Predicting the Price of a Stock

An Interactive Qualifying Project
submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfilment of the requirements for the
degree of Bachelor of Science

By:

Apiwat	**Ditthapron**
Chaiwat	**Ekkaewnumchai**
Passavich	**Nalamlieng**
Panuwat	**Kongpornjaras**

Term A, B 2017 and C 2018

Submission Date:
March 2, 2018

Project Advisor:
Prof. Mayer Humi

Table of Contents

Abstract

This project investigates the possible short term prediction of stock prices using historical data of the stock and market in general. To this end, we created several models from real data and integrated multiple tools such as autocorrelation, least square linear fit, fourier series, moving averages, correlation with Nasdaq and Dow Jones. The models provided good predictions for over half of the stocks under consideration for ten business days. Further refinements of the model should be considered.

Executive Summary

Term A: At the start of term A, we learned the basic ideas about how stock market works and how we can relate it to statistic and statistical models. Throughout the term we started to developed our model slowly and encounter both negative and positive results. At the end of the term, we are able to understand and accurately use the idea of rsi index, autocorrelation, linear least square regression and residual plot, and fourier theories in our models to predict the future stock price with our ten samples. Although there are still errors and complicated pattern that could not be understood, we hope that we will learn and accomplish new methods that could justify these confusions about our models on the upcoming term.

Term B: At the start of term B, we learned how to approach the historical data differently. During term A, we used closing price from historical data to make a prediction but in term B we decided to use the moving average of the closing price to make prediction instead and then compare the results. This new alteration widen our option and provide useful information that is beneficial to our model. We were also introduced to NASDAQ and DOW JONES index and their influence on our stock prices. These two indexes were then added to the model to see how might the correlation between them and historical data of stock prices could improve the prediction. By the end of the term we have a more solid model with support from statistical data that were collected. This an essential element to improve the accuracy of the model in the future.

Term C: The last term of IQP was all about improving the accuracy of the methods we did during term A and B. We came up with the hypothesis involving the correlation between predicted days and volatility. We predicted that volatility could help us estimate the number of days that stock prices could be predicted accurately. At the end, even though the results proved that our hypothesis is incorrect, we were still able to obtain results with data to provide great conclusion for our IQP project.

Introduction:

For some people, investing in stock market is one of the best ways to make money, but for the others, it is where people throw their money away. In order to make money from investing in stock market, people have to be able to predict the ups and downs of a stock price and know the right time to put their money down or when to start withdrawing money away. There are many factors that influence a stock price, and it is essential to know the positive and negative impact that these factor could do to the price. It is really complicated and tricky to be able to judge how each factor influence a company's stock. Although there is a way to use mathematical model as a guide to stock prediction. Therefore we form a group of four students to do an official IQP project on predicting the stock price.

The stock market IQP project also allows us to apply our courses of study at WPI into the work. Since half of the group is a computer science major and the other half is an industrial engineering major, we decided to focus on improving the efficiency of the model through computer programs. Industrial engineers at WPI are taught to work on improving the process on any system in order to minimize errors and generate satisfying result, while computer science major students are taught to generate ideas into codes. As a result this project is a great combination between the two majors. This project will also be beneficial to our future careers. Our dream jobs range from working in a corporation like facebook or a becoming engineers in a large company. Industrial engineers' work will likely require to create models that involves time into the variable, which is similar to this project. This project is also a great practice for computer science major students for solving problems based on the statistical methods and research using programs that they might encounter in their future career.

The main goal for this IQP project is to help individuals who do not really understand stock market to be able to predict stock prices using simple technical tool. Therefore, we would create a statistical model from MatLab program that can accurately predict the stock price for any company with minimum errors as much as possible to fulfill this goal. In order

to do this, each of us select ten stocks as our samples in testing the model throughout the term. Each week, our goal is to learn a new method or new way to shape our models that would minimize the errors for the prediction and test it out on our samples. By the end of this project, we hope to combine these methods throughout the terms to create a solid model that could accurately predict the price of any stock.

