CSE 578 : Computational Investment

Final Project Report

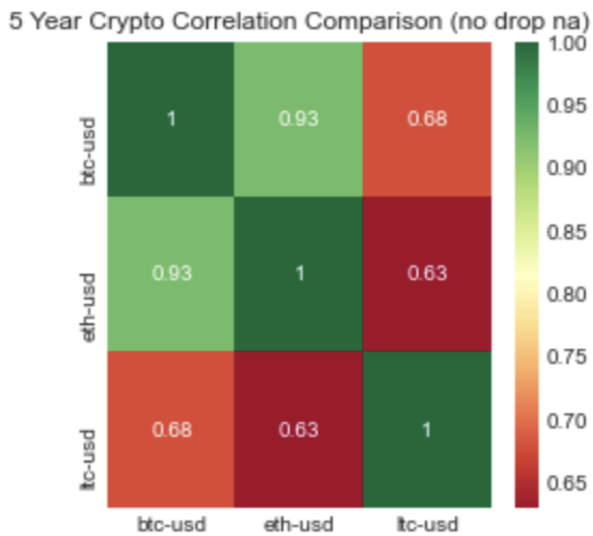
Authors:
Josh Caskie, 50425015, jmcaskie@buffalo.edu
Ashutosh Shailesh Bhawsar, 50416025, abhawsar@buffalo.edu
Zachary Giannetta, 50274012, zlgianne@buffalo.edu
Ethan Nhan, 50270070, ethannha@buffalo.edu

Instructor: **Prof. Zhen Liu**

Date Last Edited: December 9, 2022

A study and backtest of correlation moving average and its implications for creating a profitable trading algorithm.

This study was motivated by the noticeably large correlation effect in cryptocurrency markets. [fig. 1]. Our goal was to understand this situation and identify its potential utilization for a profitable trading algorithm. We were also interested in working with an unconventional idea and innovating something new. Basic research yielded numerous insights and drawbacks to strictly correlation-based trading. Mean reversion was a highlighted concept throughout many sources which suggests a strong correlation is not one-to-one with a strong chance of reverting back to a mean price–there may be unknown variance at play [1,2]. Another was the idea that observable correlations between any two stocks may not resolve to valid inter-stock dependencies but rather to the adhesive effect between those assets and a larger index [3]. We attempted to diagnose these thoughts through our hypotheses below:



1. **Development of CORRMA: Correlation Moving Average Indicator:** Determine if trading on a pair based on the indicator outperforms buying and holding either asset in the pair (crypto, stocks, sectors).
2. **Applied the SARIMA statistical model to the crypto time series:** Transparent statistical regression model to fit and predict price.

1 CORRMA

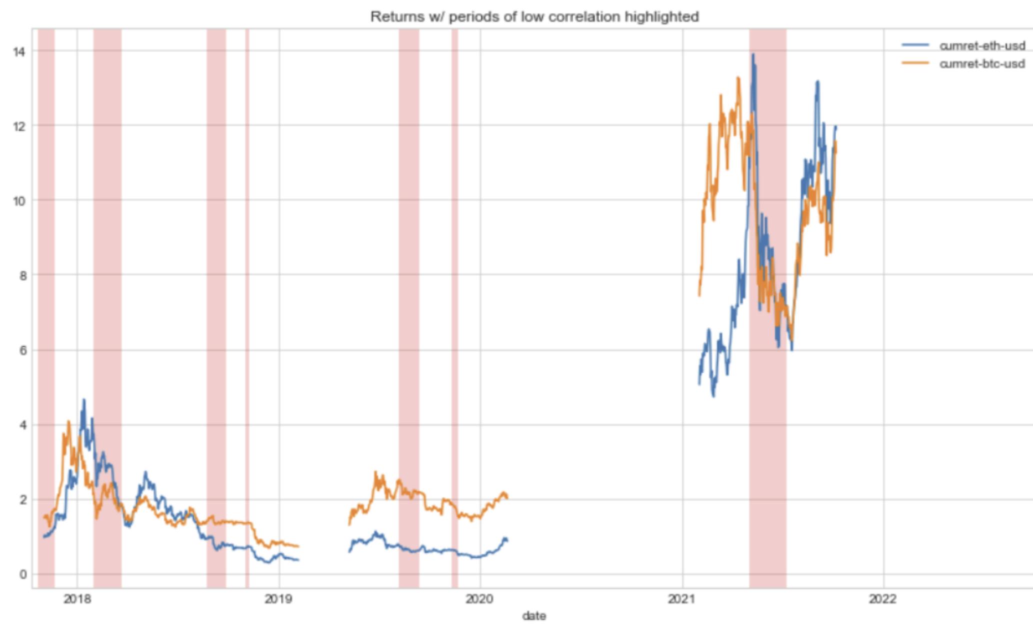
The CORRMA indicator is an extension of broad correlation calculations to a moving average which aids in tracking trends based on a rolling data window over an asset's timeframe. The initial trading algorithm relied on daily price data and two parameters:

