Natural Gas Pairs Trading without CoIntegration:
The Wild and Black Duck Strategy
Chrilly Donninger
Chief Scientist, Sibyl-Project
Sibyl-Working-Paper, Nov. 2013
http://www.godotfinance.com/

The Mallard or Wild Duck is a dabbling duck which breeds throughout the temperate and subtropical Americas, Europe, Asia, and North Africa, and has been introduced to New Zealand and Australia. Mallards usually form pairs (in October and November in the Northern hemisphere) only until the female lays eggs at the start of nesting season which is around the beginning of spring, at which time she is left by the male who joins up with other males to await the moulting period which begins in June.
(en.wikipedia.org/wiki/Mallard)

Abstract:

I introduced in a recent working-paper the Duck strategy (see [1]). The best performing variant was the Natural-Gas Duck. The strategy was considered interesting by Siddharth Batia, the chief-trader of the Sibyl-Project. When implementing the trading-code I found an embarrassing bug in the historic simulations. For longer maturities there were sometimes no trades available. These days are coded with a price of Zero (to mark the fact). But the historic simulation took them as is. Using the Zero-values creates extreme outliers. The trades exploited sometimes these artificial values. In short: The results in [1] are for Natural Gas not valid.

This working paper is a complete rework. The original CoIntegration test is in its explicit form skipped, because it adds no additional value. All interesting Natural Gas futures pairs are CoIntegrated. Instead a measure for the most interesting pair is developed. This rework is termed the Wild Duck strategy. Additionally a different model which operates on the forward-curve and the specifics of the Natural-Gas market is proposed. This new model – the Dark Duck - works slightly better than the traditional pairs approach. But the two species are closely related. The historic performance patterns are similar. They depend on the overall volatility of the natural-gas market.

Stylized Facts:

The NG Natural Gas Futures (Physical) are an outright natural gas contract for delivering 10,000 million British thermal units at the Henry Hub in Louisiana. There are other Hubs and corresponding contracts too. But the Henry Hub Futures are by far the most liquid title. The prices of other contracts differ from NG due to different transportation costs and local market specifics. For the Dec. 2013 futures the agreed amount must be delivered over the whole month. The futures expiry 3 trading days before the end of the previous month. The density of natural gas is in comparison to oil low. Hence the market is not a global one. Practically all the gas delivered at Henry Hub is produced in the US or the Gulf of Mexico. The production level is – besides interruptions due to natural catastrophes – over the year almost constant. Consumption shows a marked seasonal pattern with a peak on cold winter days. Gas is typically injected from beginning of April to end of October in depleted gas fields and withdrawn from November till March. Graphic-1 shows the details of the storage level from 1994 to 2010. The storage cycle is also reflected in the forward-curve. Graphic-2 shows the forward-curve for the years 2010-2013 in January. The peaks are at about 12 and 24 months Graphic-3 is the same for June. Now the peak is at about 6 and 18 months.
Graphic-1: Inventory Level of US Natural-Gas.
(source: Fazzio [2], p. 36).

Graphic-2: Forward-Curve of NG Futures in January
Graphic-3: Forward-Curve of NG Futures in June

Graphic-4: NG Nearest Future Price (orange) and Volatility (blue) 2007-01-01 till 2013-11-27
The storage level and the storage costs are an important price factor. The U.S. Energy Information Administration (EIA) publishes every Thursday at 10:30 a.m. EST the Weekly Natural Gas Storage Report. (http://ir.eia.gov/ngs/ngs.html)

This report (actually the deviations from the forecasts) forms a regular shock to the natural gas market. Thursday is hence the weekday with the highest volatility. This fact is exploited by the forward-curve model aka the Dark Duck.

Graphic-4 shows the development of the gas price since 2007 till recent. Like for many other commodities the price exploded before the 2008 crash. It reached a historic maximum at 2008-6-30 with 13.57. The current price at 2013-11-27 is 3.86. The price of crude oil (ticker CL) dropped in the 2008 crisis too, but it recovered in the meantime somewhat (see Graphic-5). The different behavior can be attributed to the new conveyance technique of hydraulic fracturing aka fracking. (en.wikipedia.org/wiki/Hydraulic_fracturing)

Pairs-trading is immune to the long-term trend. But there is nevertheless an influence of the declining price trend. Volatility has also drifted down (blue bottom-chart in Graphic-4). The best conditions for a pairs-trade are a strong mean-reversion effect and a high volatility of the spread (see [3]). The declining overall volatility reduces therefore the habitat of the ducks.

