An All-Weather, Multi-Strategy Approach to Pairs Trading

Abstract—We develop 3 pairs trading strategies, each of which performs well under a different set of market conditions: 1) a Kalman filter-adjusted mean reversion strategy that identifies opportunities in horizontal price movements, 2) a copula-based anomaly detection strategy that tracks return relationships particularly well under market stress conditions, and 3) a factor residual-driven Machine Learning model. Together, the strategies complement each other to form an all-weather statistical arbitrage umbrella strategy. We implement a Three-State Variance Switching Model to orchestrate the allocation to our sub-strategies. The strategy achieves a Sharpe ratio of 0.95 and Sortino ratio of 1.4 during our out-of-sample period for the pair comprising of S&P 500 Utilities and S&P 500 Electric Utilities.

1 Introduction

The origins of pairs trading and statistical arbitrage can be traced back to Morgan Stanley in the 1980s. As pairs trading gained popularity and the strategy became increasingly crowded, arbitraging between securities with obvious co-movements and sound economic grounding became less-and-less profitable. With the low-hanging fruit gone, arbitragers have been pushed towards: (1) higher leverage on the pairs, (2) shifting from an economically-based approach to one that is purely statistical, and (3) overlaying directional bets to earn risk premia on market factors. Market participants have been on a quest for *de novo* strategy innovations ever since.

Consequently, a multitude of approaches have been developed to capitalize on markets' mispricings. These can be categorized into a variety of sub-strategies including: mean-reverting, momentum, regime shifting, seasonal trending, and high-frequency trading, among others (Chan, 2008, ref[12]).

In theory, many of these strategies are market-neutral, making them ideal sources of alpha for hedge funds and proprietary trading desks alike. Yet, while they may generate apparently market-neutral returns under stable conditions, when facing market stress, simultaneous unwinding of crowded convergence trades creates a game of musical chairs for strategy participants: the last one to unwind their positions suffers the greatest slippage towards divergence.

The most prominent statistical arbitrage sub-strategy is mean-reversion, which involves examining the relative price movements of two securities and creating trading rules to exploit the empirically-observed price patterns. The strategy operates by observing the co-movements of two securities over time. When it identifies pricing anomalies (i.e., divergence), it enters a long position on the underpriced and goes short the overpriced. Theoretically, when these pairs revert back to their historical relationship (i.e., convergence) traders unwind the position and pocket the profit.

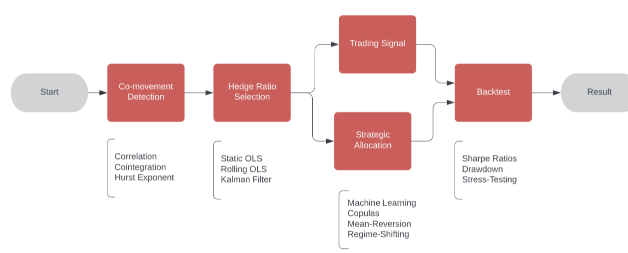

Fig. 1. Flowchart demonstrating trading strategy pipeline from data to return results.

The goal of our paper is to use non-linear methods to make improvements to traditional pairs arbitrage. In particular, we have limited our investment universe to a pair of two equity indices. A combination of linear, copula, and machine learning teachniques is used to both select a pair and generate trading signals. Fig[1] shows a schematic for our strategy beginning with daily index data and resulting in a returns profile together with performance metrics.

2 Traditional Linear Methods

Before exploring non-linear methods, we consider a baseline (linear) distance approach. We begin by identifying two co-integrated stock indices exhibiting mean-reverting behavior from which we intend to generate profit. A linear combination of the returns of such a pair is stationary as follows:

$$\epsilon = R_1 + \gamma R_2,$$

where ϵ is the spread of the returns and γ is a nonzero co-integration factor (Krauss, 2016, ref[10]). Once we have identified a co-integrated pair we assume that the return behaviors of our two assets are linearly dependent under normal conditions.

We generate trading signals from two Bollinger bands, created by adding or subtracting 2 rolling standard deviations of the spread time series. If the spread breaks the upper band, we know that the relative value of the first asset to the second is too high and vice versa. At this point we initiate a 1-to-1 long/short trade to earn a profit from mean reversion.

This baseline strategy applies a combination of the co-integration methodology and the distance approach which are often found in basic linear statistical arbitrage methods. We utilize Granger's causality test to determine whether one time series has information about the future value of another time series (Granger, 2004, ref[2]). We conduct our hypothesis test for all 56 shortlisted pairs and find that the p-values are, largely, close to 0. Thus, we reject the null hypothesis and consider the alternative hypothesis that one time series does Granger-cause the other. With these results, we can conclude that the two indices are co-integrated.