Background

Autocorrelation:

Autocorrelation is a mathematical representation of data, which a correlation between the values of similar variables is based on related object. In general, autocorrelation could be use for two purposes. The first purpose is used when dealing with data pattern or correlation, in which it is used to detect non-randomness in data. The second purpose is used when dealing with time series modeling, in which it is used to find an appropriate time series model when data is not random.

Sometimes autocorrelation could be referred as lagged correlation when use to determine the correlation between present values and their past values. The equation, given lag k, of autocorrelation function is .

$$r_k = \frac{\sum_{t=k+1}^{n} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n} (y_t - \bar{y})^2}$$

Linear Least Square Regression:

Least squares fitting, developed to estimate obtained parameters by mathematicians Karl Friedrich Gauss, Adrien Marie Legendre and Robert Adrain, is one of the most commonly used modeling method. It is a procedure for finding the best-fitting curve or line to a given set of data points through minimizing the sum of the residuals, or the square of the offset by the equation $y = mx + b$. There are two types of offsets that we can use to create the least square fit. The vertical offset where the the offsets are minimized instead of the perpendicular offsets. Therefore, it provides a fitting for the independent variable and estimates the dependent variable for a given independent variable, giving an easier generalized form of best fit line.

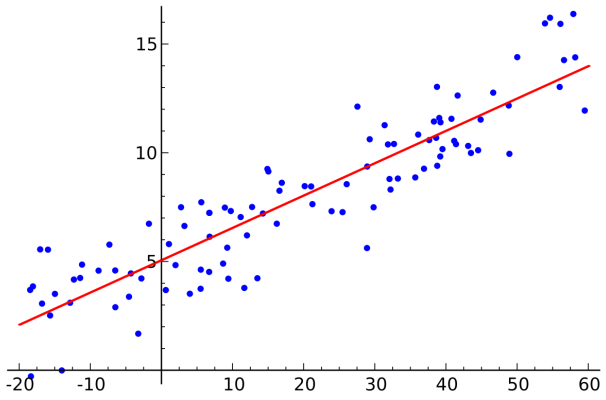


Figure 1: Linear Least Square Regression Example from

https://en.wikipedia.org/wiki/Linear_least_squares_(mathematics)

This is an example of a plot when linear fit is being used to find the correlation between two variables.

Fourier Series:

Fourier series is created by french mathematician, Jean-Baptisete Fourier, can be described as an expansion of a periodic function in terms of sines and cosines by the equation

$$g(t) = a_0 + \sum_{m=1}^{\infty} a_m \cos\left(\frac{2\pi mt}{T}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nt}{T}\right)$$

$$= \sum_{m=0}^{\infty} a_m \cos\left(\frac{2\pi mt}{T}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nt}{T}\right)$$

Sometimes it is referred to have share similarity to taylor series, which defines functions as infinite sums of monomial terms. Fourier series is useful when dealing with problems with arbitrary function. Fourier series breaks up the function into set of simple terms that may be substitute in, solved independently and then recombined to find the answer to the original problem. It is also an method to solve heat equation and spherical harmonics. For example fourier series can be used to solve the equation $\frac{\partial T}{\partial t} - \alpha \nabla^2 T = 0,$ where T is temperature, t is time and alpha is a constant. Since the fourier series equation involve sines