Graphic-6 shows the volume of NG at 2013-11-20. Nothing special happened on this day. The pattern was – besides the rollover of NGZ13 to NGF14 - in the previous and following days practically the same. The first 5 futures are very actively traded, but there is a reasonable amount of volume up to April 2015.
<table>
<thead>
<tr>
<th>Month</th>
<th>Volume</th>
<th>TRADE TYPE DETAIL</th>
<th>Deliveries</th>
<th>Open Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GLOBEX</td>
<td>Block Trades</td>
<td>EFP</td>
<td>EFR</td>
</tr>
<tr>
<td>DEC 13</td>
<td>120,235</td>
<td>3,166</td>
<td>2,587</td>
<td>45,358</td>
</tr>
<tr>
<td>JUN 14</td>
<td>91,604</td>
<td>3,870</td>
<td>5,044</td>
<td>357,207</td>
</tr>
<tr>
<td>FEB 14</td>
<td>24,160</td>
<td>24,962</td>
<td>251</td>
<td>102,449</td>
</tr>
<tr>
<td>MAR 14</td>
<td>28,346</td>
<td>20,641</td>
<td>107</td>
<td>167,751</td>
</tr>
<tr>
<td>APR 14</td>
<td>16,364</td>
<td>17,280</td>
<td>0</td>
<td>135,467</td>
</tr>
<tr>
<td>MAY 14</td>
<td>4,495</td>
<td>4,495</td>
<td>0</td>
<td>55,196</td>
</tr>
<tr>
<td>JUN 14</td>
<td>3,088</td>
<td>3,068</td>
<td>0</td>
<td>27,137</td>
</tr>
<tr>
<td>JUL 14</td>
<td>1,793</td>
<td>1,793</td>
<td>0</td>
<td>27,335</td>
</tr>
<tr>
<td>AGL 14</td>
<td>608</td>
<td>768</td>
<td>0</td>
<td>23,143</td>
</tr>
<tr>
<td>SEP 14</td>
<td>699</td>
<td>799</td>
<td>0</td>
<td>25,545</td>
</tr>
<tr>
<td>OCT 14</td>
<td>5,131</td>
<td>5,421</td>
<td>0</td>
<td>86,766</td>
</tr>
<tr>
<td>NOV 14</td>
<td>1,623</td>
<td>1,623</td>
<td>0</td>
<td>27,456</td>
</tr>
<tr>
<td>DEC 14</td>
<td>2,242</td>
<td>2,242</td>
<td>0</td>
<td>50,873</td>
</tr>
<tr>
<td>JAN 15</td>
<td>3,245</td>
<td>3,245</td>
<td>0</td>
<td>45,022</td>
</tr>
<tr>
<td>FEB 15</td>
<td>84</td>
<td>64</td>
<td>0</td>
<td>7,099</td>
</tr>
<tr>
<td>MAR 15</td>
<td>373</td>
<td>373</td>
<td>0</td>
<td>11,375</td>
</tr>
<tr>
<td>APR 15</td>
<td>410</td>
<td>410</td>
<td>0</td>
<td>16,356</td>
</tr>
</tbody>
</table>

Graphic-6: Volume of NG at 2013-11-20
(Source CBOE)
The work in [1] was originally started with ETFs in mind. The test for CoIntegration was therefore the essential selection criterion. This work was then extended to Natural-Gas futures. But the test was according the following recalculation with corrected input data redundant. Practically all futures with reasonable liquidity are CoIntegrated. There was – in contrast to the ETFs – no information in this filter. The spread was defined as the deviation from the regression line. The slope of this line is in general not 1. But trading was based – due to restrictions in the Sibyl-fund – on an equal volume. One goes 2 futures short and long. After correcting for missing data the strategy was also not really profitable anymore.

The current strategy is based on the theory developed in [3]. The potential profitability of a spread depends on the mean reversion factor and the volatility.

\[
Z(t) = X(t) - Y(t) \quad (1)
\]

\[
dZ(t) = a + b*Z(t-1) \quad (2)
\]

with

\[
dZ(t) = Z(t)-Z(t-1) \quad (3)
\]

The spread Z is in (1) simply defined as the difference of the 2 futures prices. Mean-reversion is the slope b in equation (2). b must be negative. A smaller – more negative – b is better. Volatility is defined as the standard-deviation of Z. The performance is considerable improved if all these measures are calculated with robust statistics. The linear regression is estimated with the Theil-Sen estimator. This estimator – which was already invented before classical OLS – has excellent statistical properties and is easy to calculate. It's only drawback is the quadratic run-time.