With a co-integrated pair, Fig[2] shows that the baseline mean-reverting strategy performs poorly when the underlying spread is trending. The drift, independent of the rolling mean reversion eats into the mean-reversion profits and produces a significantly larger variance. From this observation, we can infer that the strategy will not be effective without an efficient hedging ratio to offset the drift.

Fig. 2. Return of mean reverting strategy for the pair SP500-5510 & SP500-551010 with Hedge Ratio = 1

3 Pair Selection

The most basic pair selection method uses Pearson correlation to measure the linear relationship between two returns' time series and thereby rank/select pairs. However, using Pearson correlation does not account for non-linear monotonic relationships. In addition, it is easy to overlook the fact that the spread between the returns can be non-stationary. Yet, stationarity is crucial for a mean reversion pair trading strategy to be effective.

To account for non-linear association we use Spearman's rank correlation and Kendall's Tau to measure non-linear monotonic relationships and subsequently select pairs. To test for stationarity of the spread, the Hurst exponent method is used. Anti-persistence is commonly associated with mean-reversion so selecting pairs with Hurst exponent of less than 0.5 allows us to apply mean-reversion strategies to the pairs (Ramos-Requena et al., 2020, ref[8]). The Hurst exponent is also suited to model the relationship between two indices due to its dependence on the power law, which often describes histograms of stock price fluctuations (Gabaix et al., 2015, ref[7]). We therefore hypothesize that the power law will model index prices well and find that the Hurst exponent provides valuable insight into the mean-reverting relationship of an index pair.

The Hurst Exponent H is defined by:

$$\mathbf{E}\left[\frac{R(n)}{S(n)}\right] \sim Cn^H \text{ as } n \to \infty$$

where:

- $R(n)$ is the rescaled time series
- $S(n)$ is the series (sum) of the first n standard deviations
- $\mathbf{E}[x]$ is the expected value
- n is the time span of the observation (number of data points in the sample)
- C is a constant

The Hurst exponent ranges between 0 and 1. Based on the value of H, we classify time series as follows:

- $H < 0.5$ - mean reverting, anti-persistent
- $H = 0.5$ - a geometric random walk
- $H > 0.5$ - a trending persistent series

We compute H and follow standard hypothesis testing procedures as we choose to conduct a left tailed hypothesis test for H to verify that it is significantly less than 0.5.

First C_2^n unique index pairs are generated from a universe of 131 indices. Next, Spearman's rank correlation coefficients and Kendall's tau values are computed for each unique index pair and a threshold of 0.8 is used to filter index pairs. Then the Hurst exponent criteria is applied with threshold set to 0.5. Finally, pairs that satisfy Hurst exponent criteria are considered as the universe of index pairs in our study.

This filtration results in a universe of 56 equity index pairs constituting 50 unique indices. Then the optimal probability distribution that best fits each of the index returns is determined based on Akaike Information Criterion (AIC), the p-value of the Kolmogorov-Smirnoff test and the p-value of the Cramèr-von Mises Criterion. Next, the optimal copula for each pair is selected based on Akaike Information Criterion (AIC), and the p-value of the Kolmogorov-Smirnoff test for bivariate distribution.

For each index from the list of indices, the optimal marginal distribution is chosen from: (1) Normal, (2) Student's t, (3) Logistic and (4) Extreme. These distributions perform the best on financial data as they have features of a normal distribution while allowing for fatter tails.

To determine the optimal copula for each pair, 7 copulas from the families of Elliptical and Archimedean copulas: (1) Gaussian, (2) Clayton, (3) Gumbel, (4) Frank, (5) Joe, (6) N13 and (7) N14 are considered. Each pair is fit to each copula and parameters are estimated using maximum likelihood estimation (Stander et al., 2013, ref[14]). Then, the optimal copula is chosen based on AIC, and the p-value of the Kolmogorov-Smirnoff test for each bivariate distribution.

		Pearson	Spearman	Kendall's Tau	Hurst Exponent	P-Value
DJU	SP500-551010	0.978	0.966	0.855	0.296	2.16e-5
SP500-5510	SP500-551010	0.985	0.976	0.8825	0.292	1.04e-4

Fig. 3. Statistics for our selected pairs

Fig[3] shows two pairs we selected from the index universe that we will be considering for the rest of the study.

	Distribution	AIC	KS-Test P-Value	Cr-Test P-Value
SP500-5510	Student's t	-25069.19	0.131	0.143
SP500-551010	Student's t	-24853.43	0.118	0.255
DJU	Student's t	-25033.24	0.151	0.130

Fig. 4. Optimal probability distributions for selected indexes

		Copula	Copula Parameter	AIC	KS-Test P-Value
SP500-5510	SP500-551010	N14	7.952	-13636.93	0.177
DJU	SP500-551010	N14	6.372	-12080.57	0.069

Fig. 5. Optimal copula and statistics for selected pairs

4 Dynamic Hedging

Statistical arbitrage is the systematic exploitation of perceived mispricings of similar assets.