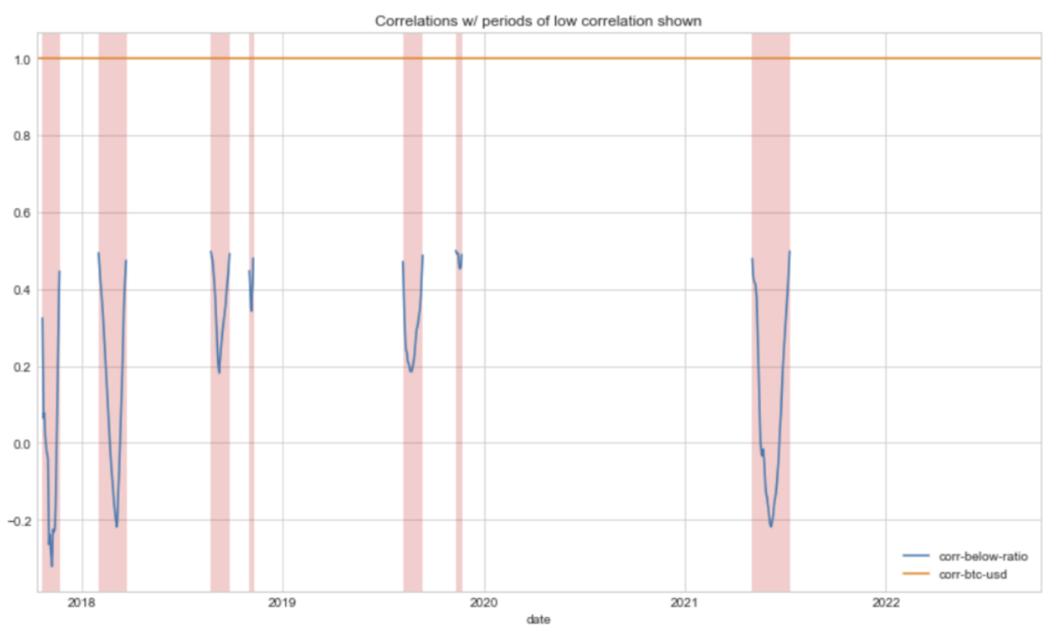
- WINDOW = rolling timeframe for correlation computations.
- THRESHOLD = correlation threshold to trigger an action.

Each day:

- Compute past CORRMA using WINDOW days.
- If correlation is $\leq$ THRESHOLD, swap the original asset.
- If correlation is $>$ THRESHOLD, swap back to the original asset.

Our initial analysis was to determine moments of lower correlation between highly correlated assets such that these moments would present a trading signal. We first created a Jupyter notebook to find periods of low correlation based on window and threshold inputs. The red regions below are where the correlation falls below the threshold, and our algorithm should make a trade (BTC vs. ETH). This graph highlights CORRMA(90 day, 0.3)





Next, we implemented the algorithm on QuantConnect for backtesting and analysis. The repo link is attached below. Through numerous backtests and optimizations, we came to various conclusions regarding CORRMA's usability as a basic technical indicator.

2 Cryptocurrencies

Bitcoin is the largest cap asset in the crypto market and carries a large correlation between it and the other top-10 coins, one of which is Ethereum. We thoroughly studied the pairing to determine if our algorithm was successful (see technical appendix for details). It started with buying BTC, waiting for the indicator to initialize, and then swapping to ETH when the threshold is triggered. Our process started with testing our novel approach and then iteratively generating new hypotheses by understanding the Sharpe ratio, average win/loss rates, and other trends we noticed through the data.

