Team Name: Solarity Capital Partners
Strategy Name: KAL-EGY

From Black to Green Gold: Arbitraging to Net Zero with the Kalman Filter

***Abstract**. Our strategy takes the well-documented classic strategy of statistical arbitrage and projects it onto a wide-ranging variation of companies within the energy space, ranging from oil majors to new technologies such as hydrogen fuel cells. Our strategy uses statistical arbitrage and applies novel forecasting methods to identify cointegrating pairs yielding a market-neutral return. The strategy achieved 14.23% CAGR in an out-of-sample back-test between Q2 2015 and Q2 2024, with an annualised Sharpe Ratio of 1.09.*

Economic Hypothesis

History of Statistical Arbitrage and the Kalman Filter

Statistical arbitrage, the method underpinning our strategy, derives its roots from more basic pairs trading strategies, improving upon them through exploiting volatility driven pricing inefficiencies in the market, returning equities to their fair market price through investing in a cointegrated pair of equities to return the spread to equilibrium through mean reversion. Initial development and adoption commenced during the latter half of the 1980s and the strategy now comprises a major chunk of the operations of many firms in the investing space, especially within quantitative hedge funds. Arbitrage implies the deviation of at least one financial instrument from its fair market value without changes in fundamentals, with the mean reversion of spread exploited also conforming to some interpretations of the efficient markets hypothesis as first developed by Paul Samuelson and Eugene Fama in the 1960s. The Kalman filter is a forecasting method developed by Rudolf Kalman in 1960 which we use to predict logarithmic future price data from a previous quarter of cointegration.

Model Novelty

Our strategy aims to exploit a niche within the energy market by considering the possibility of cointegration between an equity pair combining renewable and fossil energy companies, as well as cointegration between equities in each class, a novelty not previously detailed within relevant literature. There are strong reasons for long-run price equilibria within the oil and gas industry; companies typically have little diversification, and their products are subject to the whims of the global market price of oil, as well as sharing common equipment suppliers. Renewable energy should also see long-run equilibria due to relative dependence on government intervention to maintain viability, such as tax credits and subsidies like the Inflation Reduction Act, as well as sharing common customers through the provision of a utility, which is subject to significant public price intervention, an example being the government-mandated energy price cap in the United Kingdom. However, we also make the case that price ratio equilibrium should cut across the energy divide. The Russian invasion of Ukraine which sent oil prices skyrocketing to near-record highs had the simultaneous effect of turbocharging renewable energy investment as part of the long-run policy response to promote energy independence away from the volatile fossil fuel market. Likewise, lower oil prices should also suppress renewable energy investment due to fossil energy becoming cheaper than profitless renewable energy generation, and therefore becoming unviable without significant public financial support. Literature offers some support - Hongli (2021) finds that there is a moderate positive correlation between WTI and the WilderHill Clean Energy Index which is significant at the 1% level, and Hsiao et al (2019) find that "the returns for both international oil and China's renewable markets exhibit the characteristics of time-varying, volatility clustering, and similar motility". We believe that the novelty of this approach should allow us to harvest market-neutral returns in a section of the market as yet untested by other such funds.

The strategy should also benefit through exploiting the well-documented volatile nature of both markets. Given that our thesis establishes that there is a strong case for co-integration within our two equity classes and between them, the volatile nature of energy markets should provide ample pricing inefficiencies which the strategy is well-placed to exploit.

Investable universe

Our investable universe can be differentiated into two different groups, a series of 31 oil companies and equipment facilitators, and a smaller series of 21 renewable energy companies whose facilitate renewable energy. This equity pool is smaller due to the nascent nature of much of the technologies behind renewables but will expand over time as the industry continues to grow. We have constrained our dataset to only include companies either domiciled in the USA or have a depositary receipt listed on the New York stock exchange or on OTC markets to simplify the model and remove any FX noise. Equities were selected through an analysis of primary revenue stream, with successful candidates either providing critical minerals, producing equipment for or producing and distributing either green/renewable energy or oil and gas, deriving at least 50% of their revenue from such operations. All equities were required to have a market cap of at least $250m to minimise capital risk.