For generating trading signals one has to normalize the spread. For this step the median was used as the robust centrality measure. The MAD (Median-Absolute-Deviation from the Median) as the dispersion. The MAD was multiplied with 1.4826. This is the scale factor for a normal-distribution. 1.4826 times the MAD equals in this case the standard deviation. The scaled MAD is for the NG spread considerable smaller than the Standard-Deviation. These means that the distribution has fat tails. As in [1] the window-length for calculating the statistics was 126 trading-days. If one (or both) of the futures has on a given trading-day no entry in the time-series the day was skipped and the series was extended longer. The actual window can hence be 127 or 128 trading days (2 missing values was the maximum).

On a given day t one calculates (1) and (2) for all pairs of the next 6 Futures. These futures have according to Graphic-6 sufficient liquidity. The minimum maturity must be – in the best setting – 28 calendar days. All the pairs are sorted according to

\[
Z\text{score}(i,j) = \text{abs}(b)\times\text{volatility}(Z) \quad (4)
\]

As already noted above, volatility is 1.4826*MAD.

A high volatility combined with a very negative b is the best combination. Actually the criterion is conflicting. Futures with a smaller difference in maturity have a stronger mean-reversion but also lower volatility. The Zscore() is a simple measure which tries to find the optimal combination.

The pair i,j with the highest Zscore is the potential trading pair. Additionally the normalized spread at time t must be larger than the 1.0 (or smaller than -1.0). Additional filters are:

\[
b(t) \leq -0.03 \text{ and volatility}(t) > 0.03. \quad (5)
\]
One closes a pair if:

a) Mean-Reversion has happened. There are (at least) two different definitions of mean-reversion. Either the spread has closed or the spread has swung to the other direction. In [1] I used the first definition. But for the current calculation the swing to the other direction performed better.
b) The pair is open since 10 weeks (70 calendar days). In [1] a position was kept up to 90 days. The 70 days rule performs slightly better, but 70 or 90 days does not matter too much.
c) The maturity of the shorter future is 7 or less calendar days. One does not want to deliver at the Henry Hub. This is the typical rollover-distance of NG.
d) The normalized spread is above 3.0 (or below -3.0 if the spread was initially negative). This is a stop-loss.

One can calculate the normalized spread for rule a) and d) with the original median and MAD when the position was opened. Or one could update the two values each day and normalizes the spread with the updated values. Keeping the original values performed slightly better. But the difference is due to the robust nature of the median and the MAD only minor.

The performance index starts with 500.000$ at 2009-11-18. The final value is – after trading costs - 548.180$. The trading costs sum up to about 10.000$. The maximum relative drawdown is with 6.33% also relative modest. One fights especially in the last year but also during 2011 with the low overall volatility.

One has to increase the leverage to a volume of 16 per pair to get over the last 4 years the same performance than the SPY (Graphic-8). The max. relative drawdown increases in this case to hefty 31.3%. The maximum relative drawdown of the SPY was in the same time 18.6%. But it was higher.
during the 2008 crash.

Graphic-8: Wild Duck (orange) and SPY (yellow) from 2009-11-28 till 2013-11-27

The Dark Duck:

*Dark Ducks are close relatives to the Wild Duck or Mallard. Dark and Wild Ducks interbreed regularly. Some authorities even consider the Black Duck to be a subspecies of the Mallard, not a separate species at all.*

(en.wikipedia.org/wiki/American_Black_Duck)

S. Ohana developed in [5] a simple model for commodity forward curves which suites especially well for the natural-gas market. One models the returns and not the price-series.

\[ R_i(t) = R_m(t) + \exp\left(-\frac{K}{252.0} (T_i - T_1)\right) \cdot (R_1(t) - R_m(t)) \quad (6). \]

\( R_i(t) \) is the return of the ith future at time t.
\( R_m(t) \) is the return of the future with the longest considered maturity.
\( R_1(t) \) is the return of the nearest future.
\( K \) is the exponential factor.
\( T_i \) is the maturity in years of the ith future.
\( T_1 \) is the maturity in years of the nearest future.