Instead of using a traditional correlation approach, we have used a combination of Spearman's rank correlation coefficient, Kendall's tau and the Hurst exponent as metrics to determine the similarity.

Once we confirm that there is a comovement, we create a time series of the spread between the two indices of a pair while considering β. The spread takes the form:

$$Spread = Y - \beta X$$

where:

- X, Y are the daily returns for each index
- β is the number of units of index X we want to buy or sell

Though stationarity exists within static beta-defining processes, the spreads revert to the long-term mean every 3-10 years for our index pairs. This prolonged directional drift is unprofitable within an annual trading period. Searching for a dynamic hedging ratio that captures the relationship of two indices to form a mean reverting spread series becomes essential for a statistical arbitrage strategy.

To execute dynamic hedge ratios we produce a time series of β to form a short-term mean reverting series. We implement a Kalman filter for our dynamic beta assumptions, as it resembles the market aggregation of return beliefs.

The Kalman filter is a Bayesian inference algorithm that produces an estimate of one or more unknown variables based on noisy measurements. The Kalman filter allows several inaccurate measurements to be used to estimate a parameter better than any individual measurement (Xu, 2017, ref[13]). Mathematically, it can be represented by the following equations:

Prediction stage:

$$\hat{x}_k^- = A\hat{x}_{k-1} + B\mu_k \tag{1}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{2}$$

Update stage:

$$k_k = P_k^- C^T (CP_k^- C^T + R)^{-1} \tag{3}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - C\hat{x}_k^-) \tag{4}$$

$$P_k^- = (I - K_k C)P_k^- \tag{5}$$

where:

- $\hat{x}_{k-1}$ is previous best estimate of state variables
- A is the state transition matrix
- $\hat{x}_k^-$ is the prior estimate of state variables
- $B\mu_k$ is offset term for the transition.
- P_k^- and P_{k-1} represent the prior estimation of the current covariance matrix and previous covariance matrix respectively
- Q is the transition error
- k_k is the Kalman gain
- C is the observation matrix
- R is the observation error

We implement Kalman filters on the time series returns to estimate a dynamic beta. We adjust the sensitivity of the Kalman filter to fit a beta between 0 and 2 such that it fits the theoretical beta between two equity indexes and that net exposure fluctuates within the +1 and -1 range so no leverage is applied within the initial trading signal. After applying dynamic hedge ratios the spread series then becomes mean reverting within a 60-day time frame and would make mean reversion models viable for signal generation.

5 Signal Creation

In this paper, we have constructed a multi-strategy framework that ultimately integrates an ensemble of 3 sub-strategies into an all-weather statistical arbitrage strategy. The sub-strategies built on 3 different measures excel in capturing different movements of the markets: (1) the naive mean reversion strategy captures the fluctuations of spread under stable market conditions; (2) the copula models return probability distributions between two indices under market stress; and finally, (3) the factor-residual based machine learning model exploits non-systematic price deviations.Then we apply a regime-change filter to decompose market behaviors and navigate our strategic selection process to complete our multi-strategy framework.

Fig[6] shows the trading period, profit conditions, and strategic value which will be considered in the allocations of the multi-strategy.

	Copulas Anomaly Detection	Bollinger Band	Factor Machine Learning
Trading Period	Short-Term (1 day)	Mid-Term (60 days)	Long-Term (180-250 days)
Profit Conditions	Tail Events	Non-directional Moves	Directional Movements
Strategic Value	Crisis Events	Baseline Strategy	Return Enhancement

Fig. 6. Matrix of the trading period, profit conditions, and strategic value for each strategy in the framework

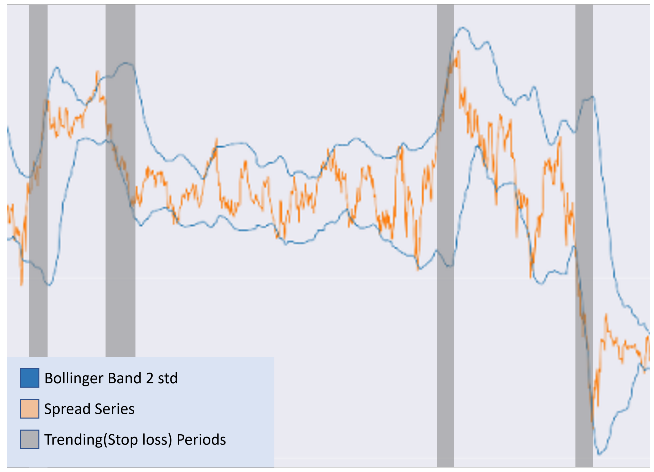

Fig. 7. An illustration of tracking the Bollinger bands and trending periods with the spread series for the European indices STOXX50E & N100

5.1 Kalman Filtered Mean Reversion

One econometric flaw of the Bollinger Band [Fang et al., 2014, ref[4]] approach is that it is applied mainly to a single stock's price series, which is non-stationary by construction. The constant drift of stocks taxes its mean reverting profits, making it unprofitable in the long run.

Applying Kalman beta to pairs goes beyond fixing the stationarity issue, and also improves the mean reversion aspect to a spread series. In our data, the combination of Kalman-filtered dynamic beta and Bollinger bands is a simple but effective way to capture mean-reverting attributes of the spread.

As shown in the Fig[7], the Bollinger bands using 60-day rolling mean and 2-standard deviations above or below to capture the mean reverting attributes for non-directional periods. Setting a stop loss (in a single strategy context) or regime shift (in a multi-strategy context) becomes key to the profitability of the sub-strategy.

5.2 Copulas

One of the major advantages of using a copulas-based approach in generating trading signals is that copulas successfully capture the mispricing and the strength of the mean reversion. The degree of mispricing is measured by evaluating the probability of the observed value of the spread in the market. We argue that events deep in the tail of the copula distribution will show stronger mean reversion.

Copulas are able to identify when the two indices differ from their historical relationship. We develop an arbitrary confidence threshold and when the random variable reaches a state space which falls outside the confidence threshold, a pairs-trading opportunity presents itself. The probabilities are generated from the conditional marginal distributions of the fitted copula as a function of V given U denoted by $C(V|U=u)$ as well as the mirror reflection of this probability on the V axis: $C(U|V=v)$.

To determine when one index is underpriced or overpriced, we compute the conditional probabilities:

$$P(U \leqslant u|V=v) = \frac{\partial C(u,v)}{\partial v}$$

$$P(V \leqslant v|U=u) = \frac{\partial C(u,v)}{\partial u}$$

If the calculated probability is too low, then the index is underpriced and opposite implies that the index is overpriced. The conditions for opening long/short positions are provided below:

$$\underbrace{P(U \leqslant u|V=v) \geqslant c}_{\text{asset is underpriced, long}} \;\&\; \underbrace{P(V \leqslant v|U=u) \leqslant (1-c)}_{\text{asset is overpriced, short}}$$

$$\underbrace{P(U \leqslant u|V=v) \leqslant (1-c)}_{\text{asset is overpriced, short}} \;\&\; \underbrace{P(V \leqslant v|U=u) \geqslant c}_{\text{asset is underpriced, long}}$$

where:

- c is the confidence interval

- u and v are transformed standard normal variables

We must ensure that for both long or short positions, both of the probability criteria are met in order to open the positions. Though (Stander et al., 2013, ref[14]) do not specify the value of the confidence band, to model extreme market conditions choosing c = 0.75 suffices. In unexpected market events, we expect there to be a sharp downward movement of the index price implying that the current price is deep in the tail of the distribution. We have presented some of the results of our out of sample test. Some of the noteworthy results we found are that the S&P 500 Utilities and S&P 500 Electric Utilities Index pair produces a Sortino ratio of 1.40 which suggests that the strategy has produced a strong return when trading on the downside.

5.3 Machine Learning

We also explored a machine learning approach for a pairs trading strategy. Machine learning methods have been found to be largely successful in capturing non-linear relationships in data with some caveats. In the case of financial time series modelling, out of sample performance degrades rapidly and it is easy to introduce look-ahead bias.

Some of the earliest proposed machine-learning approaches for pairs trading involve the use of Support Vector Machines (Montana and Parrella, 2009, ref[11]) and shallow Artificial Neural Networks (Dunis et al., 2006, ref[3]).

The recent popularity of Deep Learning has seen many methods proposed to use large neural network models to capture non-linear co-movement to estimate statistical arbitrage pair trading strategies. The most popular models range from Long-Short-Term Memory (LSTM) (Flori et al., 2021, ref[5]) to utilizing state of the art models such as Transformers (Guijarro-Ordonez et al., 2022, ref[9]).

Jorge Guijarro-Ordonez et al. (2022) demonstrate that a multi-asset arbitrage strategy combined with asset characteristics and factor residuals perform significantly better than models with simply return information in the context of Deep Learning. In our preliminary experiments with Deep Learning methodologies we found that out of sample performance tends to degrade rapidly. Given that daily data over 30 years provides roughly 8000 observations per asset, we deem the data to be too sparse for training a robust Neural Network strategy with only daily prices.

Our proposed method uses the lagged out of sample residuals of the factor models projected onto the asset return space. We combine the lagged residuals with the lagged spread of the pair in consideration and train a regression model to predict the spread. While the optimization objective is to obtain a tight regression, we do not aim for forecasting accuracy, but rather a directional movement signal with no look-ahead bias. We find that this approach is capable of picking up momentum patterns and generates the correct signal ex-ante to take advantage of a temporary spread divergence.