and cosines, its graph often look like this

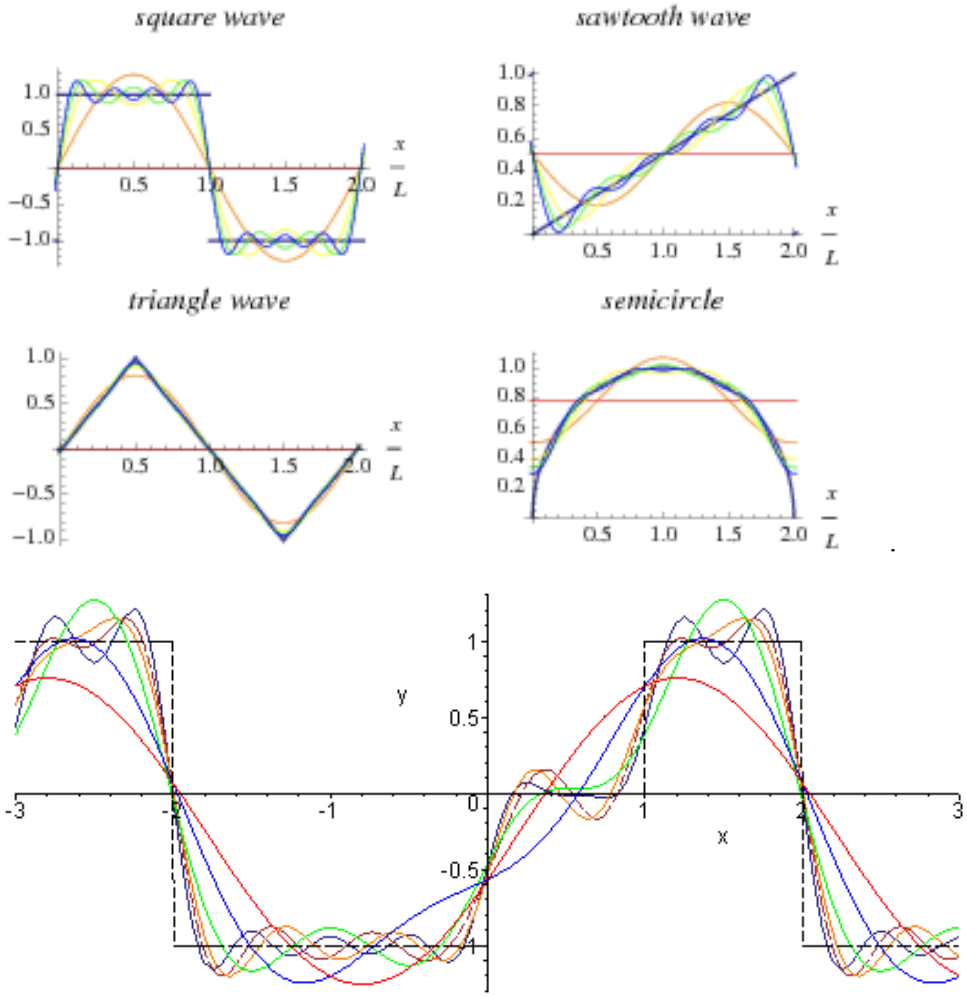


Figure 2: Fourier series fit of different waves from

http://mathworld.wolfram.com/FourierSeries.html and

https://www.tes.com/lessons/zZmFBnJtw6OM_A/fourier-series

Fourier Transform:

Fourier transform is a generalization form of fourier series in the limit for as L to infinity. It provides the alternate representation of fourier series, which deconstructs the waveform of fourier series that is based on sines and cosines into sum of sinusoidal functions. The equation that represents fourier transform is $\hat{f}(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x}\, dx.$. It is used to solve problems that fourier series could not solve like an arbitrary function over the entire real line. When talking about sound, fourier transform is also important in signal processing. The fourier transform helps people to view the signals in different domain, which a fourier transform of signals provide frequencies that are present and their proportions.

Moving Average:

The purpose of finding average is to find the rough estimate number of the set of data. However, there are more than one way to determine average of any set of data. Moving average is another way of finding the mean for overall data. It can be calculated by finding the average of a subset of data and then combine the average of each subset to see the trend of the data. Moving average could be useful in many contents related to statistics. For example, if we were to find the 5 year moving average of car accidents from 1990 to 2018 then we would find average from 1991-1994, 1992-1995 and continue until we reach 2018. When the 5-year moving averages of the accidents are plot into a graph we could inspect the trend and correlation through the line of best fit.

Some people might adopt moving average to help dealing with stock prices. It is widely used in term of technical analysis that helps smooth out price action and random price fluctuations that may be problems for predicting stock prices. Some analysts adopt moving average as trading signal, for instance, a rising moving average indicates that the security's trend is up and vice versa when it is declining.