The idea of this model is simple. The nearest Future 1 reacts strongest to shocks like the storage report or natural disasters. The future m with the longest maturity shows a damped reaction. The returns of the
futures in between follow an exponential declining curve. The factor \( K \) is determined by minimizing the squared differences of the model and the measured returns. M. Zolotko reports in [6] \( K=2.86 \) for natural-gas and \( 3.10 \) for heating oil. The model explains according these calculations 93% of the return variations.

\( K \) depends according my results on the time-horizon of the return-calculation. If one calculates daily-returns from Wednesday to Thursday \( K \) is \( 3.02 \). If one calculates weekly or biweekly returns from Thursday to Thursday the factor reduces to \( 2.7 \). The calculation was applied to the first 9 futures (\( m=9 \)). The most nearby future is rolled over 7 days before maturity. The calculations are for the time range 2009-11-24 to 2013-11-24.

The Black Duck uses \( K=2.7 \). But the performance of the trading strategy is insensitive to slight changes. A \( K \) of 2.86 like in [6] has only a minor influence.

For the Black Duck one calculates each Thursday at the close the forward curve according to (6). One selects the pair with the largest deviation from this curve. One can use the daily, weekly or biweekly return (one can of course use any time-window, but these are the interesting cases). Thursday is due to the Weekly Storage Report the logical trading day. The Dark Duck does not work at other days.

Note: Restricting the Wild Duck to Thursdays does not increase the performance.

For the best combination the spread (the difference of the pair to the model curve) must be for the biweekly return at least 1.75%. The position is closed if the pair has swung to the other direction or the position is for 28 days open. There is an additional stop loss of 5.0%.

The performance of the Dark Duck is with a final value of 556.100$ slightly better than its wild relative. But the curve is much smoother. Besides a single spike at 2010-10-29 the relative drawdown is consistently below 2%. One has to increase the leverage to a volume of 15 per pair to get over the last 4 years the same performance than the SPY (Graphic-10). The max. relative drawdown is still lower than for the SPY. One has also to check if such a leverage is from the margin-requirements feasible. There
are sometime up to 4 pairs or 120 futures open. But the Dark Duck can not overcome the problem of low volatility regimes. There is a slight decline in the last year. The decline disappears if one excludes trading-costs. But they are a fact of trading life.

Graphic-10: Dark Duck (orange) and SPY (yellow) from 2009-11-28 till 2013-11-27

Graphic-11 shows the calculation for the daily-returns from Wednesday to Thursday. One basically trades the shocks of the last storage report. One has to decrease the minimum spread to 0.75%. For a spread of 1.5% there are no trades at all. The volume per pair is set to 15 like in Graphic-10. The position is never hold longer than a week. So there is a maximum of one pair open. As can be seen in the bottom-chart of Graphic-11 one has most of the time no open positions at all. There are only a few storage-report shocks which are large enough. The final performance is 545.750$. The picture is without trading-costs much nicer. The final value is 588.950$ with a rather smooth performance curve (Graphic-12). But these are in Austrian dialect “Haettiwari” (If I would have, I could have) wins.

I tried the Dark Duck also for heating oil. The forward-curve model works for heating oil too. But the volatility is too low. It's about half of the volatility of natural-gas.
Graphic-11: Dark Duck with daily returns from 2009-11-28 till 2013-11-27

Graphic-12: Dark Duck with daily returns and zero trading costs from 2009-11-28 till 2013-11-27
Conclusion:

The Dark Duck has an edge over the Wild Duck. It exploits the over- or under reaction of different maturities to (storage-report) shocks on Thursdays. The strategy is up to my knowledge an innovation in the pairs-trading universe. But both ducks suffer from the relative low volatility of natural gas in recent time (it is in comparison to other commodities still relative high). According the theoretical results in [3] there is little chance to find considerable better trading rules which circumvent this problem.

An alternative is to try the spread between related commodities. An interesting pair should be crude-and heating oil (CL v. HO). From the visual inspecting (see Graphic-13) there are some periods with comovement. But the gap between HO and CL is over the last 4 years more than 30%. The two commodities do not stay close together in the long run. Trading this pair is hence a different problem which will addressed in a forthcoming study.

Graphic-13: Heating Oil (orange) and Crude Oil (yellow) from 2009-10-27 to 2013-11-27.
References: