

Polynomial Forecasting of Securities' Prices

Overview

In this study we develop a methodology for evaluating forecasting methods applicable to trading securities. The examples cited are all forecasts of future prices made by extrapolating a polynomial curve that has been fitted to historical data from the security in question. The methods here should be extensible to other forecasting techniques.

As discussed in the prior Summa Management study "Plots Specification for Cause and Effect Analysis for Securities Price Movements", to be valid, forecasting techniques must be evaluated in the context with which they will be used. This includes how the forecast will be interpreted and how the trading system utilizing it works.

Incorporating these concepts, this methodology entails the following steps when applied to polynomial forecasting:

1. Financial performance is measured while back-testing over a period which includes a minimum of one full economic cycle
2. Use regression analysis to fit a polynomial to split-adjusted prices for the target security for the last N trading days prior to the test day¹
3. Plug the test day and the next eligible trading day into the daily fitted curve formula and determine if the projection is for the price to rise or fall between the two
4. If the price is projected to rise, the study's funds are invested in a long position for the security; if projected to fall, the study's funds are invested in a short position^{2,3}
5. Each flip of position (to go from buy long to sell short, or the reverse) is assumed to cost 0.25%⁴
6. The effective investment yield (Compound Annual Growth Rate, CAGR) is computed for the previous M months and plotted with the security price history, and the CAGR is computed for the full term of the study as a single figure-of-merit
7. Parameters of the study are swept across a range to find the conditions which result in the peak return (CAGR)
8. At the peak return conditions, graph how the accuracy of the forecast fared short versus longer term and on rising versus falling markets, particularly for big market swings (where it likely that the most money can be made if the projection is correct)

¹ During back-testing, the test day is sequentially swept from the start to the finish of the back-testing interval.

² Note that many securities have an inverse Exchange Traded Fund (ETF) that shorts the targeted security. For the S&P 500, one such fund is ticker "[SH](#)".

³ If the prediction is correct, regardless of whether it is up or down, money will be made (neglecting spread/commissions).

⁴ This encompasses the buy/sell spread plus the commission. For this to be realistic, the securities being traded must be sufficiently liquid, their trading volumes sufficiently high relative to the trades of this study, and the size of each trade sufficiently large to amortize fixed fees per trade.

The earlier study cited above showed graphically examples of polynomial forecasting of various orders (including order zero, which is a straight line), in its discussion on Curve Fitting. Additionally, it showed the effect of using different intervals (number of trading days prior to the forecast) to train the fitting function.

The analysis runs over a 5 year period (2006 through 2010) and varies parameters across 4 different dimensions, as follows:

1. Order [1 (= straight line) to 6 (= 5-humped curve)] of predictive polynomial
2. Training time (6 to 20 trading days)
3. Trading cycle (once per trading day to once every 6 trading days)
4. Type of security (broad stock index, blue-chip, volatile small-cap stock)

Study Summary

Security (all stock-based)	Buy and hold 5 year annualized return	Best 5 year forecasting annualized return	Poly-nomial order	Training interval (trading days)	Trade interval (trading days)
Index (S&P 500)	-0.3%	+12.3%	1	25	3
Index (DJI)	+1.0%	+12.2%	2	50	7
Index (NASDAQ composite)	+2.4%	+14.3%	6	45	7
Blue chip (IBM)	+12.0%	+18.6%	6	40	7
Volatile small-cap (Vitesse)	-33.6%	+131.9%	2	10	2

Analysis

The table above shows that the right choice of polynomial prediction parameters consistently beat the market, but that the optimal order of the polynomial, the training interval, and the trade interval were not consistent across the securities tested.

Figures 1 through 5 are more detailed summaries of the back-testing on the various securities. The table on the top of these figures shows how the CAGR varies as a function of the order of forecasting polynomial and the amount of history for which it is optimized (i.e. the training period). What is evident, particularly when this table is plotted in the adjacent 3-D view, is that the CAGR does not behave in a nice smooth fashion, monotonically rising to a peak in the polynomial-order/training-day space. This implies that the solution to obtain peak CAGR is fickle, and not robust. A minor deviation in either dimension, but especially the polynomial order, results in a major change in the CAGR. In the case of the S&P 500, moving 1 notch up in polynomial order and two notches down in training period moves one from the maximum CAGR to the minimum CAGR. Also note that multiple (though lesser) CAGR peaks exist.

The trade frequency table, in the middle of each summary, shows a strong dependency on how frequently one trades. Look particularly at the DJI, where changing from trading

every 6 days to trading every 7 days moved the annual return from near zero to over 12%. Also note that these algorithms predicted the correct market direction just barely better than a flip of a coin (up to 56% correct), and many times they were worse (down to 41% correct). This also shows that this prediction method is not systemically robust, and what we are seeing is noise and coincidence.

The chart at the bottom of each of these summaries shows what points in time this trading system made money, and when it lost money; all in comparison to what the price of the underlying security was doing. For the S&P 500, there were two regions where the system consistently made money (as evidenced by the 20 day CAGR plots all being above 0% (out of the red-shaded region). These are July 2006 through February 2007, and October 2008 through December 2008, with one gap. These regions are magnified in Figures 16 and 17. The S&P was slowly but consistently trending up in the first, and falling rapidly in the second, except for an uptick that corresponded to the previously mentioned gap in positive returns (because the trading system was holding a short position).

The next set of graphs, Figures 6 & 7 for the S&P, and corresponding pairs for the other securities evaluated, show how the prediction algorithm performed versus how far out in the “future” the prediction was made, and how well the prediction algorithm performed in getting the market movement direction correct for various slopes of market movement (% change at the next eligible trading day). As would be expected, most market movements between trading days were small (between $-\frac{1}{2}\%$ and $+\frac{1}{2}\%$), while a few were $\pm 5\%$ or greater. The character of the curves did not change significantly as one looked forward from 1 to 6 days out, although a higher percentage of the changes were greater (no longer in the $-\frac{1}{2}\%$ to $+\frac{1}{2}\%$ bin) the further out one went.

What is curious is that down-market movements were predicted by these polynomials much more accurately than up-market movements. This is shown clearly on the bottom chart in Figure 7, which looks at how the predictions worked on big (5% or more per trading interval) changes. I attribute this to the character of the market over this interval. When it fell, it fell fast and with conviction; when it rose, it rose slowly and more uncertainly.

Areas for Further Work

Potential other forecasting methods that can be evaluated for their performance within a trading system are enumerated on the website <http://www.ipredict.it/ForecastingMethods.aspx> and listed below. These all extrapolate future datapoints algorithmically from historical datapoints, as opposed to using external data, such as news items, to form the basis of the forecast.

It should be useful to determine if any of these can consistently produce good trading returns, and understand the conditions under which promising algorithms perform well and where they don't. This additional insight should allow for even better returns.

Classical Algorithms

1. Simple Moving Average
2. Geometric Moving Average
3. Triangular Moving Average
4. Parabolic Moving Average
5. Double Moving Average
6. Exponential Moving Average
7. Double Exponential Moving Average
8. Holt's Double Exponential
9. Triple Exponential Moving Average
10. Holt's Triple Exponential
11. Adaptive Response Rate Exponential Smoothing
12. Holt Winter's Additive
13. Holt Winter's Multiplicative
14. Holt Winter's Modified Multiple Seasonalities
15. Additive Decomposition
16. Multiplicative Decomposition
17. Sparse Series Croston's Exponential
18. Linear Trend / Regression

Curve and Bayesian Model Fitting

19. Linear Trend And Additive Seasonality
20. Linear Trend And Multiplicative Seasonality
21. Linear Trend And Multiple Seasonalities
22. Polynomial
23. Logarithmic
24. Exponential

Wavelet Smoothing and Forecasting

25. Frequency Identification
26. Haar De-noising
27. Daubechies Linear De-noising
28. Daubechies Exponential De-noising
29. Wavelet Forecasting

Additional Functions and Algorithms

30. Random Number Generation
31. Wavelet Transforms
32. Fourier Transforms
33. Hurst Exponent

Advanced Algorithms

34. Fractal Projection
35. Active Moving Average

Kernal Smoothing

36. Gaussian Kernel Smoothing
37. Hilbert Kernel Smoothing
38. Triangle Kernel Smoothing
39. Epanechnikov Kernel Smoothing
40. Quartic Kernel Smoothing
41. Triweight Kernel Smoothing
42. Cosine Kernel Smoothing
43. Savitsky-Golay Smoothing
44. Spline Smoothing