Nasdaq and Dow Jones Index:

NASDAQ and DOW JONES are indexes that help indicate how well the current stock market is doing. DOW JONES index follows the performance of 30 companies that are considered significant players in the industries and mainly in New York Stock Exchange, while NASDAQ index tracks about 4,000 stocks that are traded in the NASDAQ exchange (electronic exchange where investors can buy and sell stock). It can be seen that the two indices have different perspective in determining the overall performance of the stock market. DOW JONES only focus on the main players that they believe are influential to the stock market, while NASDAQ prefers to focus on the quantity, which they believe high number of stock quantity could better identify the overall performance and trend of stock market. A lot of people might be confused that they could trade stocks in Nasdaq and Dow Jones index, but these two indices are only a mathematical average that stock analyst use to assess the trend and performance of stock market.

Volatility:

The word "volatility" alone could be defined as "liability to change rapidly and unpredictably." However, it could be used in many contexts to help describe certain states or identify characteristics.

In mathematics, volatility is a statistical measure of the dispersion of returns for given set of data by using standard deviation or variance to determine it. In economics, volatility helps identify degrees of variation of the stock prices. Generally, high volatility means high risk and low volatility indicates low risk since stock prices might be easier to predict. Analysts often find volatility of their target stocks prices to analyze the risk that he or she might need to consider into making decisions. It could be another essential factor in the decision to invest in certain stock or not.

Research

Stock Portfolio 1:

This portfolio stocks was picked from technology section at the beginning of September 2017 by looking at top-40 most active stock on NASDAQ ranking website. RSI and closing price range were used as main indicators to determine a proper stock. However, volatile stocks and penny stocks were avoid in this study. The final list of chosen stocks are eBay Inc.(EBAY), Yelp Inc.(YELP), TripAdvisor, Inc. (TRIP), Corning Incorporated(GLW), QUALCOMM Incorporated (QCOM), Symantec Corporation (SYMC), CyberArk Software Ltd. (CYBR), AT&T Inc. (T), Kyocera Corporation (KYO), Wix.com Ltd. (WIX)

eBay Inc. is a big online auction and shopping company that acquired many digital payment system companies such as, Paypal and Braintree. Paypal and Ebay have closed relations in revenue and stocks trend. In last two quarters, company stock price has increased more than 25% without any retracement, and RSI trend is over 60 containing only a few frustrations. Regardless of no new released products, it is a worth stock to buy.

Yelp Inc., known as online restaurant review website, Year-To-Date has a declining trend, which it is inclining in this quarter which might be confounded by other e-commerce companies that are growing rapidly in past few months. However, the stocks are overbought from seeing the RSI is above 60, but the stock price is not increased. There is a chance, in recent future, that it will be decreased back to its normal trend, which short-selling before that time probably give profit.

TripAdvisor, Inc. is a travel and restaurant website company providing hotel and restaurant reviews, accommodation bookings and other travel-related content. In last few quarters, TripAdvisor's quarterly revenues reflect an increase of 14.0% sequentially. However, the stock dropped down 14.8% in early June along with other companies in E-commerce market. RSI is frustrating between 50 and 70 during pass two months, which can predict that the price of stock is going up continually in short-time prediction. Since TripAdvisor is a

traveling company, it usually get boost up two or three months before December, the month that has long weekends and most people are on the vacations which is its main source benefit.

Corning Incorporated is an manufacturer of glass, and related materials, primarily for technology industrial. The company is known as local LCD and fiber optics supplier. In past year, the company has a strong growing since Apple announced that it will be investing $200 million in their technology. Shares of Corning have increased 32.4% year to date. Now, U.S.-based service providers overhaul networks to improve 4G coverage and accelerate 5G deployments, which mainly uses Corning major productions, the Optical Communication lines. The RSI currently is above 70, but the stock price is about average of this quarter. Buying this stock and holding for one or two quarters seem to give a profit.