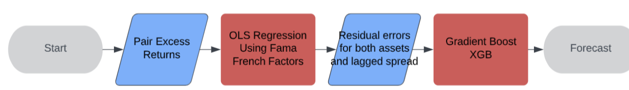

Fig. 8. Workflow of the XGBoost approach

5.3.1 Factor Residuals

We follow Jorge Guijarro-Ordonez et al. (2022) and define our primary features as the residuals from a linear projection of the factor model space onto the return space.

$$R^e_{2,t} = \beta^T_{2,t-1} F_t + \epsilon_{2,t} \qquad (6)$$

$R^e_{2,t}$ denotes the vector of excess returns of two indices at time t, with $t \in [0, T]$. The K factor models $F_t \in R^{TxK}$ capture systematic risk. $\beta^T_{2,t-1}$ represents the information set available at the previous time step, allowing us to obtain an out-of-sample forecast for the residual at time t. In our analysis we use the the Fama-French 5-factor model (Frama & French, 1993, ref[6]) as it is an observed and traded portfolio, setting K=5.

As the residual components should annihilate systematic asset risk, a spread forecasting method using the lagged residuals as features should display trend characteristics that are more driven by non-systematic factors.

5.3.2 Regression Model

We experiment with gradient boosted trees for our forecast model. XGboost (Chen et al., 2016, ref[1]) is a particular implementation of gradient boosted trees that is known to be robust to overfitting. The model has an inbuilt statistical method for assessing feature importance by evaluating predictions across trees and voting on the best predictors. This is a more meaningful approach to evaluating feature importance than stating the coefficient of a best fit curve across the feature space.

We use L lagged residuals and spread returns available as of t–1 as features for a forecast at t. We arbitrarily choose $L = 10$, but observe that lags beyond 6 are not considered useful by the model in most pairs. The results discussed are for the pair S&P 500 Utilities and S&P 500 Electric Utilities.

5.3.3 Model Results

We examine the out-of-sample predictions against the realized spread and note that the forecast is a stationary series with a mean close to zero and low variation. The deviations from near zero are observed to coincide with the divergence in the spread of returns. This suggests a very simple strategy - use any significant movements of the forecasted spread as a signal for entering a momentum position. Fig[9] illustrates the out-of-sample spread prediction from 2008 to 2014.

We use XGboost's in-built functionality to assess the feature importance of our regressors shown in Fig[10]. Unsurprisingly, we note that the residuals at $t-1$ for both indices have the highest F scores for feature importance. This aligns with experiments conducted by Jorge Guijarro-Ordonez et al. (2022).

5.3.4 Trading Strategy

We experiment with a strategy that is parsimonious and implementable. We check if the forecast of the cumulative return of spread at time

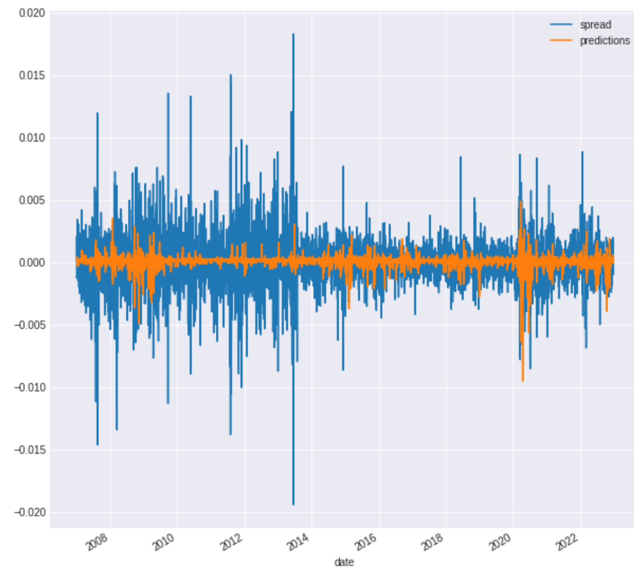

Fig. 9. Out of sample spread forecast of the Model compared to realized spread for SP500-5510 & SP500-551010

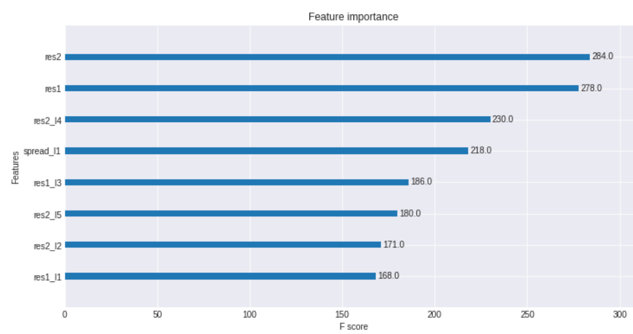

Fig. 10. F-Scores of Features Evaluated by XGBoost

t exceeds the 100 day moving average of the cumulative forecast up to $t-1$. Exceeding this threshold implies a buy signal and being under this threshold implies a sell signal to profit off of momentum based divergence. Fig[11] is the cumulative performance of the strategy with the cumulative return of spread as reference.