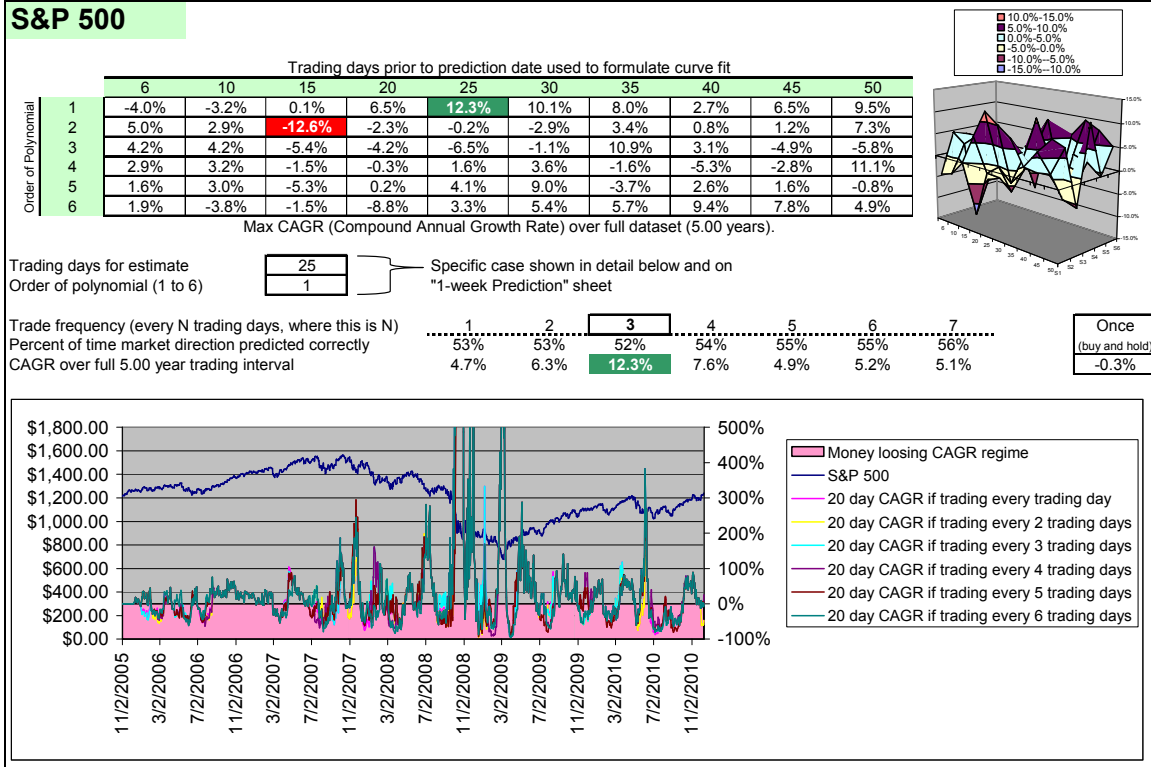


Figure 1

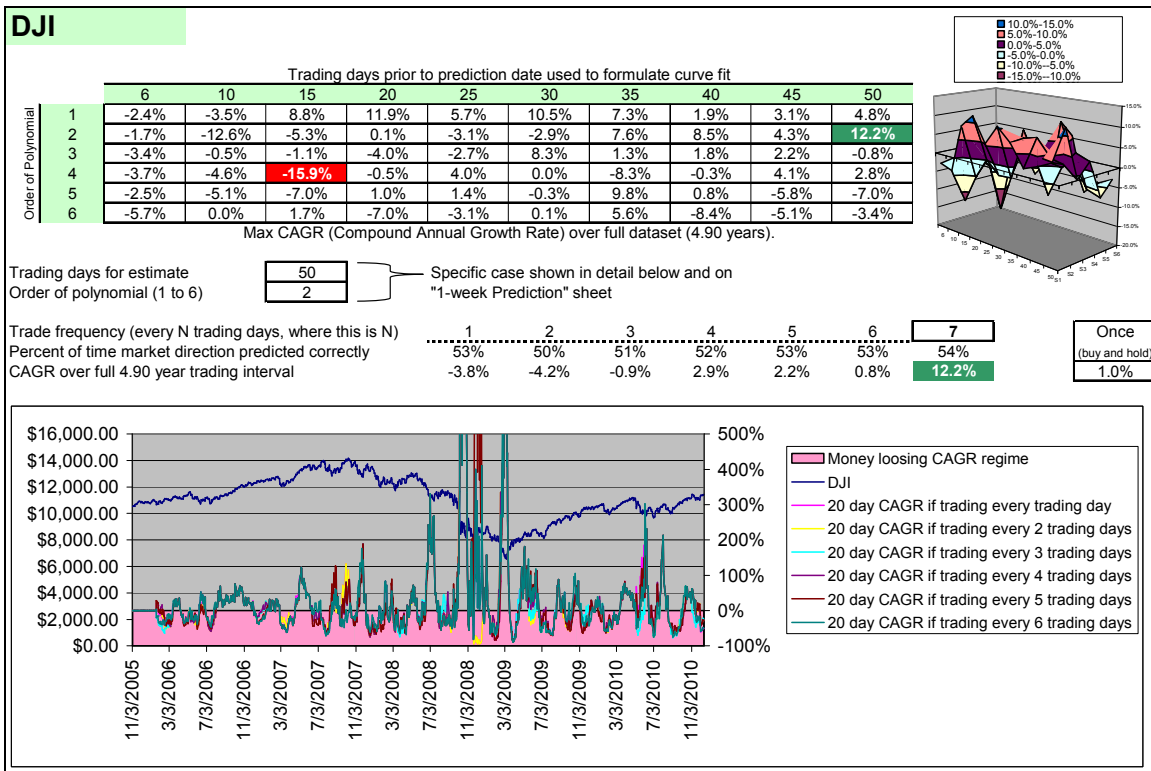


Figure 2

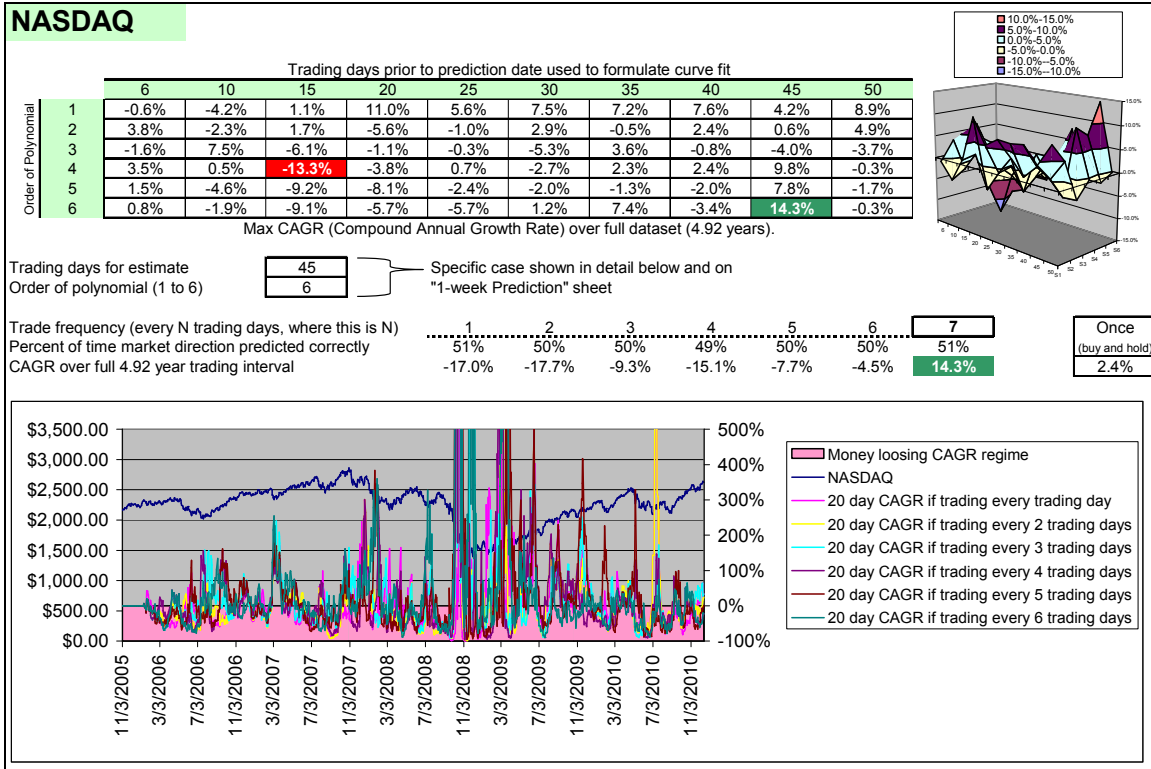