Qualcomm is a semiconductor equipment and wireless related telecommunications producer company. At the beginning of A term, It is on downside with shared down 22% year-to-date because of the last dispute with Apple Inc over manufacturing license that Qualcomm holding preventing other companies to improve their products. The patent interferes with new 5G communication system, which once boosted its stock price up to highest point. RSI is also frustrating between 30-50% signal that the stocks are being sold overtime, so short-selling will give profit before the result comes out.

Because of Equifax hack on September 7,2017 leaking customer privacy informations such as social security numbers, birth date, address, and phone number, it raised competitor company stock, the **Symantec Corporation**, priced up 10 percent in two days. Main company production is software for security, storage, backup and availability. From this situation, traders who holding SYMC stock got a lot of profit in short time and seem to be got more.

Another company that affected by Equifax hack is **AT&T**, the world's largest telecommunications company, but in a different direction. It went down 10 percent on that week, because it uses Equifax security system product to manage internal system, regardless of its claiming that there is no customer information leaking from the company. However, luckily apple released new IPhone devices, that would be sold by at&t, a few days after the attacks that which pulled up the stock price back a little bit.

CyberAsk company is security software company focused on eliminating cyber threats using insider privileges. It has a different customer target from Equifax, so there was no effect observed. Last year, CYBR stock had a strong declining trend, but the company itself gained profit according to company report. After mass Equifax hacking, cyber security has become a certain issue and CyberAsk is one of companies that got directly benefit. The stock price is climbing up with RSI gradually raising above 50.

Last stock is **Kyocera**, a Japanese multinational ceramics and electronics manufacturer headquartered in Kyoto, Japan. According to Kyocera fiscal first-quarter report, the company shares have risen more than 17 percent and the stock has risen 21 percent during the last 12 months. Kyocera seems to be in a good status. However, RSI and the price are diverging which indicates the retracement. During this retracement, stocks can be bought at the lower cost than normal trending price and sold after the retracement.

Most of the stocks in this portfolio are known as immense companies in online marketing and cyber security software.

Stock Portfolio 2:

The ten stocks are chosen at the beginning of A term 2017 from the technology section, which the penny stocks are avoided due to its unpredictability. After filter down to one sector, the stocks are chosen by their characteristics of predictable prices trend based on statistical data in the past and strength of the companies in term of doing business.

Aaron's Inc is a lease-to-own retailer that focuses on electronics and computers. The current price of the stock is rising as well as the rsi. The price of the stock shares a positive relationship with the rsi. The recent pattern shows that as rsi increase, due to increase in stock buyers, the price of the stock also increases. At the moment the rsi index is still below 50, which means there are still more sellers than buyers. However, the increasing number of rsi index may suggest that the price of the stock would continue to rise.

Blackbaud Inc is a supplier of software and services mainly to non profit organizations. For the past months, the stock price of Blackbaud has been fluctuated from $83 to $85 which makes it kind of difficult to predict. The past records indicates that when the stock of Blackbaud is overbought the price seemed to jump rapidly. Recently, the rsi index for this company is increasing and becoming close to the overbought line, so there is a high chance that the stock price for Blackbaud will stop fluctuating and start increasing continuously in the future.

Amdocs Limited is a company that supplies software and services to media and communications companies. Currently, the stock price and rsi index for this company is dropping. However, BT Group, one of UK's leading communication companies, chose Amdocs to be its service platform. This could create positive impact to the stock price since Amdocs just landed on a huge client. Amdocs just introduced "smartbot", which would help them with customer services and enhance its technological advancement. This would also influence the stock price in a positive way.

Black Knight Financial Services is a corporation that provides services, integrated technology and analytics data solution to real estate industries. For this stock, there is no strong relationship between rsi index and price of the stock. Recently the company announced that they would complete a transaction of equity interest to holders of FNF group common stock. This would strengthen the capability of the company therefore it might be the influence factor that escalated the stock price. It would also suggest that both the rsi and stock price would continue on the positive trend.

Ebix is a leading supplier of software and ecommerce services to the insurance industry. For the past months, Ebix's stock price has been fluctuating between $57 to $58.5. Currently, the price is increasing but due to fluctuation it cannot concluded that the price will continue to rise as well as the rsi. Therefore, it would be risky to keep investing on this stock.

