Application of Design of Experiments to an Automated Trading System

Ronald Schoenberg, Ph.D.

Trading Desk Strategies, LLC

ronschoenberg@optionbots.com

www.optionbots.com

A Design of Experiments (DOE) method was applied to an automated technical trading system described in Connors and Alvarez [1]. A cubic response curve was fit to results from 56 plus 6 trials generated by an I-Optimal five factor design and an optimal "sweet spot" was found that improved the outcome by 388% over a result using settings suggested in their book.

Automated trading systems, especially based on technical indicators, are becoming popular (for example, www.cool-trade.com). These systems are highly dependent on settings such as moving average periods which are usually determined in some unsystematic fashion. Some are set by tradition such as trading long only when an index is above a 200 day moving average. Others are found by brute force back testing, generally by what is called "a single factor design" testing, that is, each setting is tested on its own. All of these methods fail to capture the interconnectedness of the settings, nor are they likely to find something optimal. These methods usually sift through many attempts before finding one that generates any profit at all.

One systematic method that avoids these problems is the grid search method. This is not generally used for a good reason. For a five factor model, the type we'll be discussing here, the number of computer runs covering the parameter space could run into the millions. For some types of trading systems this could take weeks or even months for a single investigation.

The Design of Experiments method is a way to solve all these problems. First, a set of points in the parameter space are selected based on a statistical criterion, 56 points plus 6 additional ones for the experiment discussed here. Runs are conducted at each of those set of points and an outcome, profit/loss in this case, are recorded. Then a polynomial function is fit to the outcome data, the trial points as independent variables, and the outcome as dependent. We are now able to express the outcome, the profit/loss, as a simple calculation throughout the entire parameter space called a

response surface. Finally, one of several types of hill-climbing methods can be applied to the response surface to find the "sweet spot", the optimal parameter settings.

For our problem here, a single run on 20 years of historical data takes .25 seconds to run on a PC. A grid search on the entire parameter space with resolution of one decimal place would require 43,000,000 runs which at a quarter of a second comes to 17.7 weeks. On the other hand for the DOE, the 62 points in the parameter space takes 12.84 seconds, and the hill-climbing for the sweet spot was another 45 seconds, a total of one minute as opposed to 17.7 weeks.

DOE makes it possible to find optimal settings quickly and easily that take into account the relationships among the parameter settings in producing the outcome. This dramatic increase in efficiency also has another important benefit – it makes it easy to test modifications to the structure of the trading system. For example, should it be a simple average, a moving average, or an exponential moving average? A conclusion to that question can be found directly by repeating the trial runs and the hill-climbing for each method, in three minutes for the problem we're looking at here.

The polynomial model used for the problem in the paper is a cubic. The cubic response surface has a greater ability to fit observed outcomes but it also contains a lot of terms relative to the number of factors which then requires many trial runs. For a problem with run times on the order of a second or two, this isn't a difficulty. For trading systems with longer run times, for example on the order of an hour or so, a simpler polynomial model, such as a quadratic or an interactive (a polynomial without the power terms) can be used which require fewer trial runs. A comprehensive grid search for one hour trial runs is completely out of the question, but a quadratic model with 15 one hour trial runs is feasible.

A Technical Trading System – The VIX RSI Strategy

On page 96 Connors and Alvarez [1] describe a trading system for the SPY:

1. The SPY is above its 200-period moving average.
2. The 2-period RSI of the VIX is greater than 90
3. Today's VIX open is greater than yesterday's close.
4. The 2-period RSI of the SPY is below 30.
5. Buy on the close.
6. Exit when the 2-period RSI of the SPY closes above 65.

The motivation for this system is to use the VIX to detect a weak and oversold market. In a back test on three and a half years of historical data they found 92 signals with an average hold under five days with a frequency of success of 79.35%.

A DOE Analysis

The trading system above contains five factors:

1. Period of the SPY moving average
2. RSI setting of the VIX
3. RSI setting of the SPY
4. Exit setting of the SPY RSI
5. Period of the RSI

To define the parameter space we choose minimums and maximums for the first four factors:

1. [50 – 250]
2. [80 – 100]
3. [15 – 45]
4. [35 – 65]

The fifth factor is a discrete integer which will be defined to take on the values of 2, 3, 4, 5, and 6.   The first factor is also a discrete integer but its large interval allows us to treat it as continuous. Any value we find will be truncated to an integer without serious effect to the investigation.

Using the Gosset [2] computer program we generate parameter settings for 56 trial runs (see Table 1). Daily data for the SPY and the VIX from January 29, 1993 through October 17th 2003 was selected for the trial runs. A computer program written in the GAUSS programming language [3] was developed for the simulation of the trading system given the parameters.

To produce a more realistic result, each simulation started with an account of $500,000, paid a $20 transaction cost per trade, and a bid/asked spread was also included. The results can be found in the last column of Table 1.