Figure 3

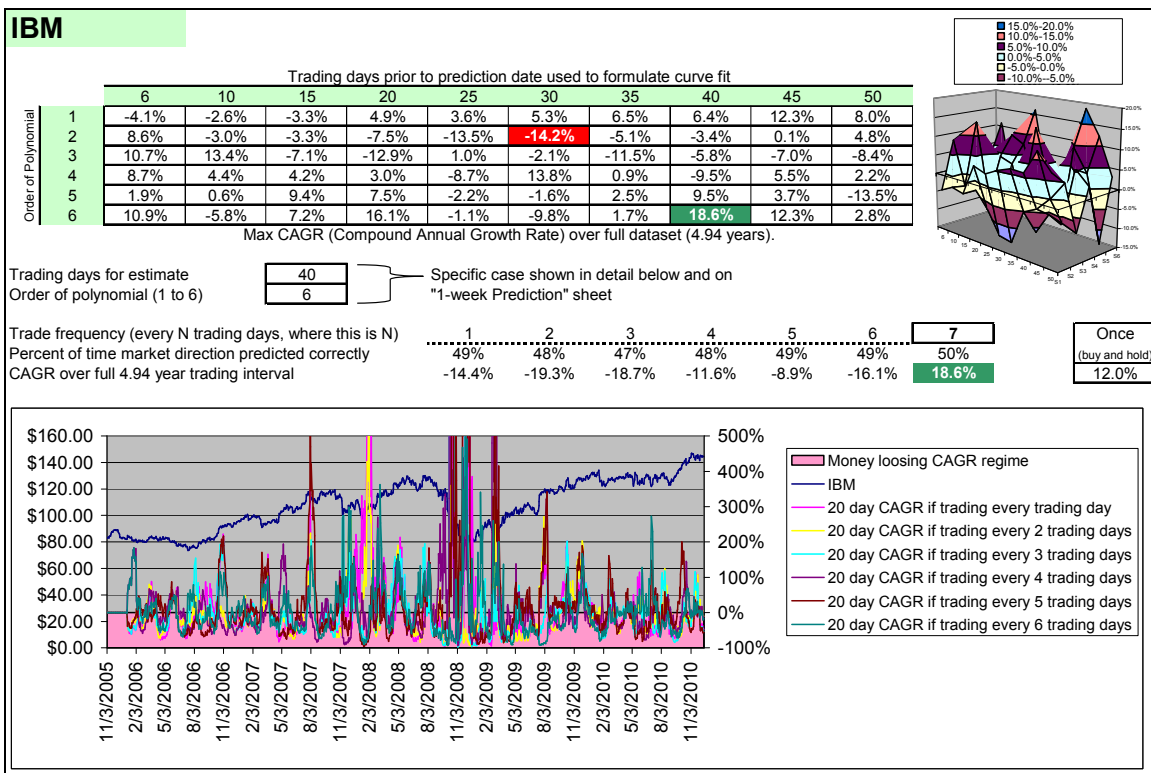


Figure 4

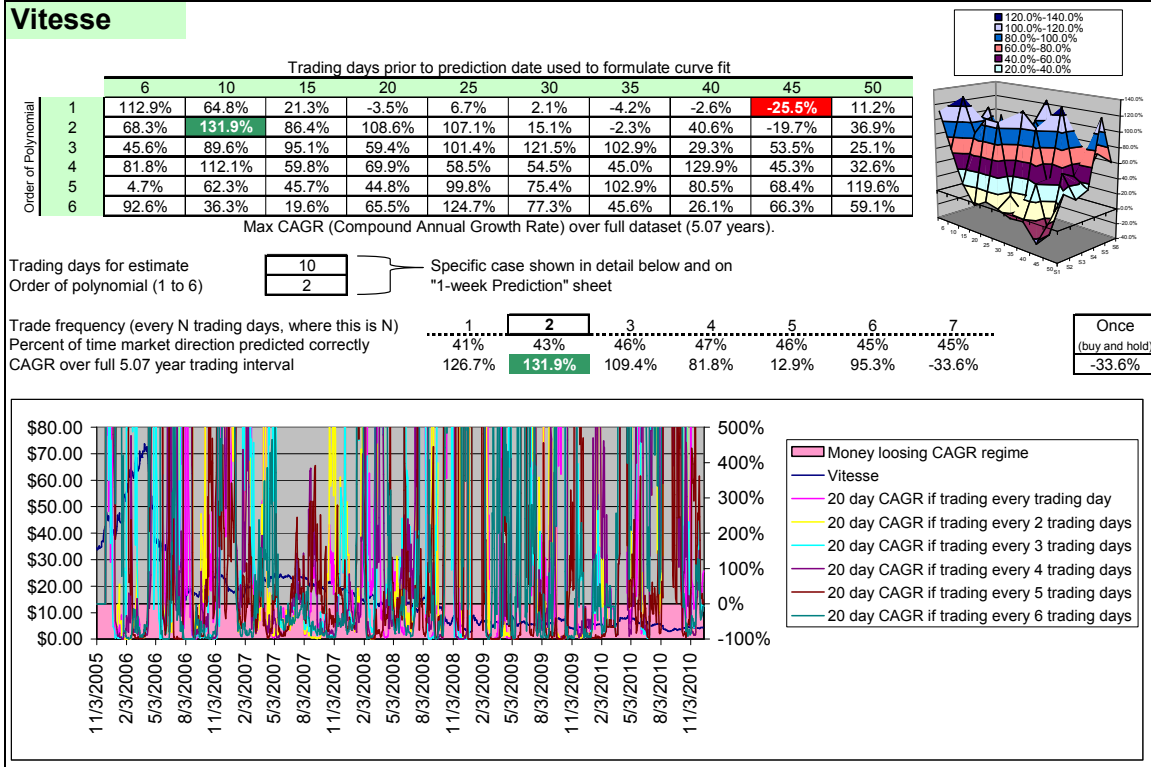


Figure 5

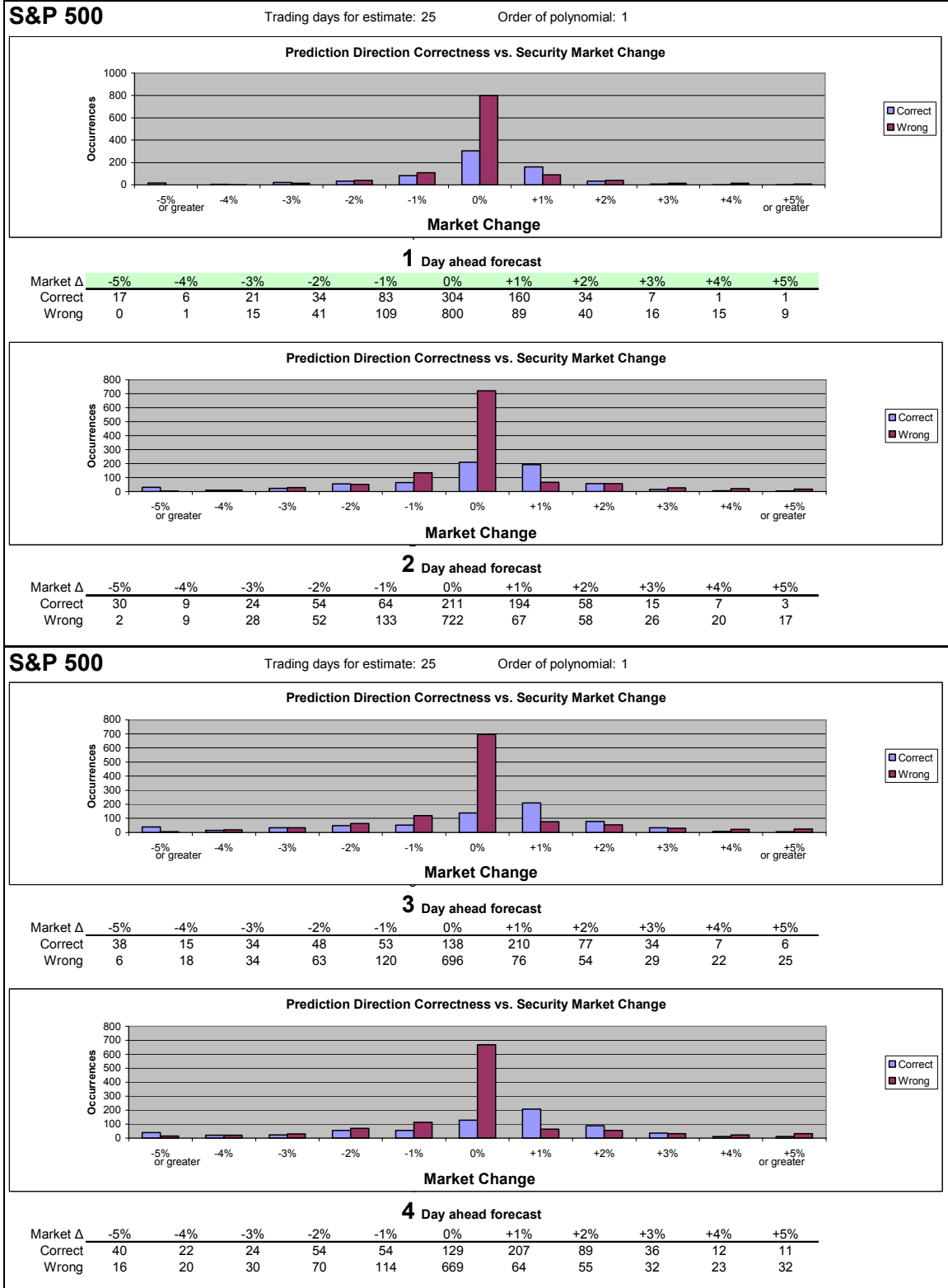


Figure 6

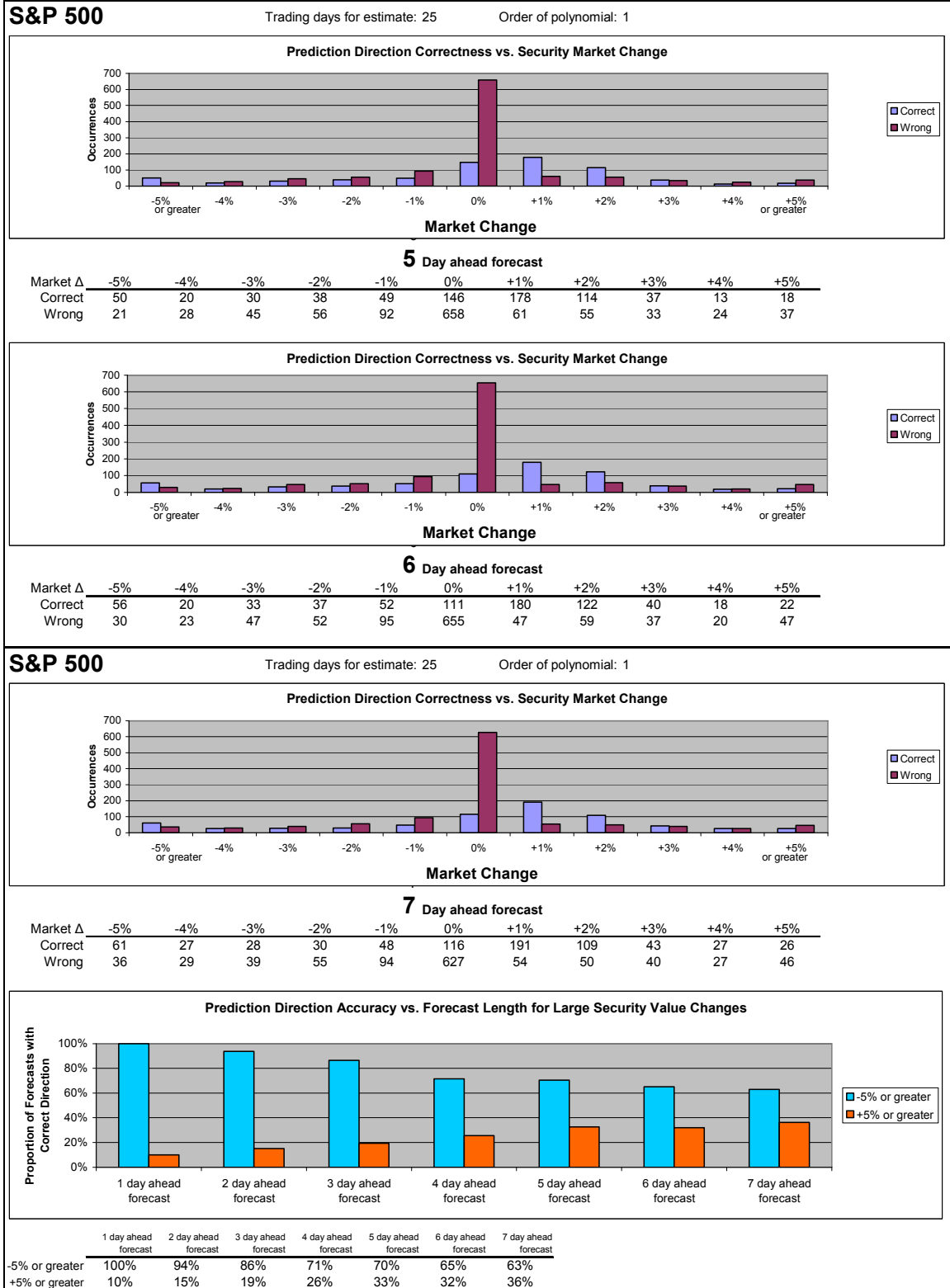


Figure 7

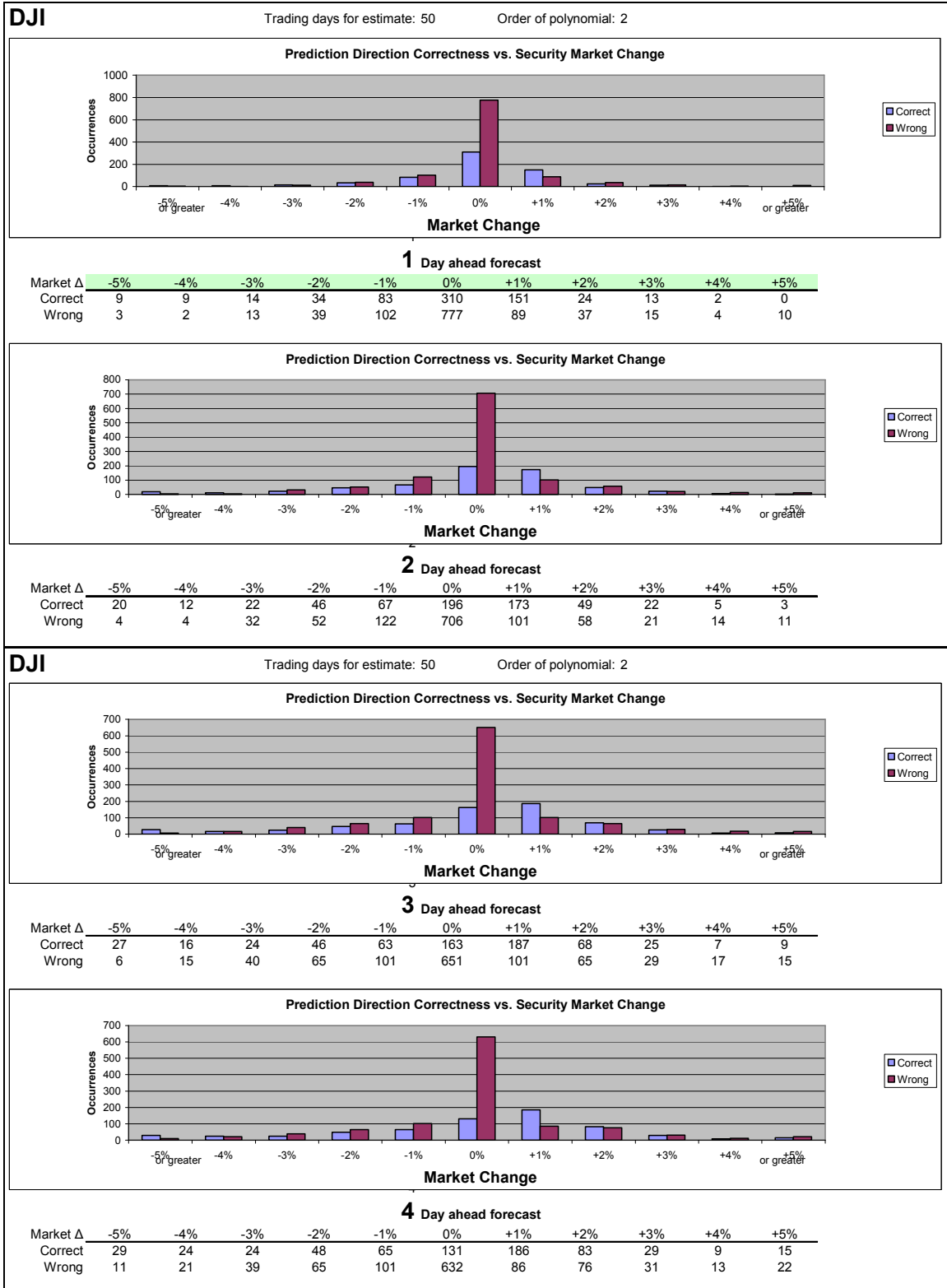