Interestingly, we note that the strategy is able to pick up on lagged indicators to gain the upside of the spread during the 2008 crisis and the 2020 Covid-19 crisis, and tracks the spread for the most part during the bull run period in between. One disadvantage of this strategy is that it always enters a new position re-balanced daily and could thus benefit from stricter trading criteria to only enter the market when large movements are anticipated so that

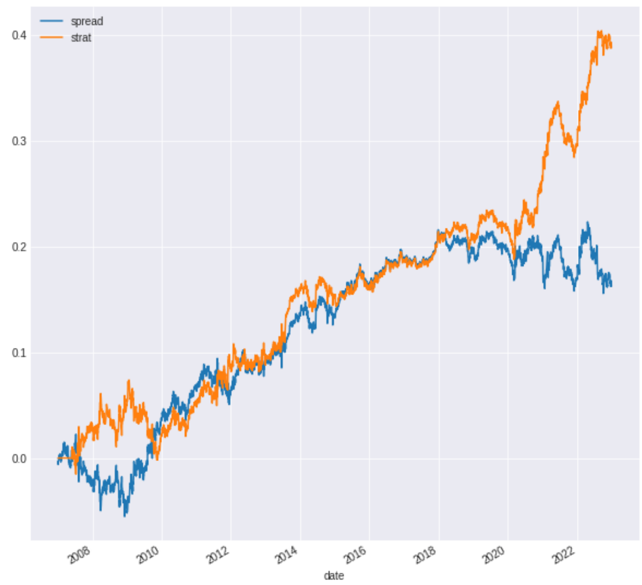

Fig. 11. Performance of a simple strategy on forecast mean reversion compared against the actual return of spread of SP500-5510 & SP500-551010

trading costs, and holding costs if considered, would not stack up significantly. The copula strategy is more conservative with its entry signals in this regard.

6 Regime Shift

Knowing the strategic attributes of our three sub-strategies, we construct a regime shift mechanism, where we utilize each sub-strategy during periods with their respective favorable market conditions. The mean-reverting strategy performs well under restricted price movements, while the copulas anomaly detection model favors extreme times where short-term reversions are expected. The machine learning model observes the patterns of unexplained residual on the factor model projections on the return space. One prominent parameter to determine regime shift turns out to be the volatility of the spread series.

From the analysis above, we conclude the dominance of mean reverting strategy in low volatility environments, the copulas anomaly detection model prevalent in high volatility markets, and the machine learning model for middle volatility times. The Three-State Variance Switching Model thus becomes the underlying regime shift driver for our model.

The results we obtained from the Three-State Variance Switching Model is shown in Fig[12]. Given that the sum of probabilities under all regimes equals 1, we will directly assign the probability of each state as the strategy weight in our regime shift model. Fig[12] illustrates the probabilities assigned by a Markov Switching Autoregressive model to each of the three volatility regimes over time.

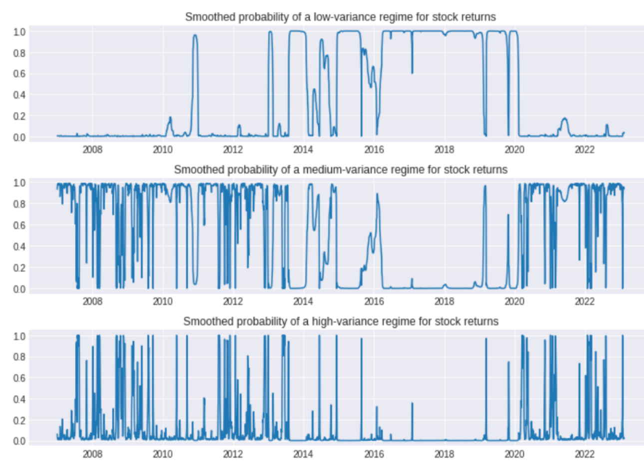

Fig. 12. Probabilities of three volatility regimes over time for the pair SP500-5510 & SP500-551010

7 Results

Our criteria for a successful all-weather strategy encompasses the following traits:

- A set of ”periodically” outperforming sub-strategies
- Low correlation across sub-strategies
- A filter that successfully captures the winning strategy or avoids the losing period of sub-strategies.

We then evaluate the above-mentioned criteria, their relative performance to peers, and their correlation against major asset classes to depict the effectiveness of our strategy.

