sample (year of 2025) portfolio performance utilising refined CV params from training+validation.

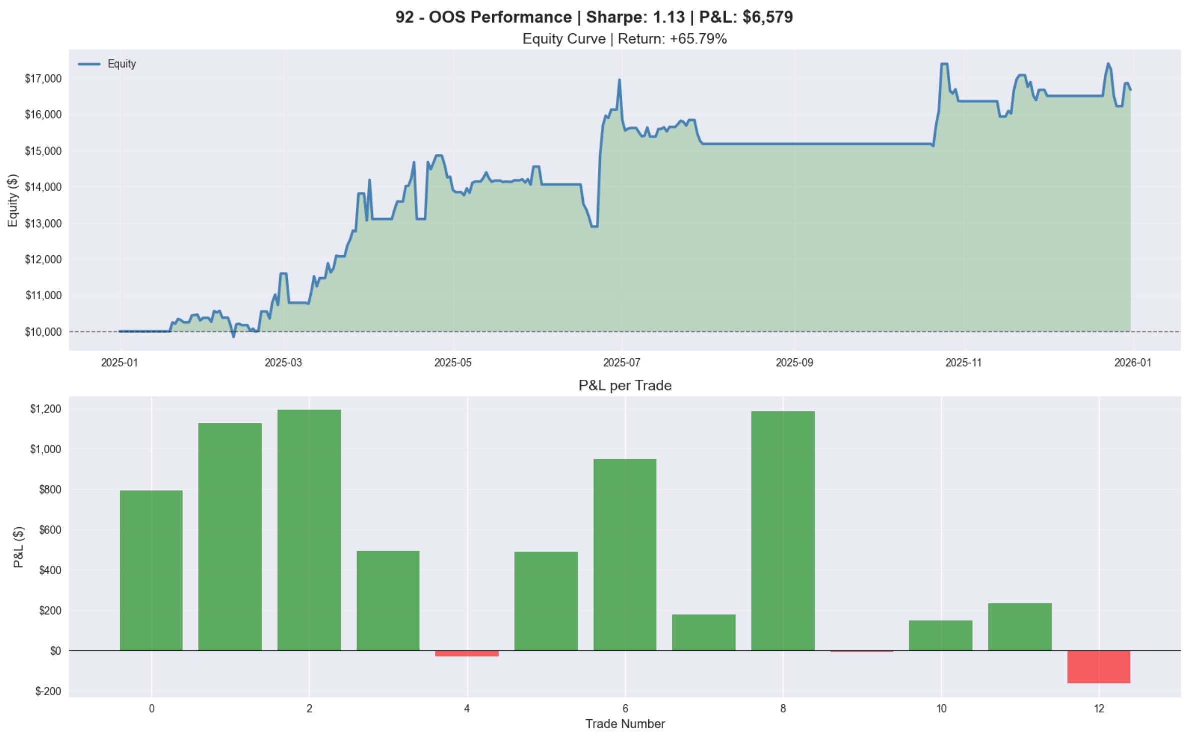
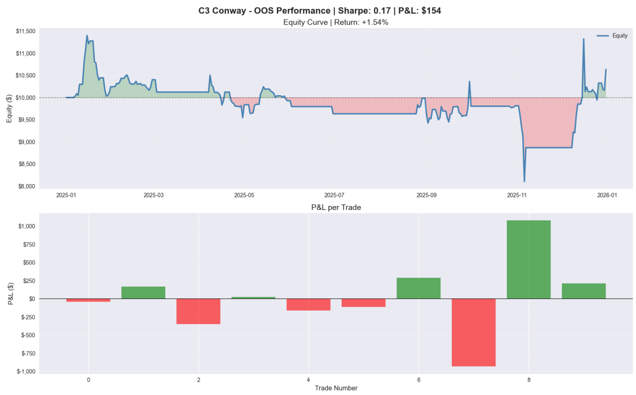
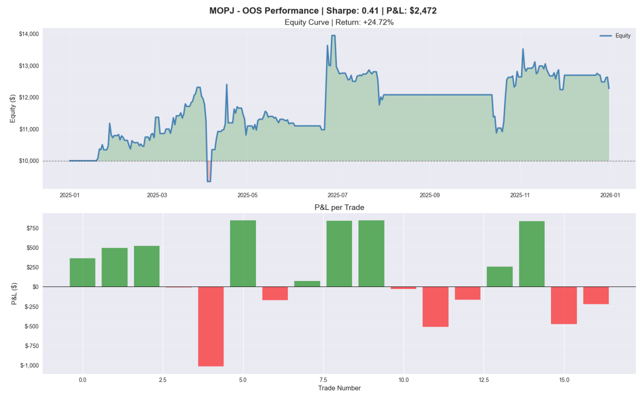