Figure 8

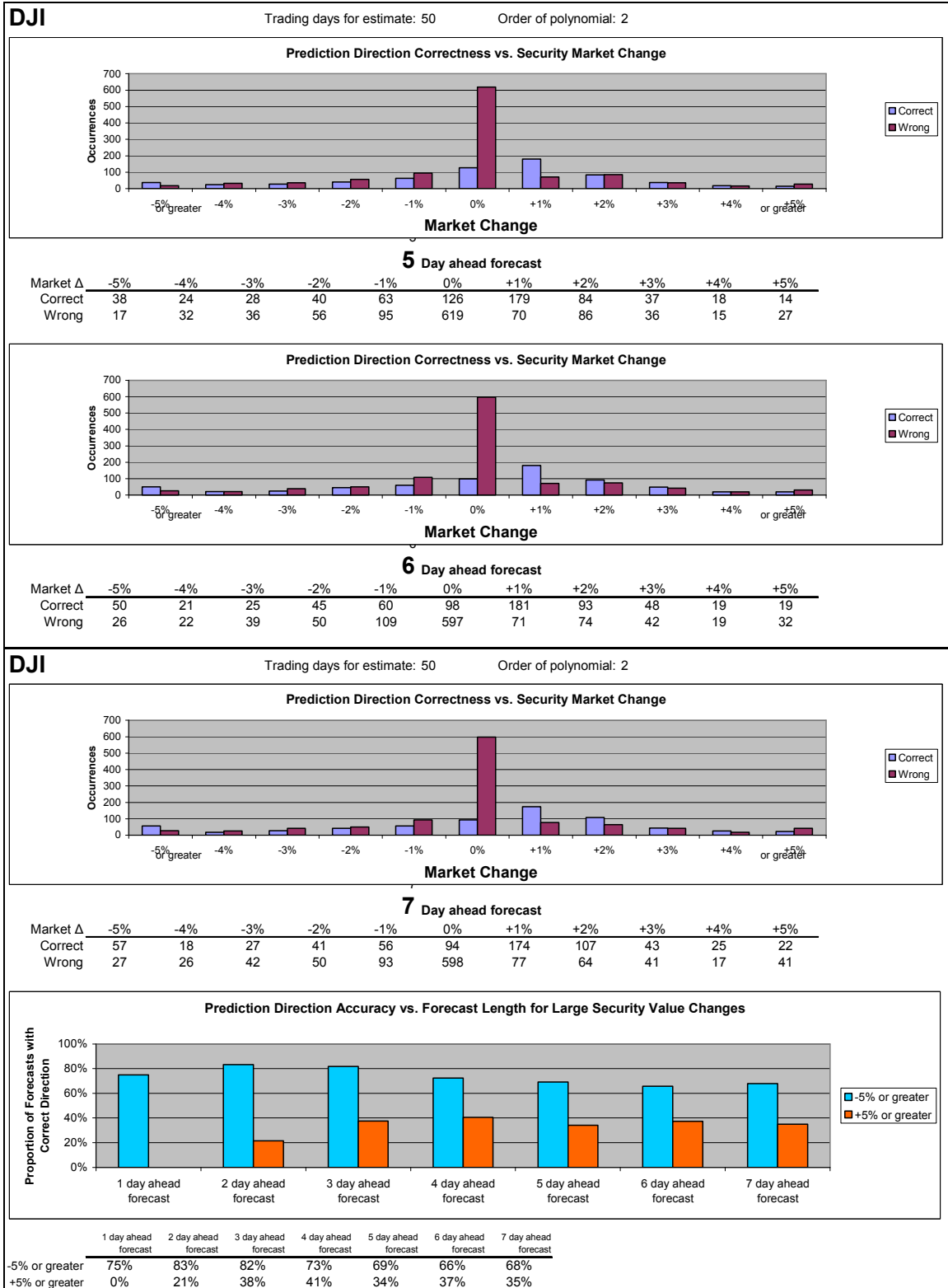


Figure 9

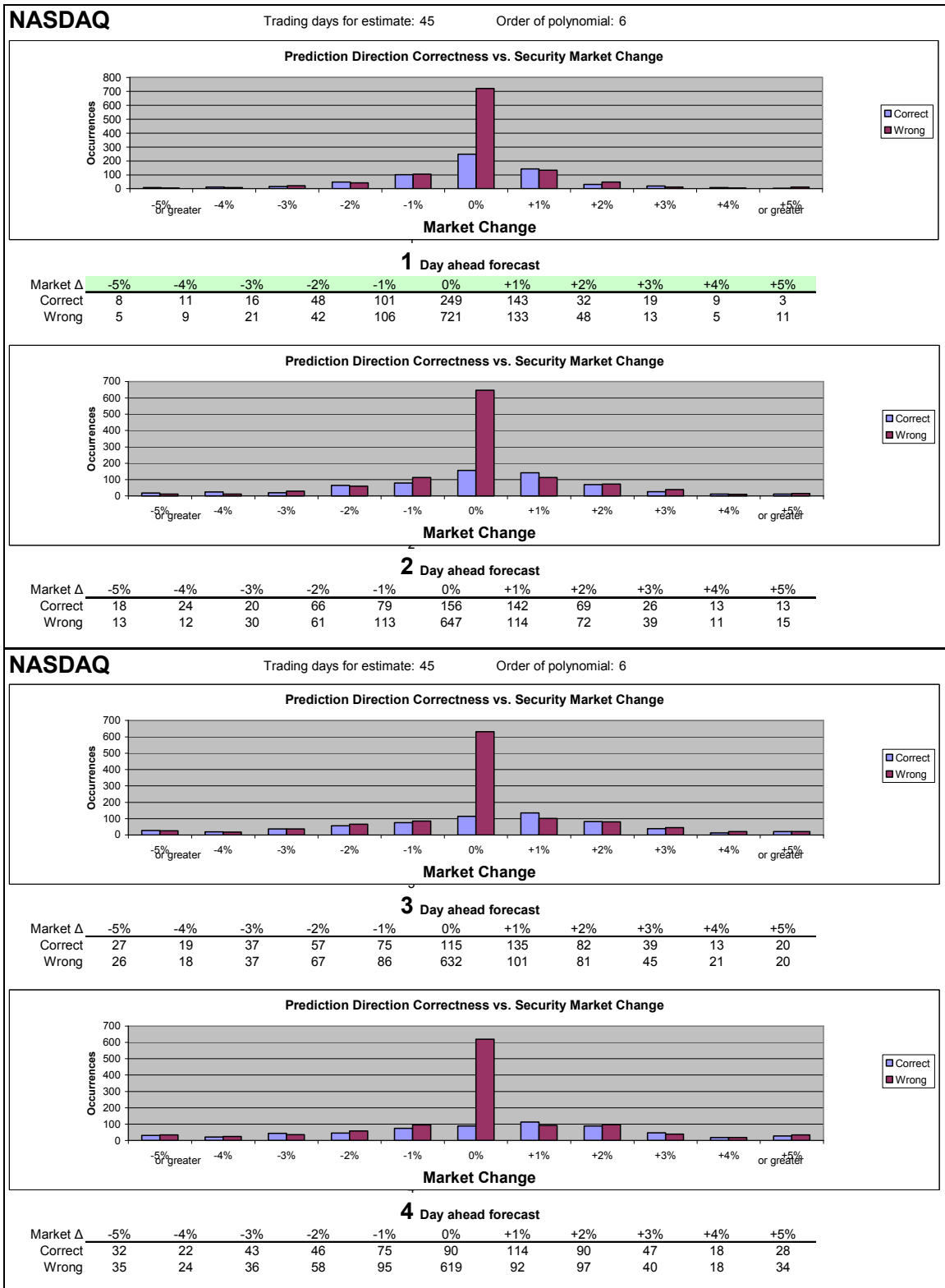


Figure 10