Implementation

Searching for Viable Pairs Opportunities

Our investable universe consists of 52 US listed equities within the energy sector, offering 1326 potential unique pairs that we could trade. In our back-tests, we split each pair by quarter, providing a maximum of 47,736 unique opportunities. Given the nature of some of our companies becoming emerging contenders in the energy industry, some of our renewable energy companies within our investable universe have had their initial public offerings at varying times during our back-test so this number in reality is smaller.

The general aim of the model was to find a linear combination using the price series of our chosen pair of equities X_1 and X_2 to solve for Y such that:

$$Y = \beta_1 X_1 + \beta_2 X_2$$

where the parameters $\beta_1, \beta_2 \in \mathbb{R}$ are unrestricted and cointegrating. The unrestricting property is implied from no optimising constraints being placed on either of our parameters. Assuming that this model holds, it further implies that $Y \in I(0)$.

Verification for this relationship was established by regressing the pairs of equities against one another using Ordinary Least Squares, similar to that first shown in (Engle and Granger, 1987) creating the following model:

$$X_1 - \beta_2 X_2 = \alpha + \varepsilon$$

where β_2 is the OLS estimator of the second asset when predicting X_1. We then performed two tests on the residuals to verify that these identified pairs were both cointegrated and also stationary.

Firstly, the residuals were tested to assess whether the model exhibited serial correlation by calculating the Durbin-Watson test statistic. Here, we included four lags in the model as the data was separated on a quarterly basis, with the sample size T being the number of trading days in that quarter. T was allowed to be variable as the number of trading days differ every quarter due to public holidays and other global events, however, the number of trading days in our sample seemed to not significantly differ from 255 annually (or 63.75 per quarter). The DW statistic took the null hypothesis that the regression was correctly specified (i.e. did not exhibit autocorrelation) and the alternative hypothesis that it was mis-specified, hence rejecting the null within a 95% confidence interval allowed our model to discard pairs that were confirmed to be mis-specified with statistical significance. A general risk of specification tests, however, is that the cointegrated pairs we accept here *could* be correctly specified as we lack sufficient evidence to prove that is not. Therefore, although the DW test increases the accuracy of our pairs, it does not eliminate all of those that are mis-specified.

Secondly, an Augmented Dicky-Fuller test was used on the residuals to verify whether our model of cointegrated pairs was also now stationary, where the test statistic took the null hypothesis of $\gamma = 0$ against the corresponding alternative hypothesis of $\gamma < 0$. Again, we used four time-lags in our test and if the null hypothesis was not rejected – our pair was deemed invalid in this time period and therefore discarded.

Forecasting Future Values of Pairs

When forecasting future values of our pairs, we used novel techniques from forecasting to trade execution to create our signals. Firstly, we applied the Kalman Filter to the log price data of our identified quarter and used this to forecast our next quarter of log stock price data. We proceed to generate long/short signals using the hedge ratio of our forecasted log prices by extrapolating our OLS model and using z-scores for these future forecasted values using the following calculation:

$$z = \frac{x - \bar{x}}{\sigma_x}$$

where x is the value of the forecasted log hedge ratio on a given day, $\bar{x}$ is the mean value of the forecasted log ratios and σ_x is the corresponding standard error of the series. We proceeded to long this spread when the $z \leq -2$ (so our trade executes with the weightings β_1, β_2) and shorted this spread when $z \geq 2$ (so our trade executes with the inversed weighting, i.e. $(-1(\beta_1, \beta_2))$. This represents our model forecasting prices that were outside of the 95% confidence interval and would be a statistically significant investment opportunity.

Our model prevented trades from executing when this relationship has suddenly broken down by discarding signals when the log spread of the forecasted prices differed to the true log spread on a given trading day by at least 2.5%. We proceeded to exit our positions in the following scenarios:

- When the z-score returned to 0 and proceeding to take profit.
- When our position reached a maximum 30% take profit or 15% stop loss.
- It was the last day of the quarter and none of the above have been met.