1. Novel analysis → **failed**
2. Compete divergence CORRMA → **failed**
3. Changing time frame → **small breakthrough**
4. Swapping the original asset → **major breakthrough**
5. Adjusting the algorithm trading criteria → **no change**
6. QuantConnect Optimizations → **success**
7. Comparing results with other cryptocurrency pairs → **success**
8. In-sample vs. out-of-sample testing → **failed, needs more analysis**

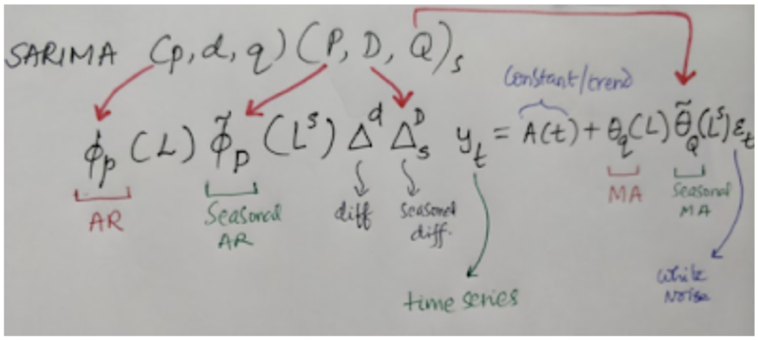
Our conclusions indicate that an in-sample backtest/optimization for an ETH v. BTC algorithm improves upon a 5-year B&H strategy by 3.75x, which, when applied to LTC v. BTC and ZEC v. BTC pairings, also netted 1.15x and 1.6x returns respectively. However, out-of-sample prediction does not give the same results and leaves room for future development.

3 Stocks vs. ETFs vs. Large Market

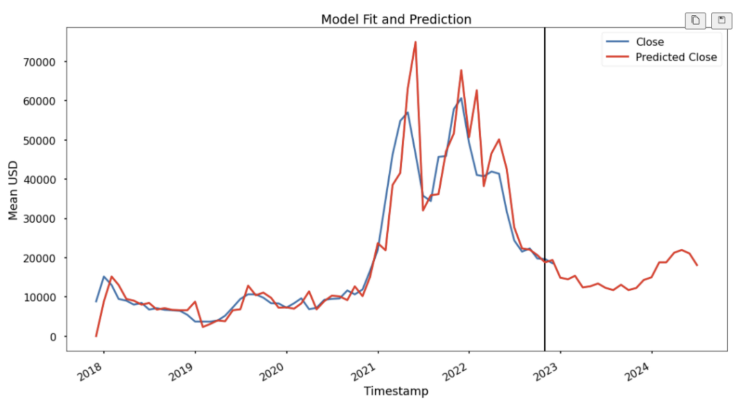
We started by comparing stock-to-stock in the same market and sector-to-sector. After reviewing academic reports from peers, the idea shifted towards comparing individual stocks to their sectors as a whole. We also compared individual sectors to the larger market. This evolved into the final version where we picked a single stock Amazon vs. its sector market XLY. For sectors to larger markets, we used XLK vs. SPY. We created B&H benchmarks for Amazon, XLY, XLK, and SPY and tested CORRMA against the benchmarks to see if we outperformed. After testing different variations against the benchmark, we found that a window of 90 days and a high to medium CORRMA threshold gave the best results for stocks-to-sectors. For sectors-to-larger markets, it was found that a window of 150 and a high threshold worked the best. These optimal variables were then tested against different time periods to see if they stayed consistent.

4 ARIMA

We acquired small-cap crypto data, including BTC, for the POC and preprocessed it to determine a suitable timeframe showing if the data is stationary (ADF test). We determined the best parameters using autocorrelation/partial autocorrelation plots and determined the best fit using AIC. Finally, we created a future data frame for prediction and visualization. Following is the description of each parameter in SARIMA equation:



Prediction and model fit on monthly BTC data:



Model Fit and Prediction

Backtesting confirms ARIMA's inability to outperform B&H, while SARIMAX is on-par.