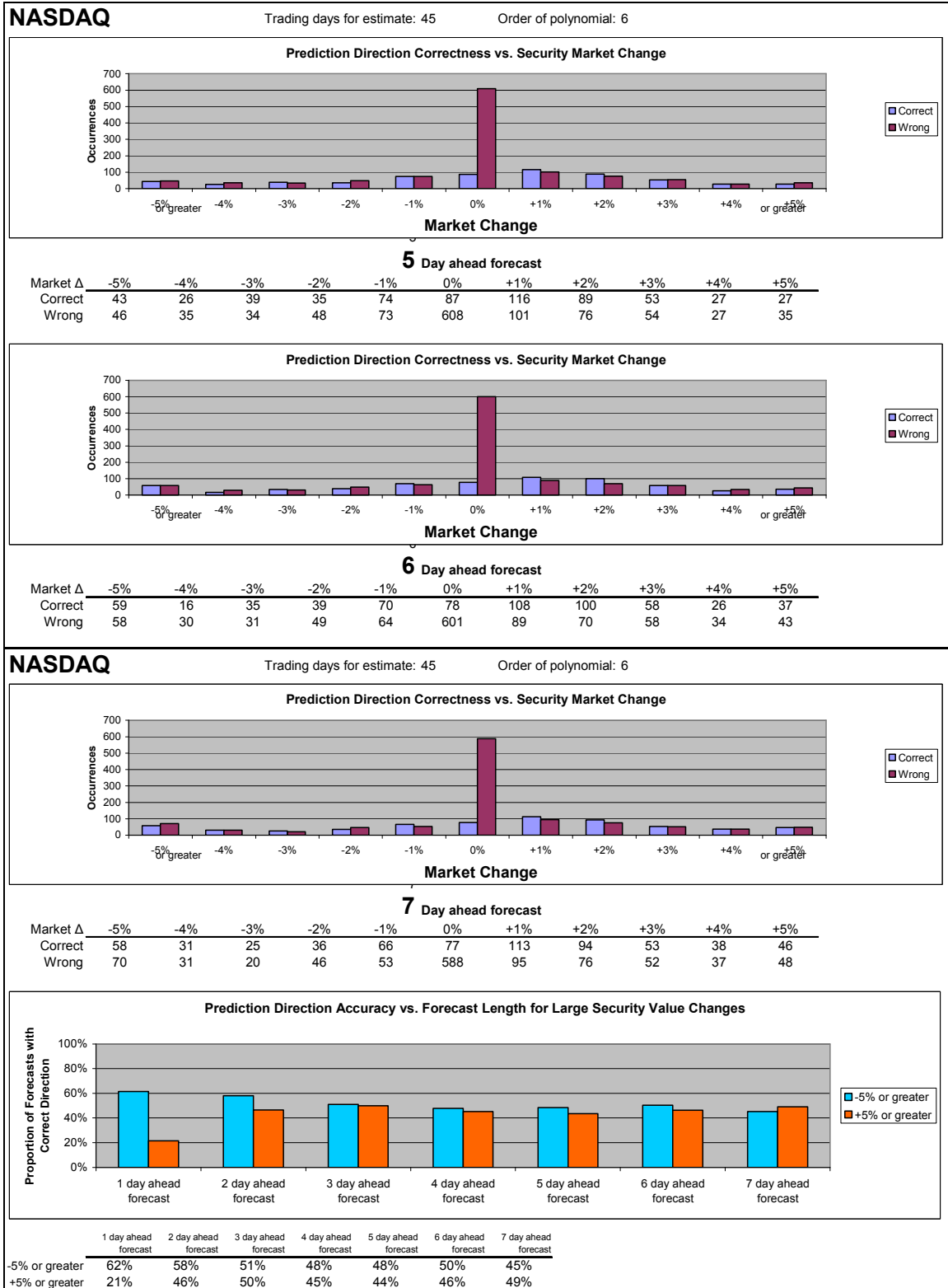


Figure 11

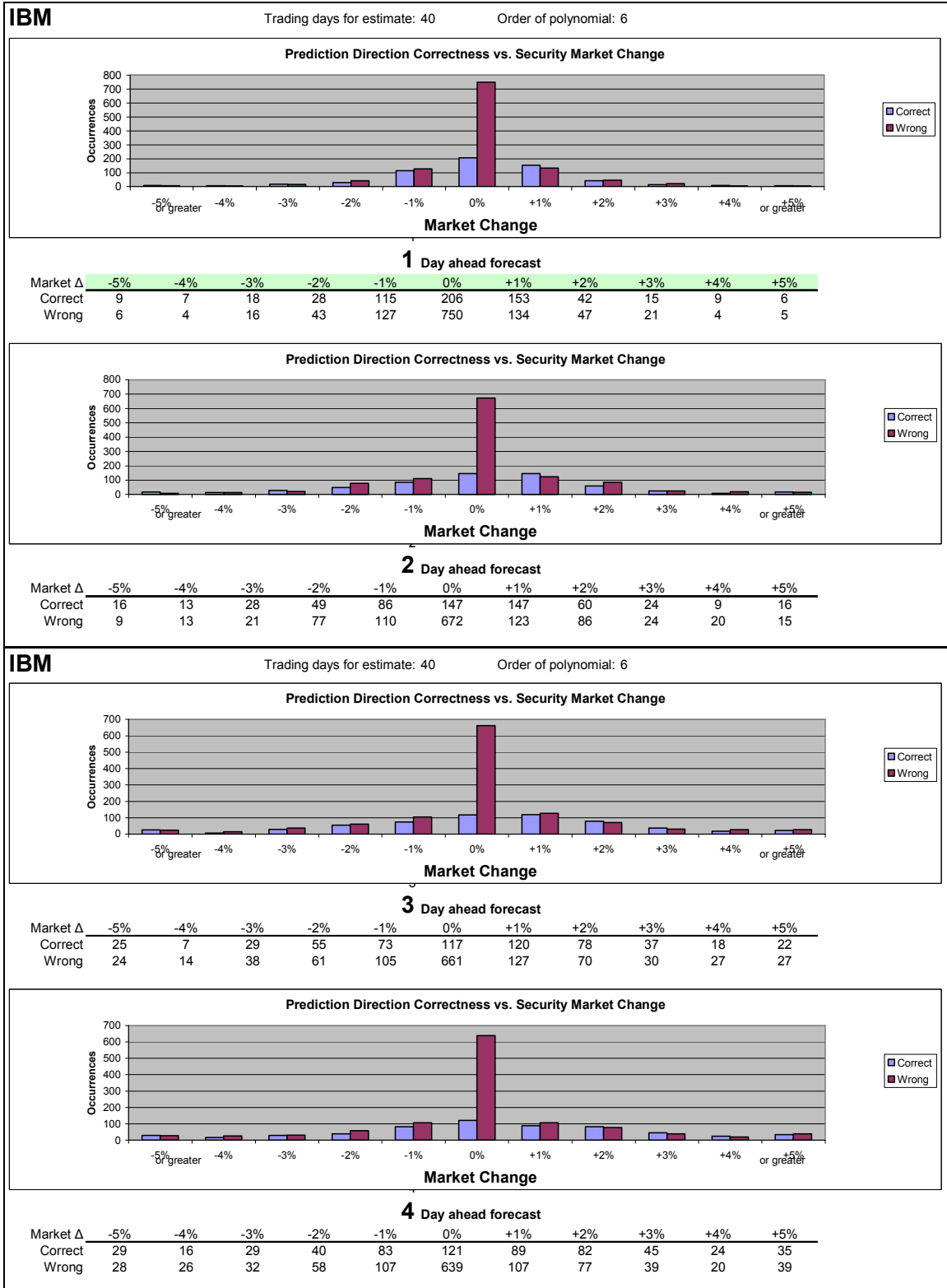


Figure 12

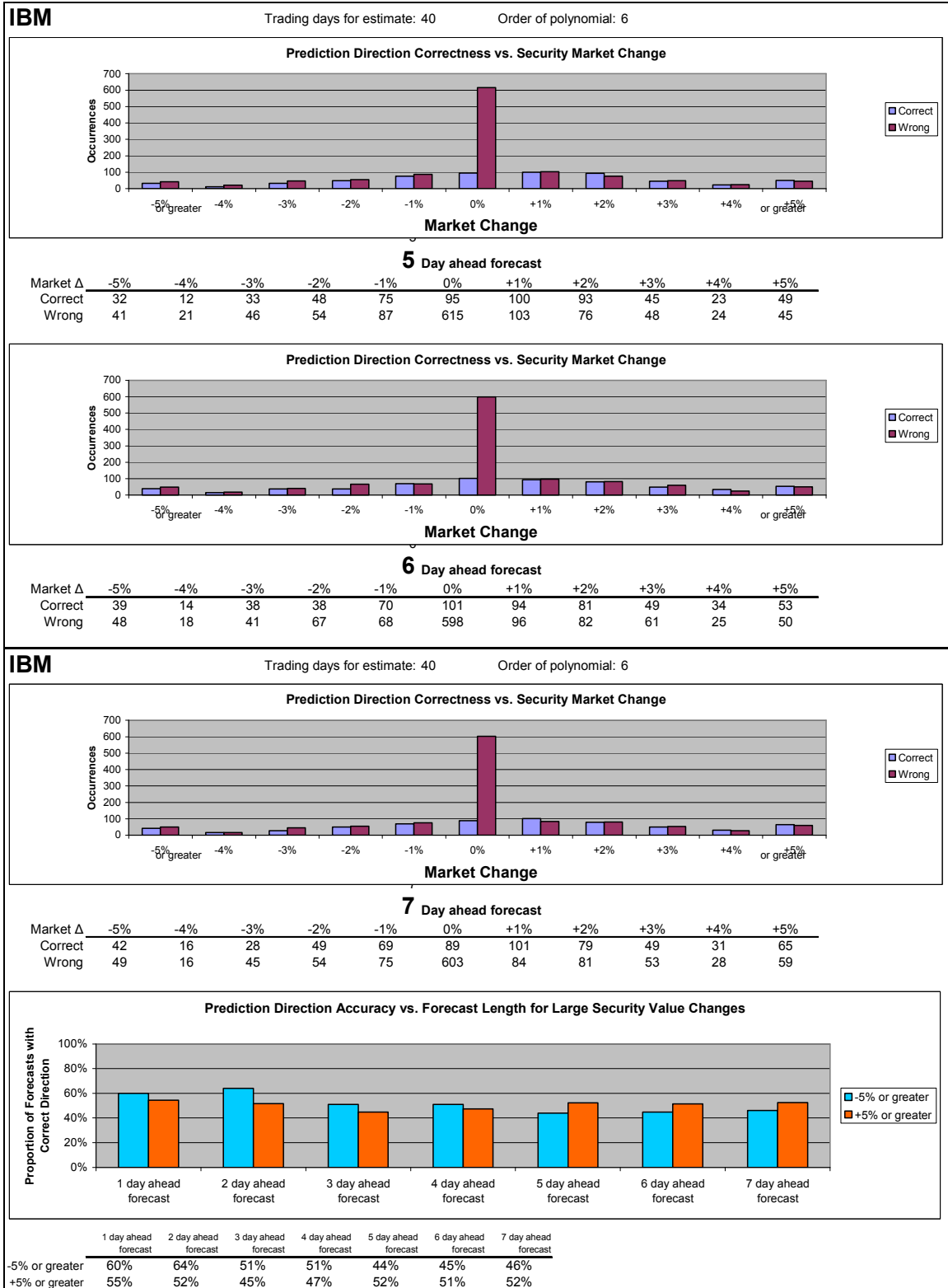