Portfolio Construction

When constructing the wider portfolio, we assumed that the fund was unleveraged, with our outstanding cash being the sole means of collateral (i.e. existing holdings were not able to be used to cover loss-bearing trades). This meant that our strategy always held sufficient cash in its brokerage account to cover a margin call from loss-bearing short positions, should it occur. Each pairs trade was treated equally and so an equal amount of capital was allocated to each position. The amount of capital deployed was also restricted to 3.5% of its 1-day average trading volume, which was particularly important when considering the lack of liquidity some of the depository receipts within our investable universe, such as Subsea 7 SA (SUBCY).

Liquidity and Capital Considerations

As with all arbitrage strategies, there is an upper limit on how much capital can be deployed to any trade. As economic agents exploit these market inefficiencies, they increase overall market efficiency, hence further reducing the scope for arbitrage. If there were more entrants within the market with a similar strategy (or a fund that is simply willing to deploy more capital), they may augment these statistical relationships and decrease potential profits. This poses a risk to our strategy through financial cannibalism, although we feel that this is minimised due to the novel cross-class niche of our investable universe.

Regarding capital deployment into individual securities, the filtering of companies within our investable universe should eliminate most problems arising from deploying too much capital into small-cap companies, with the smallest company in the dataset being Piedmont Lithium Inc, with a market capitalisation of $279.7m. This potential problem with illiquidity is mitigated through our restrictions outlined within Portfolio Construction. Within our wider investable universe, the 52 equities have a 1-year mean intraday trading volume of $6.34m per equity, however, this is somewhat skewed upwards by BP and Plug Power, despite this our 1-year median intraday trading volume still reaches $3.34m per equity.

Assuming that a maximum of 3.5% of daily trading volume deployed has a negligible impact on the respective markets and that an average of 8 pairs trades were open at any given time, we should be able to deploy over $56m before the order books are severely impacted. Contrastingly, our trades should impact these market prices as a whole, as our actions help to correct market inefficiencies, bringing them back to their true fair value.

Risks

Fundamental Shocks

Macroeconomic shocks have the potential to adversely affect our strategy should they damage the price ratios between our cointegrated pairs, breaking down the relationship and invalidating a key tenet of the strategy. This may also occur with company specific fundamental shocks. Most macroeconomic shocks should affect our cointegrated pairs evenly and maintain the relationship between the price ratios, but in cases where this does not hold, our strategy stops out any trade should it hit our hard stop loss.

Investable Universe Disruption

The nature of our investable universe is such that many constituents focus on nascent technologies with structural profitability of their products yet to be proven, and as such there will inevitably be situations where companies will face bankruptcy. Conversely, many will become attractive candidates for acquisition. Given healthy throughflow of renewable energy start-ups, we believe that it should be easily possible to replace lost constituents as time passes.

Long-Run Volatility

Extended volatility resulting from severe macroeconomic shocks may pose risks for open positions in our strategy, straining liquidity and preventing the mean reversion key to ensuring our profitability. In this case, margin calls and liquidation of open positions at undesirable prices may lead to significant losses without triggering safeguards such as our spread limit. The unleveraged nature of our fund means that margin calls are not an issue, and any open positions without a sell signal are automatically closed out at the end of each quarter.

Liquidity Shortages

The increasing prevalence of other statistical arbitrage funds operating within the same markets as our strategy may cause liquidity strain as other funds hoover up the little available to deploy within their own often highly leveraged strategies. This may force our strategy to liquidate positions at unfavourable prices without being able to see through the arbitrage. Whilst the novel approach of investing in renewable and fossil fuel pairs should avoid targeting by likeminded strategies, the more conventional pairs within our strategy purview are likely to be at risk from such trades. Slippage may also lead us to be executing trades at suboptimal prices. Mitigation of these risks is difficult but our limits on position sizes should help restrict some of the possible damage.

Model Inaccuracies

Finally, inefficiencies and errors within the model may cause spurious correlations to be identified and a failure of safeguards to activate in case of a breakdown of the price ratios between pairs. Our model is based on a sound statistical framework as detailed in the economic hypothesis and together with rigorous development this risk has been minimised. However, the trades made by the model will still be closely monitored to ensure none of the constraints and safeguards are violated, with trades being manually stopped out in this case and model implementation paused until the error is found and fixed.

Risk-Adjusted Performance and Future Potential

Across our 9-year out-of-sample back-testing period, the strategy generated a total return of 231.22%, or a CAGR of 14.23%. Our daily returns had a standard deviation of 0.86% which generated an annualised Sharpe Ratio of 1.09.

Figure 1 shows our full out-of-sample back-test performed between 2nd April 2015 and 14th February 2024. The strategy begins to meaningfully return in 2020, with Covid-19 providing ample market inefficiencies but likely also breaking some of the cointegrated pairs within our model due to the drastic collapse seen especially in oil prices relative to renewables. The market recovery appears to renew many of these relationships and plentiful speculation especially in the renewables sector created much for the strategy to exploit. The general market boom of mid-2021 sent renewables flying away from fundamentals relative to oil majors and really allowed the strategy to shine, as can be seen in Figure 2. Returns since then have been choppy, with the strategy unfortunately being unable to benefit from volatility from the Russo-Ukraine war, perhaps due to a muted effect of the oil majors themselves relative to oil price.

Looking to the future, the amount of renewable energy equities in our dataset should significantly expand as hydrogen fuel cells and other nascent technology begins to break profitability. Future government intervention against major oil producers as legally binding net-zero targets begin to be enforced and the continued evolution of renewable energy through sodium-ion batteries and other disruptive technologies should provide abundant opportunities with regard to short-term market inefficiencies that our strategy is excellently placed to exploit in the future.

With access to high-frequency data, we could exploit this volatility even more efficiently than as of now and generate even greater alpha than our strategy currently provides.

Team Name: Solarity Capital Partners
Strategy Name: KAL-EGY

Figures

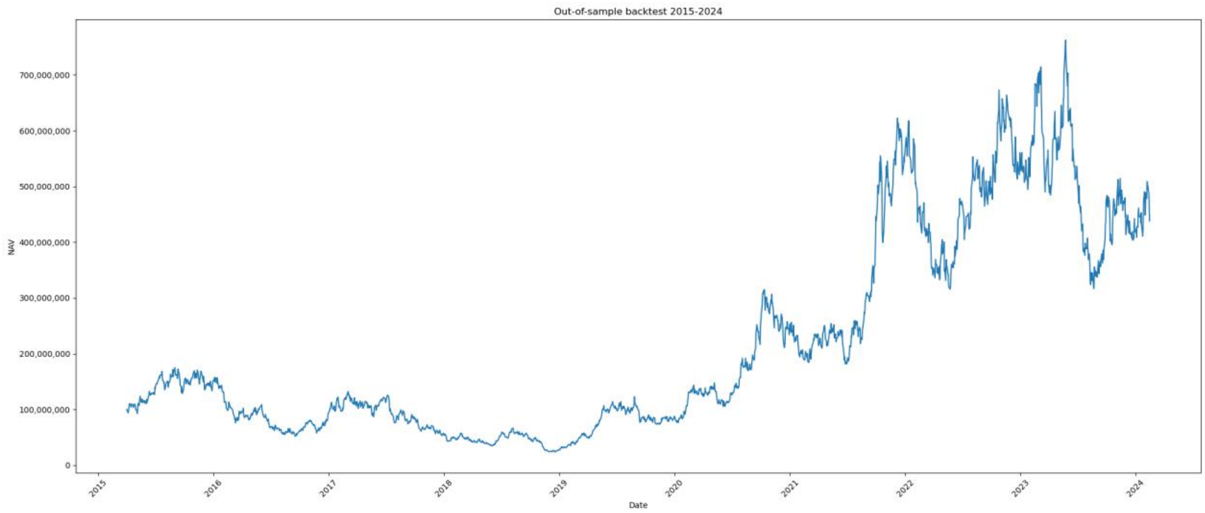

Figure 1: Out-of-sample backtest between 2015-2024.