For hill-climbing, the *Sqpsolvemt* program in the GAUSS Run-Time Library was used. *Sqpsolvemt* uses a nonlinear sequential quadratic programming method allowing us to constrain the search to the parameter space defined by the minima and maxima.   It doesn't handle integers however and thus a separate hill-climb will be conducted within each value of Factor5 and the sweet spot will be the best result among those separate runs.

The 56 trials for the cubic polynomial model provide no degrees of freedom and an initial run found sweet spots with large error deviations. To gain degrees of freedom and a better fit to the observations,

a center point was added, and the sweet spots from the initial run were also added for a total of 62 trials.   The results for each value of Factor 5 are

Predicted	Factor 1	Factor2	Factor3	Factor4	Factor5
439654.88	250.0000	89.3615	15.0000	95.0000	2
639913.19	250.0000	80.0000	45.0000	95.0000	3
534355.83	141.2619	80.0000	45.0000	95.0000	4
365909.82	128.1201	80.0000	45.0000	35.0000	5
125703.03	102.7956	96.3250	15.0000	63.0219	6

Rather than rely on the back testing for our choice of sweet spot, we will forward test each of them in another data set, the SPY and VIX from October 17th 2003 through October 12th, 2009. The results for these runs are

Profit/Loss	Annualized Return	Sharpe Ratio	Maximum Drawdown	Factor5
43930.35	1.47%	0.2687	-20468.70	2
101995.79	3.27%	2.0457	-24108.84	3
131039.61	3.82%	0.9835	-24006.78	4
11610.98	0.37%	0.5874	-20825.91	5
7320.80	0.23%	0.9562	-10596.30	6

Opinions might reasonably differ here. One the one hand, the RSI period of 4 clearly wins the profit contest. On the other, the greater theoretical prediction for a period of 3 might suggest that it would be superior in the long run.   None of these sweet spots generate much of an annualized rate of return but to be fair the forward test period includes the Meltdown.

The forward test for the original settings in Connors and Alvarez [1] are

Profit/Loss	Annualized Return	Sharpe Ratio	Maximum Drawdown
28329.22	0.93%	0.5819	-20434.68

The Design of Experiments method has succeeded in finding settings that produce from three and half to four and half times greater profit, and more than three times the annualized return.

For the back test the period 3 sweet spot produced the following result,

Profit/Loss	Annualized Return	Sharpe Ratio	Maximum Drawdown
764341.35	10.01%	1.5362	-69778.50

And the back test for the original settings in Connors and Alvarez [1] are

Profit/Loss	Annualized Return	Sharpe Ratio	Maximum Drawdown
156503.42	2.78%	0.5893	-39988.32

Again in the back test the DOE sweet spot produces more than four times the profit and more than three times the annualized return.

Findings

The regression results are presented in Table 2. There aren't enough degrees of freedom in the model for very much statistical power. However, some tentative findings are possible. First, the results don't seem to be very sensitive to the period of the SPY moving average. In fact, a period of 200 for the sweet spot for the RSI period of 3 produces results in the forward test which are comparable:

Factor 5	Profit	Annualized Return	Sharpe Ratio	Maximum Drawdown
3	131987.39	4%	1.6738	-25140.78
4	130783.56	3.96%	1.2546	-24001.11

While the DOE results suggest the longer the period the better, a 250 day moving average is a full year's worth of data and longer periods may be impractical in some situations. It does appear that the traditional period of 200, while still long, should do quite well in practice.

Connors and Alvarez [1] strongly hold to an RSI period of 2 in contrast to the traditional period of 14. The DOE results first support their contention that a period of 14 is too long. On the other hand the results also suggest that a modification to a period of 3 or 4 could produce significantly greater profits.

The regression analysis also suggests that the RSI period (Factor5) and the VIX RSI setting (Factor2) are interrelated. A future study might produce better results that incorporated separate RSI periods for the VIX versus the SPY.

Table 1

trial	Factor 1	Factor 2	Factor 3	Factor4	Factor5	Profit/Loss
1	97	99	35.96	73.00	6	36963.78
2	250	92	41.30	95.00	3	442391.29
3	50	91	19.79	76.65	3	152935.57
4	185	80	22.95	52.25	6	36588.3
5	71	100	27.41	43.89	4	-2605.81
6	57	80	45.00	35.00	2	-102887.56
7	250	100	45.00	95.00	6	31972.4
8	164	100	42.21	36.16	2	33763
9	161	90	43.92	78.53	2	82352.07
10	250	100	45.00	95.00	2	151022.55
11	50	80	15.00	95.00	2	96021.45
12	184	92	22.70	35.00	2	-63011.38
13	53	82	42.33	59.02	5	64807.41
14	250	100	15.00	37.29	3	11373.27
15	50	82	44.96	95.00	3	551780.73
16	152	97	15.00	59.90	5	33534.78
17	250	85	22.46	56.48	4	135522.19
18	235	85	36.54	78.85	6	-499936.44
19	250	96	34.10	61.45	2	140999.78
20	187	80	15.00	80.02	3	268324.76
21	237	100	45.00	71.97	4	4321.72
22	136	100	18.07	35.00	6	0
23	59	95	31.53	95.00	2	51648.44
24	103	84	35.58	35.00	6	59420.6
25	69	82	26.76	95.00	5	162501.49
26	243	92	15.00	78.90	2	255643.5
27	250	80	15.00	95.00	6	24722.76
28	91	80	31.09	75.54	2	11636.23
29	176	100	37.78	95.00	4	6470.07
30	153	80	45.00	90.90	5	316046.77
31	50	87	34.87	42.42	2	-111804.83
32	250	96	15.62	49.12	6	0
33	176	93	21.89	93.05	6	31972.4
34	250	80	45.00	35.00	6	78705.39
35	192	80	38.38	43.16	3	167178.97
36	95	84	15.00	78.55	6	0
37	50	99	15.00	35.00	2	-48289.34