Figure 13

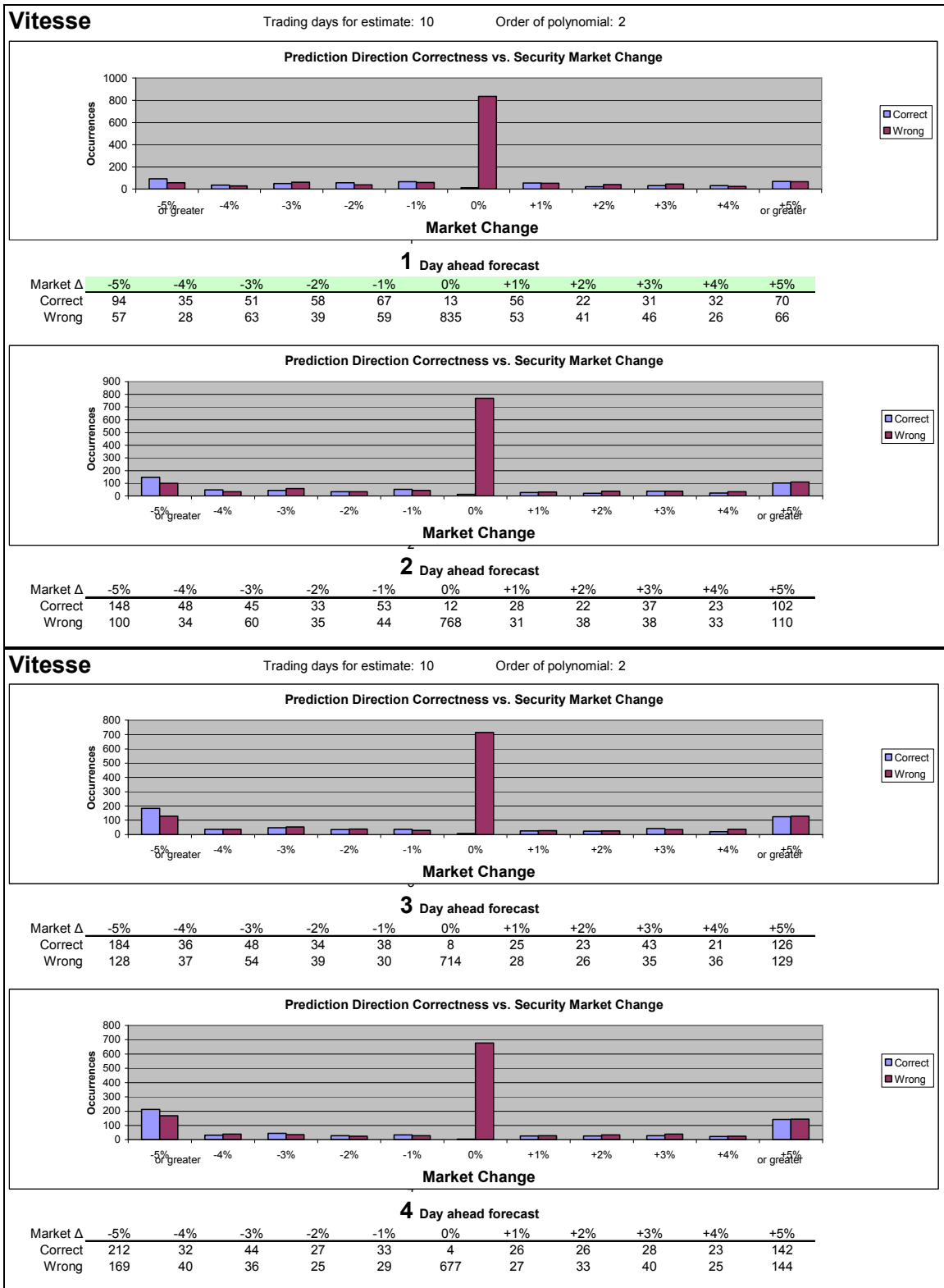


Figure 14

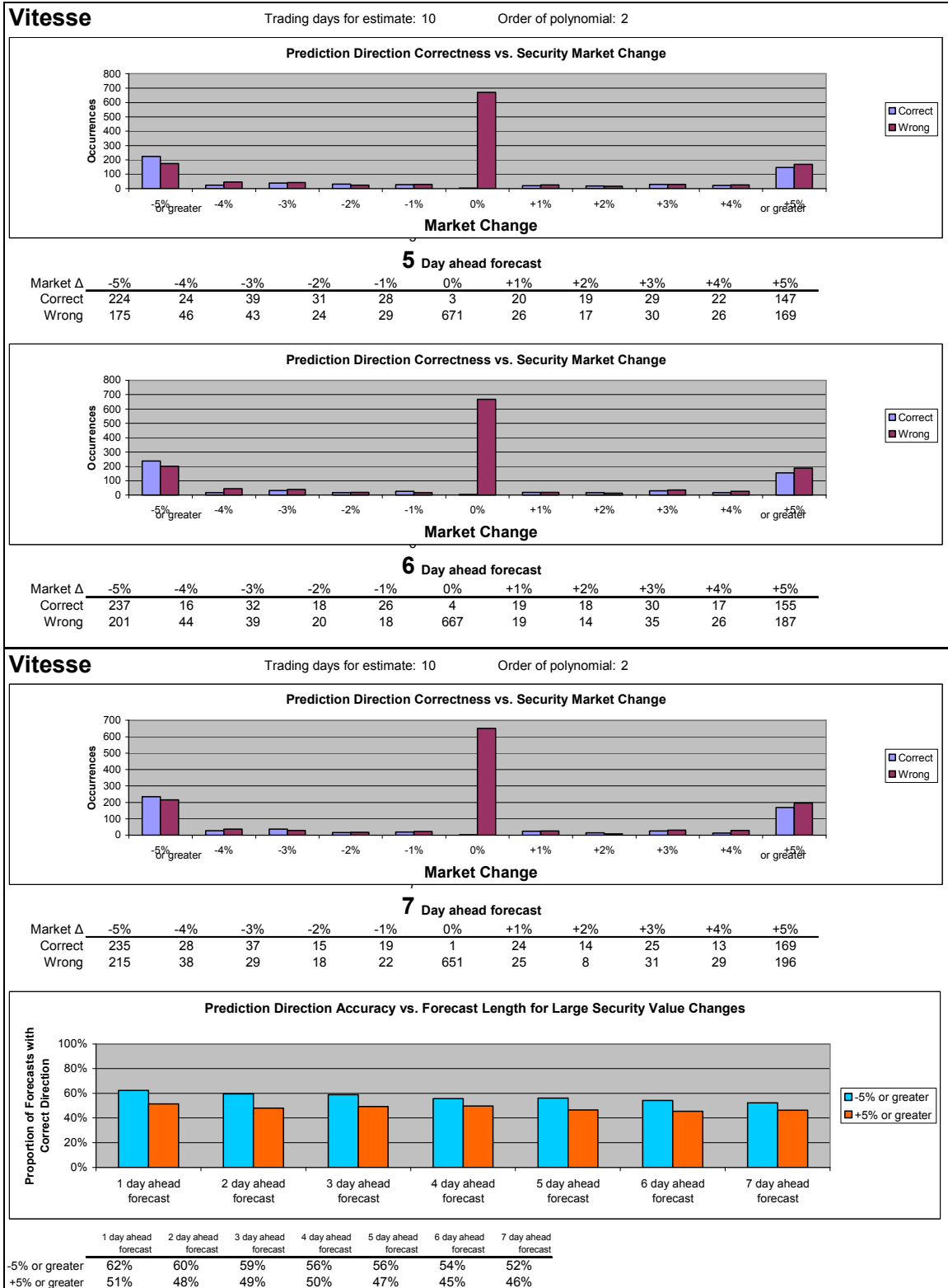


Figure 15

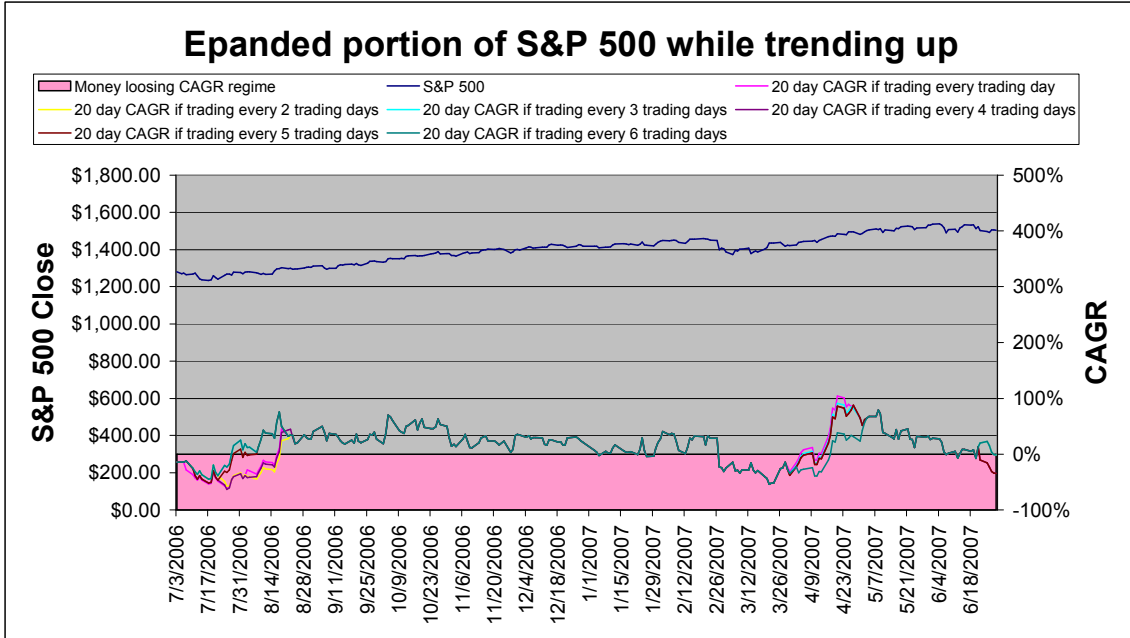