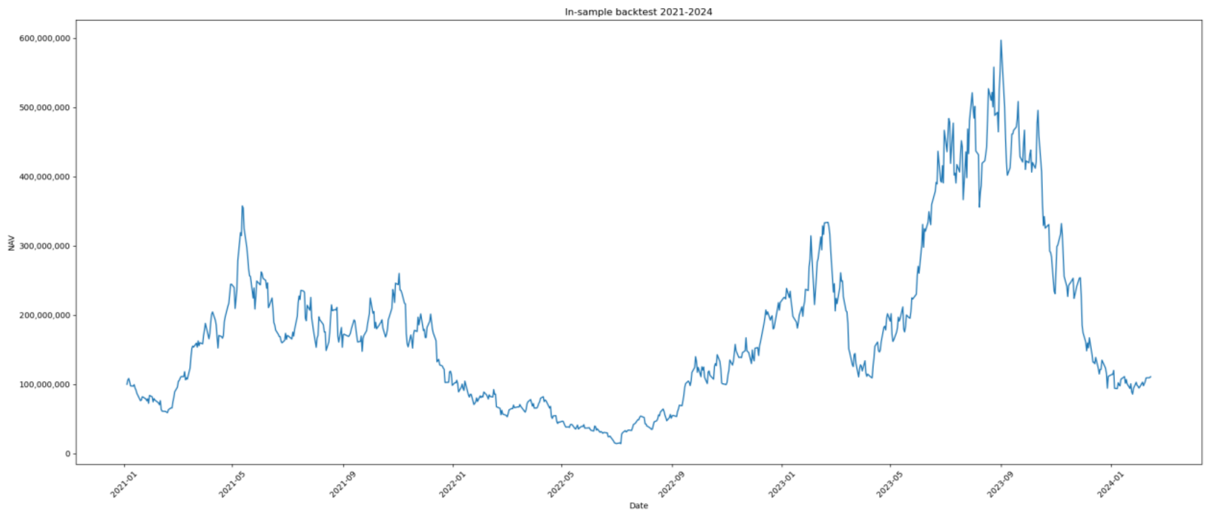

Figure 2: In-Sample backtest between 2021-2024.

Investable Universe

Company	*Ticker*	*Market Cap (USD)*	*Turnover (USD)*
Albemarle Corp	ALB	16,183,996,520	102,188,387,305
Ameresco Inc	AMRC.K	1,094,127,930	4,039,927,615
Anglo American Platinum Ltd	ANGPY.PK	10,150,335,643	120,760,393
APA Corp (US)	APA.O	8,983,276,650	45,945,333,670
Baker Hughes Co	BKR.O	29,616,080,804	60,185,717,446
Bloom Energy Corp	BE	1,973,014,245	15,748,079,861
BP PLC	BP.L	99,046,684,495	58,828,826,314
Chevron Corp	CVX	282,323,485,012	337,973,192,618
Clearway Energy Inc	CWEN.K	4,286,238,301	7,267,383,294
Conocophillips	COP	132,392,997,735	157,206,409,597
Coterra Energy Inc	CTRA.K	19,382,626,797	45,422,249,174