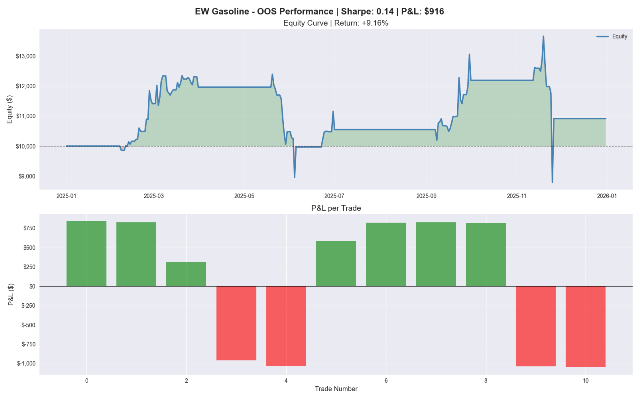
Figure 10: Highest in sample performer, Singapore 92, demonstrates strong performance out of sample, indicating good generalisation amid a favourable regime.



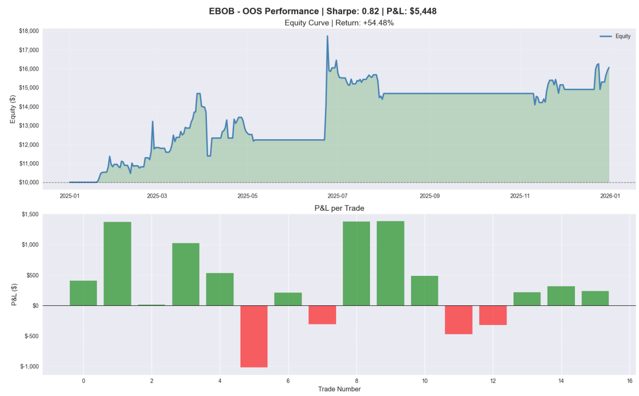
(a) C3 Conway Propane full test set performance



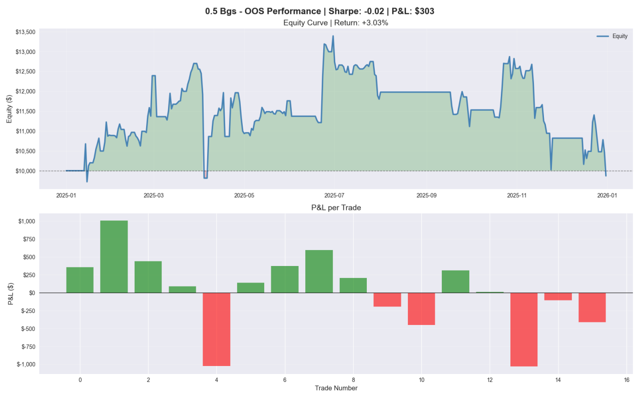
(b) MOPJ full test set performance



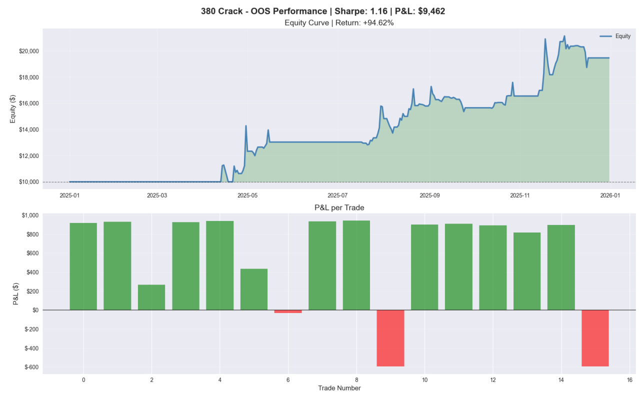
(c) East/West Gasoline full training set performance



(d) EBOB full test set performance



(e) 0.5% Barges full test set performance



(f) Singapore 380 Crack full test set performance

Figure 11: Refined parameter set following CV across training+validation periods

Out-of-sample results showcase good generalisation at a portfolio level with slight Sharpe decay from 1.63 in-sample to 1.2 in the testing set. This reduction is well within the bounds of expected performance and confirms that our cross-validation approach was largely successful in mitigating overfitting risks. However, at product specific granularity, it becomes clear that performance stability was not universal. C3 Conway and East/West Gaosline saw notable performance degradation, suggesting some overfitting in sample and poor generalisation. We note also that in the case of the East/West, the portfolio trades the individual flat prices directly, and retains good performance OOS, suggesting that this contract may be a lower quality signal of the same underlying behaviour in each of its legs.

8 Conclusion and Future Work

We developed a systematic mean-reversion strategy combining multi-timeframe MPI with an enhanced three-stage confirmation for contracts with strong liquidity, and validated it through a rigorous train-validation-test protocol. We verify that this approach is capable of generating high quality alpha signals, with clear, robust edge across regimes.