	Capacity	Equity	Fees	Holdings	Net Profit	PSR	Return	Unrealized	Volume
Buy and Hodl:	$2000000B	$1,640,017.87	-$3,484.66	$0.00	$1,543,502.69	31.031%	1,540.02 %	$0.00	$1,742,208.77
ARIMA:	$240000B	$165,844.04	-$532.70	$0.00	$66,376.75	4.043%	65.84 %	$0.00	$266,151.74
SARIMAX:	$110K	$1,634,062.66	-$3,472.73	$0.00	$1,537,535.54	30.941%	1,534.06 %	$0.00	$1,736,241.62

Backtest results

Summary for best model selected using AIC:

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                           Close_box   No. Observations:                   61
Model:             SARIMAX(1, 1, 0)x(2, 1, 0, 12)   Log Likelihood                  102.596
Date:                             Sat, 03 Dec 2022   AIC                            -197.192
Time:                                     20:39:53   BIC                            -189.707
Sample:                                 11-30-2017   HQIC                           -194.364
                                      - 11-30-2022
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.4797      0.077      6.244      0.000       0.329       0.630
ar.S.L12      -0.4682      0.221     -2.120      0.034      -0.901      -0.035
ar.S.L24      -0.5576      0.121     -4.622      0.000      -0.794      -0.321
sigma2         0.0007      0.000      4.794      0.000       0.000       0.001
===================================================================================
Ljung-Box (L1) (Q):                   1.13   Jarque-Bera (JB):                14.45
Prob(Q):                              0.29   Prob(JB):                         0.00
Heteroskedasticity (H):               0.19   Skew:                            -0.67
Prob(H) (two-sided):                  0.00   Kurtosis:                         5.33
===================================================================================
```

Best model output

We conclude that SARIMA is close to B&H because it considers a rolling window for auto/partial correlation along with seasonal changes to regression parameters; thus, it maintains the latest trends for prediction. Also, ARIMA is not profitable for unstable (non-stationary) markets like crypto.

5 Conclusions and Future Work

Overall, each study of CORRMA and the respective trading algorithm showed strong potential. There are various future considerations that we'd like to explore:

- Deeper in-sample vs. out-of-sample backtesting.
- Develop more algorithms (i.e., time to stay out of the market).
- SARIMA model extended to include eXogenous features.

As a retail trader, if you are fine with the risks associated with holding EITHER asset for the long run, it might be better to use this CORRMA algorithm. Based on our optimized results, we at least perform about as well or much better than B&H with large upside potential.

REFERENCES

Our Data and Resources

- GitHub: https://github.com/joshcaskie/478-marketbusters
- Strategy Backtesting Sheet: Docs
- Final Presentation Slides: Docs

Primary References

[1] QuantConnect. Intraday Dynamic Pairs Trading using Correlation and Cointegration Approach. https://www.quantconnect.com/tutorials/strategy-library/intraday-dynamic-pairs-trading-using-correlation-and-cointegration-approach.

[2] Pollet, Joshua M, and Mungo Wilson. "Average Correlation and Stock Market Returns." Journal of Financial Economics, ScienceDirect, 24 Feb. 2010, https://www.sciencedirect.com/journa of-financial-economics.

[3] Shapira, Y., et al. "The Index Cohesive Effect on Stock Market Correlations." The European Physical Journal B, vol. 72, no. 4, 2009, pp. 657–669., https://doi.org/10.1140/epjb/e2009-00384-y.

Secondary References

Bhat, Shripad. "ARIMA/SARIMA with Python." Data Science, 1 Jan. 2019, www.datasciencesmachine in-python.html.

Flavin, Thomas J, et al. "Explaining Stock Market Correlation: A Gravity Model Approach." The Manchester School, vol. 70, no. S1, 2002, pp. 87–106., https://doi.org/10.1111/1467-9957.70.s1.5.

Kwapień, J., et al. "The Bulk of the Stock Market Correlation Matrix Is Not Pure Noise." Physica A: Statistical Mechanics and Its Applications, vol. 359, 2006, pp. 589–606., https://doi.org/10.1016/j.physa.2005.05.090.

Wichard, Jörg D., et al. "Detecting Correlation in Stock Market." Physica A: Statistical Mechanics and Its Applications, vol. 344, no. 1-2, 2004, pp. 308–311., https://doi.org/10.1016/j.physa.2004

Data Resources

CoinCodex - Cryptocurrency data

fmpcloud - Cryptocurrency data

AlphaVantage - Stock, ETF, and Cryptocurrency data

Documentation References

QuantConnect:

Custom Indicators

Custom Indicator Algorithm

Custom Indicators

QuantConnect indicators Introduction

Class Resources and Example Code by Prof. Zhen Liu