7.1 Regime shifts and sub-strategy returns

The return series of our sub-strategies highlights the low-correlation aspect of our three sub-strategies, with cross-correlation of 0.023, 0.089, and 0.021 making this a diversified toolbox for multi-strategy implementation. The strategic allocation bands, on the other hand, show that the regime shift filter is able to capture the upsides of an outperforming strategy and reduce the downside risk of the overall strategy, fulfilling our expectations for an all-weather strategy. Fig[13] shows that although high volatility periods are sparse, the copulas anomaly detection strategy in high volatility periods successfully captures the majority of the returns in the middle to high volatility periods which are key drivers for return. We consider the period before 2007 to be our in-sample period for calibrating our models and the period after 2007 as our out-of-sample performance evaluation period. It is worth noting that the Kalman-Filter updates on a rolling basis, the copula method re-calibrates its parameters on a semi-annual basis, and the machine learning model is found to show improvements on a yearly rolling basis. For the machine learning model a yearly rolling basis shows an incremental improvement in strategy performance however we note that the initial in-sample period is the most important for generating a viable forecast model. Retraining the machine learning model at a higher frequency tends to overfit the training data resulting in worse out-of-sample performance.

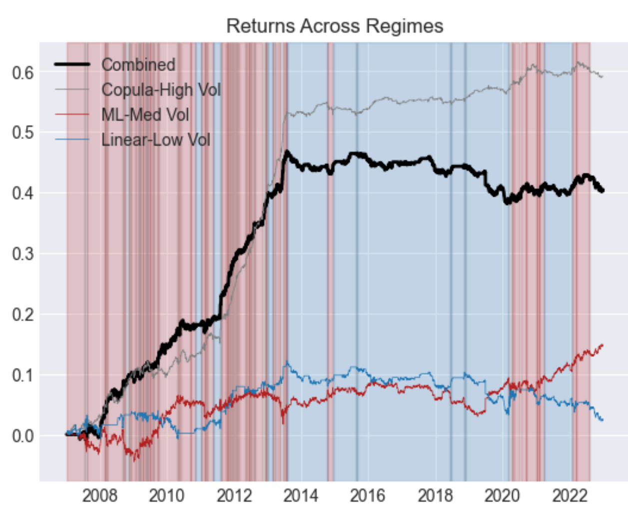

Fig. 13. Sub-strategy performance and allocation band for SP500-5510 & SP500-551010

7.2 Risk and Return profile

Compared to Hedge Fund Research, Inc.(HFR) index peers, our All-weather strategy is generating comparable returns to the Equity Market Neutral index and Multi-strategy index, while suffering significantly lower maximum drawdowns due to our regime shift mechanism in varying volatility markets shown in Fig[14].

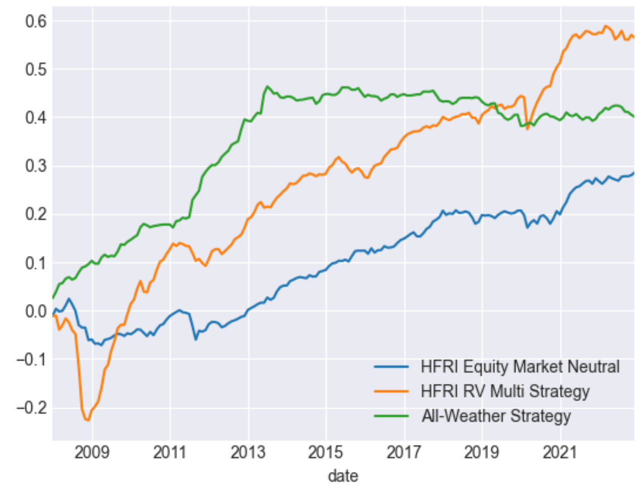

Fig. 14. Benchmarked returns for SP500-5510 & SP500-551010

		Average Excess Annual Return	Annualized Sharpe	Annualized Sortino	Max Drawdown
SP500-5510	SP500-551010	2.64%	0.95	1.39	8.65%
DJU	SP500-551010	2.22%	0.69	1.06	5.49%

Fig. 15. Returns summary for final multi-strategy on selected pairs

7.3 Diversification profile

Measuring the strategy's return profile against major asset classes, we discover low correlations against Stocks (IVV), Bonds (AGG), Real Estate (VNQ), and commodities (DBC) as Fig[16] shows, making the strategy an incredible diversifier for asset allocation.