The MPI signal at present aims to capture the level of over or underpositioning pressure across products with good accuracy. However, there is strong potential to improve the underlying model, and we aim to construct an improved MPI with enhanced elements in future projects. First, in the current approach, we implement adjustments for open interest and positioning as secondary steps in MPI construction, where these could be incorporated into the original price deviation, weighting long or short entry levels by their corresponding level of positioning when building the net deviation. Furthermore, positioning at more varied frequencies could be investigated instead of one overarching net positioning metric. This was originally implemented as a way of maximising the positioning-related information available to each index, and allows for straightforward comparison to the open interest, but aligning the positioning scaling factor or weighted long/short deviation with the same timeframe, for example blotteredmarketnet7d with the 7d rolling entry price, may produce a more representative signal.

Regime analysis on training and validation data suggested minimum ATR thresholds or vol-regime classification (which would permit trading only in medium/high vol) was a key factor in signal accuracy. This alignes with intuition around contrarian trading, in that rangebound, low volatility price action is highly difficult to trade, with the total absence of a meaningful trend to attempt to fade. The Max ADX threshold incorporates this insight, but more could be done to analyse regime classification and/or confirm technical trend exhaustion as part of the systematic signal generation procedure.

From a portfolio construction and trade-structuring perspective, a key constraint on growth of portfolio equity is the use of a fixed dollar risk per trade. While this approach was appropriate for backtesting and performance evaluation, to gauge the reliability of MPI, live trading and production signal deployment would require a more robust and dynamic bet-sizing framework.

A ADX Details

ADX (Average Directional Index, Wilder 1978) measures trend strength via directional indicators:

Plus Directional Indicator (+DI):

$$+DI_t = 100 \times \frac{\mathrm{EMA}_{14}(+DM_t)}{\mathrm{ATR}_t} \tag{24}$$

where $+DM_t = \max(H_t - H_{t-1}, 0)$ if $H_t - H_{t-1} > L_{t-1} - L_t$, else 0.

Minus Directional Indicator (-DI):

$$-DI_t = 100 \times \frac{\mathrm{EMA}_{14}(-DM_t)}{\mathrm{ATR}_t} \tag{25}$$

where $-DM_t = \max(L_{t-1} - L_t, 0)$ if $L_{t-1} - L_t > H_t - H_{t-1}$, else 0.

Directional Index (DX) and ADX:

$$DX_t = 100 \times \frac{|+DI_t - (-DI_t)|}{+DI_t + (-DI_t)} \tag{26}$$

$$ADX_t = \mathrm{EMA}_{14}(DX_t) \tag{27}$$

ADX > 25 indicates trending, ADX < 25 suggests ranging. Declining ADX signals trend exhaustion. System detects reversals: +DI crossing above -DI (bullish) or vice versa (bearish) during falling ADX.

B Product Universe Detail

Complex	Products (N)
Gasoline (7)	92, 92 Crack, Arb, EBOB, EBOB Crack, EW Gasoline, Gasnaph
Naphtha (5)	MOPJ, MOPJ Crack, EW Naphtha, Naphtha, Naphtha Crack
NGLs (14)	C3 Conway, C3 CP, C3 ENT, C3 EW, C3 FEI, C3 LST, C3 NWE, C3/C4 CP, C4 CP, C4 ENT, C4 EW, FEI/CP, FEI/MOPJ, Pronap
Fuel (7)	0.5 Bgs, 0.5 Bgs Crack, 3.5 Bgs Crack, 3.5 Bgs, Sing 380 Crack, Sing 380, 380 E/W
Crude (1)	Dated/Brent,

Table 2: Complete product universe spanning 3 complexes, 26 products total.

C Bid-Offer Modelling

The dictionary below highlights the fee imposed during entry and exit of a trade, intended to simulate bid/offer.

```
# Product-specific bid-offer crossing fees, in units traded on Flux Terminal, eg c/gal, $/bbl,
PRODUCT_FEES = {
    'Gasoline': {
        '92': 0.06,
        '92 Crack': 0.05,
        'Arb': 0.125,
        'EBOB Crack': 0.05,
        'EBOB': 1.0,
        'EW Gasoline': 0.5,
        'Gasnaph': 0.5
    },
    'Naphtha': {
        'MOPJ': 1.0,
        'MOPJ Crack': 0.05,
        'EW Naphtha': 0.5,
        'Naphtha': 1.0,
        'Naphtha Crack': 0.05
    },
    'NGLs': {
        'C3 Conway': 1.5,
        'C3 CP': 3.0,
        'C3 ENT': 1.5,
        'C3 EW': 3.0,
        'C3 FEI': 3.0,
        'C3 LST': 1.5,
        'C3 NWE': 3.0,
        'C3/C4 CP': 1.5,
        'C4 CP': 3.0,
        'C4 ENT': 1.5,
        'FEI/CP': 1.5,
        'FEI/MOPJ': 1.5,
        'LST/FEI': 1.5,
        'Pronap': 1.5
    }
}
```

Acknowledgments

Special thanks to Frederick Davy on the Onyx Quant-Desk for his valuable review of the systematic approach and insightful feedback throughout the development of the MPI. Thanks also to the Flux Insights Research team, particularly Mita Chaturvedi, for her assistance in market research for and data gathering for the case studies, which significantly strengthened the practical dimension of this project.