However, when stock was overbought last month, the price escalated from $57 to $63. Currently, the rsi is at 60 so a huge jump might occur again when it reaches 70.

Fortinet Inc is multinational corporation that develops and sell cybersecurity software. The stock and rsi of Fortinet seems to share a strong relationship. Recently, the company announced that they would expand its research and development operations facility to Vancouver area which will create 1000 more jobs. This news could be the influential factor in increasing the price of stock and attracting more buyers. This trend could continue to grow, as a result it might not be a bad idea to invest in this stock.

Open Text Corp is a company that develops and sell enterprise information management software based in Canada. The past records seem to indicate that the stock price tend to increase when there are a lot of buyers, especially during the overbought period. Currently, the rsi of this stock is at 65, which means it is about to enter the overbought period. There is a possibility that this stock price would escalate after exiting its fluctuating period.

QAD Inc is a company that supplies enterprise resource planning software to manufacturing companies. QAD's stock price shares a relatively positive strong relationship with the rsi index. About a week ago, it was at an overbought period, which means the price was rising very quickly. Currently, the rsi number started to drop below 70, therefore it can be predicted that stock price will fall as well.

2U is an educational technology company that partners with some of colleges and universities in the US. Recently, the company announced that they would offer 3,000,000 shares of common stock. This would be an influential factor that brought the company stock into a overbought period with rsi of almost 80. The trend of the price is still positive and it might continue while the rsi is above 70.

Zendesk is software company that help other companies deal with customer services. The report of the resignation of the chief revenue officer last month caused the stock price to drop drastically. The resignation seems to be effective a few days ago, which means the company might be going through some changes and might create negative impact to the stock price. Currently, the company stock seems to recover from last month drop but it is possible that the drop could happen again especially after the rsi drops below overbought point (70).

Stock Portfolio 3:

Cisco is a worldwide leader in networking for the internet. The pricing of the stock of Cisco plummeted is due to the Equifax hacks. Equifax hackers were able to access around 150 million accounts containing name, date of birth, SIN, address and as well as credit information. Due to Cisco-owned Open DNS blocking access to the site. Through the hacks and Cisco's relation with Equifax, it can be seen why Cisco's stock price dropped. Prior to these hacks it would've been impossible to predict that the price of the stock will drop this far. Next week judging from the RSI index and the direction that the stock is going I predict that the stock will stay at this current price since the RSI is almost going at a straight line.

Intel is one of the biggest CPU producers in the world. Prior to AMD's new CPU product line, Intel had a huge market share over every other competitor within its own respective market. However, after AMD's new Zen CPU chip released, AMD has gained 10.4% of the market share from Intel all in 2nd quarter of 2017. Therefore, it can easily be deduced the reason why Intel stock prices are have dropped. If AMD stays on track with this new CPU, releasing new lines to compete with Intel's then we will see Intel's market share drop more and more. However, in the past week, Intel's stock has increased drastically. The reason behind this is Intel's plan towards investing 1 billion dollars to AI startups to build a base for product innovation within that market. RSI for intel is also getting higher signifying

that people are buying more stocks than selling which will further increase the stock value of Intel.

Momo is a software company that offers free location-based instant messaging application for smartphones and tablets. The app allows user to chat with nearby friends and stranger. I wasn't able to find many news for this company but I believe that the reason why the stock prices are dropping is because of its user base. In terms of users, there wasn't a considerable amount of growth within this market which potentially causes the stock to drop since the application isn't expanding at a significant rate. I believe that this is the main reason why the stock price is dropping. However, in the following weeks I believe that MOMO prices will start to rise again judging from the positive RSI index trend.

Dell is a computer company that produces various technologies in relation to desktop PC, laptops and various other lines. As of early September, Dell and GE, the world's largest digital industrial company, has signed a multi-year commitment to use Dell's infrastructure to support GE. This would be the main reason as of why Dell's stock has increased greatly. If we look at the RSI index, we can see that it reached around 90 and is dropping. This is a sign that people are overselling the stocks and the stock price will come down.