Figure 16

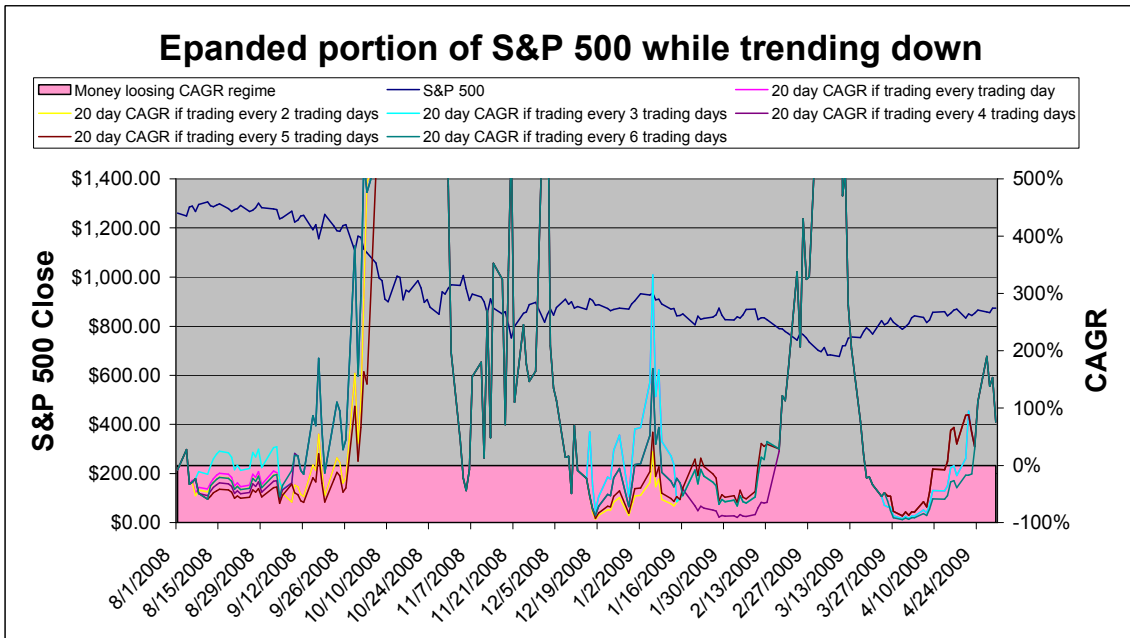


Figure 17

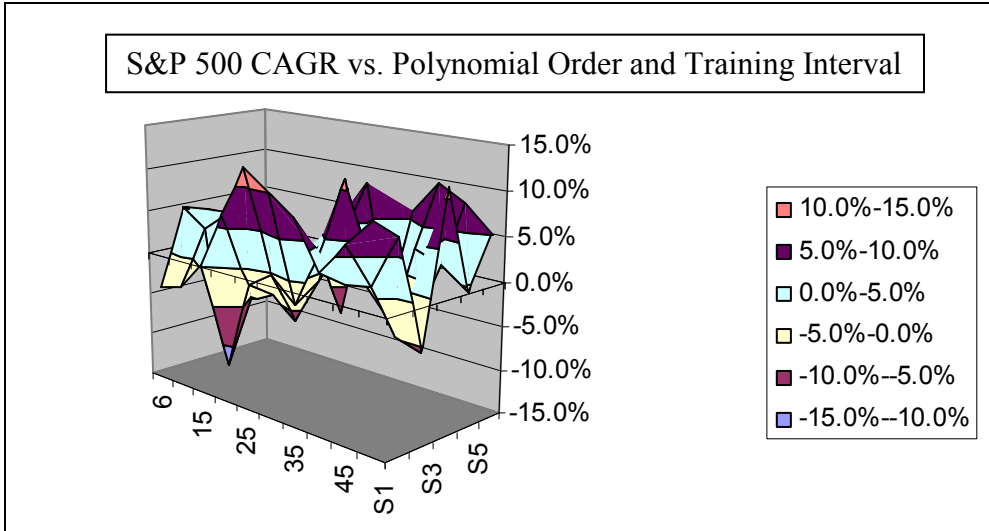


Figure 18