38	106	84	15.00	49.51	2	-28863.42
39	50	100	45.00	65.19	2	43134.47
40	50	100	45.00	35.00	6	0
41	50	80	19.21	35.00	4	29610.96
42	250	80	45.00	78.26	2	143786.78
43	90	93	45.00	36.94	3	46744.44
44	50	80	45.00	95.00	6	-499958.08
45	209	83	26.16	95.00	2	309655.71
46	180	91	45.00	49.67	6	21889.9
47	148	100	20.01	72.07	2	78568.21
48	50	100	15.00	95.00	6	0
49	250	87	45.00	35.00	2	-56868.06
50	50	93	21.90	46.99	6	0
51	250	100	21.25	87.74	5	0
52	110	95	15.00	95.00	3	368372.77
53	250	80	15.00	35.00	2	-9787.38
54	195	85	15.00	35.00	5	15092.02
55	230	97	35.75	35.34	5	30531
56	63	92	45.00	91.07	5	122935.42
57	150	90	30.00	65.00	4	123083.67

Table 2:

term	coefficient	t-statistic
Constant	-22.635	-0.128
B(1)	-2.852	-0.039
B(2)	61.490	0.913
B(3)	-35.665	-0.510
B(4)	-14.282	-0.210
B(5)	16.728	0.109
B(1)^2	11.897	0.356
B(1)*B(2)	0.144	0.006
B(1)*B(3)	-2.576	-0.125
B(1)*B(4)	-0.540	-0.025
B(1)*B(5)	-3.369	-0.092
B(2)^2	-4.123	-0.119
B(2)*B(3)	-10.342	-0.459
B(2)*B(4)	-21.883	-0.890
B(2)*B(5)	-33.861	-1.008
B(3)^2	3.898	0.109
B(3)*B(4)	1.006	0.046
B(3)*B(5)	18.065	0.531
B(4)^2	-5.698	-0.161
B(4)*B(5)	13.498	0.390
B(5)^2	-1.215	-0.029
B(1)^3	7.722	0.261
B(1)^2*B(2)	2.411	0.140
B(1)^2*B(3)	-0.076	-0.005
B(1)^2*B(4)	0.783	0.045
B(1)^2*B(5)	-3.491	-0.449
B(1)*B(2)^2	2.707	0.139
B(1)*B(2)*B(3)	2.752	0.242
B(1)*B(2)*B(4)	-0.209	-0.017
B(1)*B(2)*B(5)	-0.150	-0.027
B(1)*B(3)^2	1.107	0.054
B(1)*B(3)*B(4)	1.432	0.133
B(1)*B(3)*B(5)	0.262	0.050
B(1)*B(4)^2	9.460	0.513
B(1)*B(4)*B(5)	0.160	0.029
B(1)*B(5)^2	0.270	0.059
B(2)^3	-7.231	-0.253
B(2)^2*B(3)	0.009	0.001

B(2)^2*B(4)	-10.495	-0.558
B(2)^2*B(5)	0.523	0.062
B(2)*B(3)^2	-1.285	-0.065
B(2)*B(3)*B(4)	-1.397	-0.119
B(2)*B(3)*B(5)	2.479	0.464
B(2)*B(4)^2	-5.851	-0.323
B(2)*B(4)*B(5)	5.222	0.924
B(2)*B(5)^2	4.354	1.043
B(3)^3	5.148	0.183
B(3)^2*B(4)	5.112	0.263
B(3)^2*B(5)	-0.043	-0.005
B(3)*B(4)^2	5.378	0.297
B(3)*B(4)*B(5)	-1.224	-0.230
B(3)*B(5)^2	-2.429	-0.580
B(4)^3	11.868	0.421
B(4)^2*B(5)	2.364	0.279
B(4)*B(5)^2	-2.287	-0.530
B(5)^3	-0.193	-0.056

References

[1] Connors, L. and C. Alvarez, *Short Term Trading Strategies That Work*, TradingMarkets Publishing Group, 2009.

[2] R. H. Hardin and N. J. A. Sloane, "A New Approach to the Construction of Optimal Designs"*, Journal of Statistical Planning and Inference,* vol. 37, 1993, pp. 339-369

[3] Aptech Systems, Inc., http://www.aptech.com.