EA is an electronic gaming company that has released various high rating games in the past. I believe that it is the nature of the gaming stocks to constantly go up and down since games and patches are released so often and will ultimately affect the price. However, judging from the news that EA will be releasing a DLC for Battlefield, a game that has high count of players, and releasing a new yearly game of NBA 2k18 I can predict that the price of EA stock will increase in the next few weeks. Judging from RSI there is no trend whether EA's stock will increase or decrease since the RSI is around 50 for a few weeks now.

Oracle is a computer corporation that specializes in developing and marketing database software and technology, cloud engineered systems and enterprise software products. Prior to its stock price downfall, Oracle's stock price was increasing. However, the stock plummeted very intensely within this past week due to its wrongly reported forecast. The company's forecast was expected to bring in more money than in reality. The forecasted

revenue should be 9.2 billion dollars but in reality the revenue is 9.02 billion. Judging from the RSI which decreased greatly I believe that the price of the stock will continue to drop within the next week but will eventually regain back up again.

Stock Portfolio 4:

I focused on 9 selected stock (lexicographically sorted): Activision Blizzard Inc. (ATVI), Comtech Telecommunications Corp. (CMTL), Intel Corp. (INTC), Logitech International SA (LOGI), Microsoft Corp. (MSFT), Motorola Solutions Inc. (MSI), Oracle Corp. (ORCL), Seagate Technology (STX), and Western Digital Corp. (WDC). In this part of the project, we will focus on prediction after September 20; that is, we will use the data from January 1st, 2017 to September 20th, 2017 to predict the trend of the stocks on and after September 21st, 2017. Hence, news presented in this report is mainly after September 21st.

ATVI is an American game developer which contains 5 business units. Its popular games *are Call of Duty* (first-person shooting), *Destiny* (first-person shooting), *Hearthstone* (card game), *Overwatch* (first-person shooting), and *Candy Crush* (mobile game). Throughout this year, there was one huge jump of price on February; the forth-quarter dividend was released, where its revenue and net income increased by more than 50 percent. Recently, most of the game stock companies such as Electronics Art (EA), Take-Two Interactive (TTWO), and Zynga Inc. (ZNGA) had a big drop on September 20th to 27th. Although on the ending of September, many game stock companies released quarterly dividend and announced new games, ATVI almost had no update. However, ATVI planned to launch its famous series *Call of Duty*. Moreover, on the third week of September, some analyst websites: The Motley Fool and Zacks, suggested it be worth to invest this stock. On the beginning of October, ATVI announced eSport League for its game, *Overwatch*, but some people did not believe it would succeed and they got nervous about the situation. Also, another game of ATVI, *Destiny 2*, showed a bad sign of sale; its sale dropped 50 percent. Therefore, its price has stably dropped since October, and its trend continues to drop.

CMTL is telecommunications equipment company for both commercial solution and government solution. Since January, there were three rapid changes occurred; on March, June, and September. First, it reported second quarter dividend on that day and declared $0.10 per share quarterly cash dividend, where its stock price was fairly low making investors get more volume. Similarly, on June, it released third quarter dividend; it also declared $0.10 per share. Last, on September, its quarterly dividend was announced, but this time it did not go as planned as it fell 3 percent short compared to what had been expected. Hence, the stock price dropped immediately in that day. On the predicted period, CMTL has shown a positive trend since the rapid drop in September. One of the reasons is that it signed new contacts with government such as international 911 mobile service and GPS-based tracking system.

INTC is an American multinational corporation and technology company. Its famous technology is semiconductor chip such as microprocessor and disk. INTC stock price pretty swings over this year, but there is no obviously rapid increase or decrease. Since our set date, INTC is also showing positive trend. Most of recent news of INTC is about intelligence; one of self-driving car creator company decided to move investment from Nvidia to Intel. Also, Intel put more effort in investing in artificial intelligence (AI). On another system, it also invested in quantum computing which may become more affordable.