8 Conclusion

Throughout this paper, we have covered multiple non-linear methods to account for the non-linearity and fat-tailed nature of co-integrated

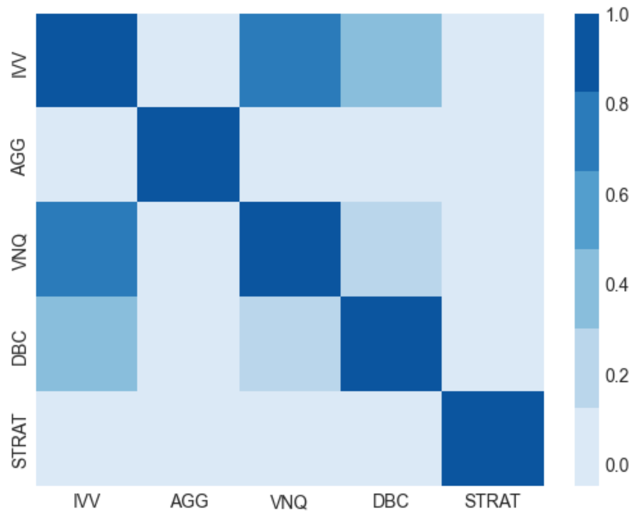

Fig. 16. Correlation heatmap across major asset classes

equity indices. We have used state-space models such as Kalman Filter and Three-State Variance Switching Model to capture the dynamic aspects of time series financial returns. We explored Copulas and Machine Learning models to capture the hidden higher-order relationship between two assets over time.

In our attempt to approximate the underlying dynamics of the ever-changing financial markets, we conclude that the non-linear methods we have applied work in tandem to capture the structural dependency of one financial time series on another. By combining the copula and machine learning methods our model will smooth out any implicit prediction bias that a singular model would hold. Our out-of-sample returns show a prominent absolute return profile which has low correlation with major market indices while matching performance with our HFR peers.

Real world implementation of our strategy would require accounting for trading costs and other market frictions. Trade execution would need to be conducted using ETFs and Futures.

Our multi-strategy framework yields fruitful and robust results as can be seen for the S&P 500 Utilities and S&P 500 Electric Utilities index pair. It is interesting to note that our framework performs best on indices that have a large number of overlapping constituents. Intuitively one would expect these indices to have similar prices regardless of market shocks and yet our strategy is able to anticipate mispricing opportunities. When comparing our trading strategy with the HFR peers, we find that our trading strategy has performed considerably better. We found our average excess returns to be 2.64%, Sharpe Ratio of 0.95 with a Maximum Drawdown of 8.65%. When compared with our HFR peers, we see that our peers have average excess returns of 2.59%, a Sharpe Ratio of 0.69 and a Maximum Drawdown of 9.15%. We can conclude that our Multi-Strategy Framework can be used as an effective stand-alone strategy or as a diversifying portfolio.

References

[1] Chen, Tianqi, and Carlos Guestrin. “XGBoost.” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016

[2] Clive W. J. Granger, ”Time Series Analysis, Cointegration, and Applications”, The American Economic Review, Jun., 2004, Vol. 94, No. 3 (Jun., 2004), pp. 425

[3] Dunis, Christian L., et al. “Higher Order and Recurrent Neural Architectures for Trading the EUR/USD Exchange Rate.” Quantitative Finance, vol. 11, no. 4, 2011, pp. 615–629.

[4] Fang, Jiali, et al. “Popularity versus Profitability: Evidence from Bollinger Bands.” SSRN Electronic Journal, 2014

[5] Flori, Andrea, and Daniele Regoli. “Revealing Pairs-Trading Opportunities with Long Short-Term Memory Networks.” European Journal of Operational Research, vol. 295, no. 2, 2021, pp. 772–791.

[6] Fama, E. F. and K. R. French, ”Common risk factors in the returns on stocks and bonds”, 1993

[7] Gabaix X, Gopikrishnan P, Plerou V, Stanley HE. ”A theory of power-law distributions in financial market fluctuations.” Nature. 2003 May 15;423(6937):267-70.

[8] Ramos Requena, J.P.; Trinidad Segovia, J.E.; Sánchez Granero, M.A. ”An Alternative Approach to Measure Co-Movement between Two Time Series.” Mathematics 2020, 8, 261.

[9] Jorge Guijarro-Ordonez et al., ”Deep Learning Statistical Arbitrage”, SSRN Electronic, 2022

[10] Krauss, Christopher. “Statistical Arbitrage Pairs Trading Strategies: Review and Outlook.” Journal of Economic Surveys, vol. 31, no. 2, 2016, pp. 513–545.,

[11] Montana, Giovanni, and Francesco Parrella. “Learning to Trade with Incremental Support Vector Regression Experts.” Lecture Notes in Computer Science, 2009, pp. 591–598.

[12] Ernest P. Chan, ”Quantitative Trading: How to Build Your Own Algorithmic Trading Business”, 2008

[13] Xu, Han. “High Frequency Statistical Arbitrage with Kalman Filter and Markov Chain Monte Carlo.” University of Waterloo, 2017.

[14] Yolana Stander, Daniël Marais, Ilse Botha, ”Trading Strategies with Copulas”, University of Johanessburg 2012