Team Name: Solarity Capital Partners
Strategy Name: KAL-EGY

Devon Energy Corp	DVN	27,978,100,000	105,356,601,957
Diamondback Energy Inc	FANG.O	32,570,070,329	74,182,603,271
Ecovyst Inc	ECVT.K	1,120,528,037	2,139,730,049
Edison International	EIX	26,155,341,254	33,358,035,480
EnerSys	ENS	3,713,419,783	6,280,244,021
Enphase Energy Inc	ENPH.O	17,242,793,646	152,432,481,101
EOG Resources Inc	EOG	66,387,014,269	98,091,917,024
EQT Corp	EQT	16,361,863,050	54,451,940,079
Equinor ASA	EQNR.K	74,213,235,854	22,147,281,128
Exxon Mobil Corp	XOM	414,719,086,968	479,249,640,393
First Solar Inc	FSLR.O	16,442,981,684	106,512,770,958
Freeport-McMoRan Inc	FCX	54,235,004,668	114,100,419,347
Fuelcell Energy Inc	FCEL.O	536,314,707	5,846,885,725
Galp Energia SGPS SA	GLPEY.PK	12,205,467,089	93,789,629
GrafTech International Ltd	EAF	452,024,091	1,718,778,050
Halliburton Co	HAL	31,215,863,147	72,292,261,959
Hess Corp	HES	44,767,413,328	87,495,081,516
Iberdrola SA	IBDRY.PK	73,918,226,559	1,156,347,503
Kinder Morgan Inc	KMI	38,594,843,778	57,766,250,727
Marathon Oil Corp	MRO	13,997,036,198	66,266,645,836
Marathon Petroleum Corp	MPC	61,152,738,216	119,227,571,621
Nextera Energy Inc	NEE	114,675,165,009	166,602,411,825
Occidental Petroleum Corp	OXY	53,306,460,998	160,432,473,812
ONEOK Inc	OKE	43,806,937,584	59,872,180,939
Ormat Technologies Inc	ORA	3,932,381,488	8,684,364,935
Petroleo Brasileiro SA Petrobras	PBR	107,024,450,376	67,261,311,167
Phillips 66	PSX	60,969,259,377	94,019,607,680
Piedmont Lithium Inc	PLL.O	279,687,801	4,226,392,745
Pioneer Natural Resources Co	PXD	54,945,821,828	131,472,021,181
Plug Power Inc	PLUG.O	2,124,962,138	51,354,953,820
Repsol SA	REPYY.PK	19,424,778,545	385,035,182
Schlumberger NV	SLB	69,244,425,016	128,917,215,028
Shell PLC	SHEL.L	200,870,737,269	84,954,579,344
Shoals Technologies Group Inc	SHLS.O	2,193,844,076	15,492,071,791
Solaredge Technologies Inc	SEDG.O	3,837,154,965	74,276,440,494
Subsea 7 SA	SUBCY.PK	4,453,234,111	78,948,701
Sunrun Inc	RUN.O	2,623,261,932	46,160,675,386
Targa Resources Corp	TRGP.K	21,922,782,861	33,690,458,812
TotalEnergies SE	TTE	152,570,919,978	23,356,211,156
Valero Energy Corp	VLO	47,032,890,706	123,793,226,779
Vestas Wind Systems A/S	VWDRY.PK	28,188,284,842	467,503,718
Williams Companies Inc	WMB	43,730,001,182	56,015,730,116

Team Name: Solarity Capital Partners
Strategy Name: KAL-EGY

Bibliography

PricewaterhouseCoopers (2022), *ESG-focused institutional investment seen soaring 84% to US$33.9 trillion in 2026, making up 21.5% of assets under management: PWC Report*, *PwC*.

Fama, E. (1970) Multiperiod Consumption Investment Decisions. The American economic review. Vol. 60 (1) p163-174

Kalman, R. (1960) A new approach to linear filtering and prediction problems. Journal of basic engineering. Vol 82 p35-45

Hongli, N. (2021) Correlations between crude oil and stocks prices of renewable energy and technology companies: A multiscale time-dependent analysis. Energy (Oxford). Vol 221, p119800

Hsiao, C, Weishun, L, Xinyang, W, Gaoyun, Y , Siqi L, and Ni, S (2019): The Impact of International Oil Prices on the Stock Price Fluctuations of China’s Renewable Energy Enterprises. Energies (Basel). Vol 12 (24), p4630.

Engle, R. F., Granger, C. W. J. (1987): “Co-integration and error correction: Representation, estimation, and testing. *Econometrica,* 55(2):251-276.