LOGI is a Swiss global provider of personal computer and mobile accessories such as keyboard and mouse. There was a big drop on the late of July because of quarterly dividend; even though the report was positive in almost all aspects, investors still thought that it was not big enough to invest in the stock, so they retrieved the money back. LOGI has no update since the beginning of September, so the price is steady and depends on investor management.

MSFT is an American multinational technology company which own computer software and hardware. Microsoft Windows and Microsoft Office are ones of its well-known software which most of people are using. On the other hand, Microsoft also released Microsoft Surface, hybrid of computer and tablet, and Microsoft Xbox, game console which is popular in America. MSFT reflects its good standing since January; its stock price constantly grows up. The recent news about MSFT contains good and bad news. MSFT allies cloud computing service with Amazon to fight against Google. For bad news, MSFT's mobile

operating system, Windows 10 Mobile, got abandoned after it tried to combine the system to Nokia's phone, which had formed the partnership to Microsoft earlier. Furthermore, MSFT decided to use Android system instead, which many of its users want. There is a big update to its famous operating system, Windows 10. Because of this positive news, the stock price seems to be in good shape.

MSI is an American data communications and telecommunications equipment provider. One notorious product is mobile named Motorola; it used to be as popular in the past as nowadays iPhone. Even though Motorola is rarely used compared to iPhone, it still gets update every year. MSI stock gets plenty spikes throughout this year. One of the biggest drop came from alleged price gouging on February. There is little news about MSI; it is trying to launch old-style flipping mobile. Nevertheless, it got the most gain on October 6th because of the suggestion from notable to buy. In my opinion, the stock is manipulated by investor.

ORCL is an American multinational computer technology corporation which specializes in database, cloud computing service, and enterprise software. One of the well-known software to programmers is Java. There are two huge changes in this year: in June and September; they had the same origin, that were quarterly dividend. While both in June and September it was able to get profit and positive revenues, it could not reach the expectation of analysts, losing around 2-5 percent of what they had estimated, in September, which caused the huge drop in the month instead of the huge jump as in June. After the big lost in price on the late of September, ORCL is gradually recovering; its price is relatively steady between $48 and $49. ORCL's update on cloud computing is the important news to follow similar to Microsoft situation. To emphasize, currently ORCL holds the better situation than its rivals.

STX is an American data storage company whose well-known product is hard disk drive (HDD). Two massive drops occurred after quarterly dividend released. The first time was not as bad as the second time because the revenue of third quarter was fine but less than expected in the first time, while the second time it simply lost profit compared to last year at the same period. On the initial days of our observation, there was a bid which was near to finish with the predicted winner being STX's rival, WDC. However, Toshiba, who is the

bidder, agreed upon Bain Capital consortium, including STX who joined later, to sale its technology to them. This affected STX in a good way as it got advantages against its rival. This made the price slightly increased a little.

WDC, similar to STX, is also an American data storage company who mainly sales hard-disk drive. Different from STX, WDC only suffered from big drop only once right after the revenue's plummet on July. Its stock price is relative unchanged compared to its competitor STX. Unlike STX, WDC lost on the bid causing its price drop for a week, but it could recover from launching the largest storage hard-disk drive after struggling for long time. Also, recently Toshiba, another competitor in hard drive market, is discussing in joint investment with WDC, which may make investor more believe in WDC.

Modeling

Autocorrelation

To predict the stock by fitting Fourier series on residual of stock price that does not following linear trend line, we applied Autocorrelation to determine the period of time serie, which compute the degree of similarity between the stock price and lagged version of itself over successive time intervals. Autocorrelation output is varied in range of -1,representing absolutely negative correlation, and 1, representing absolutely positive correlation. Positive correlation in time series means that one increase seen in one time serie leads to increase in another time serie. In the opposite, negative correlation means that one increase seen in one time serie leads to decrease in another time serie. To determine the “vital” data that we will use to fit the fourier series and its errors, we calculated when the autocorrelation first reaches the midpoint, or 0, then used the independent value to apply the predictions.

Figure 3 and Figure 4 indicates that the significant period that can be used to create fourier fit is 63 days and 58 days respectively. If we decide to use the data beyond that it could create conflict within the function.

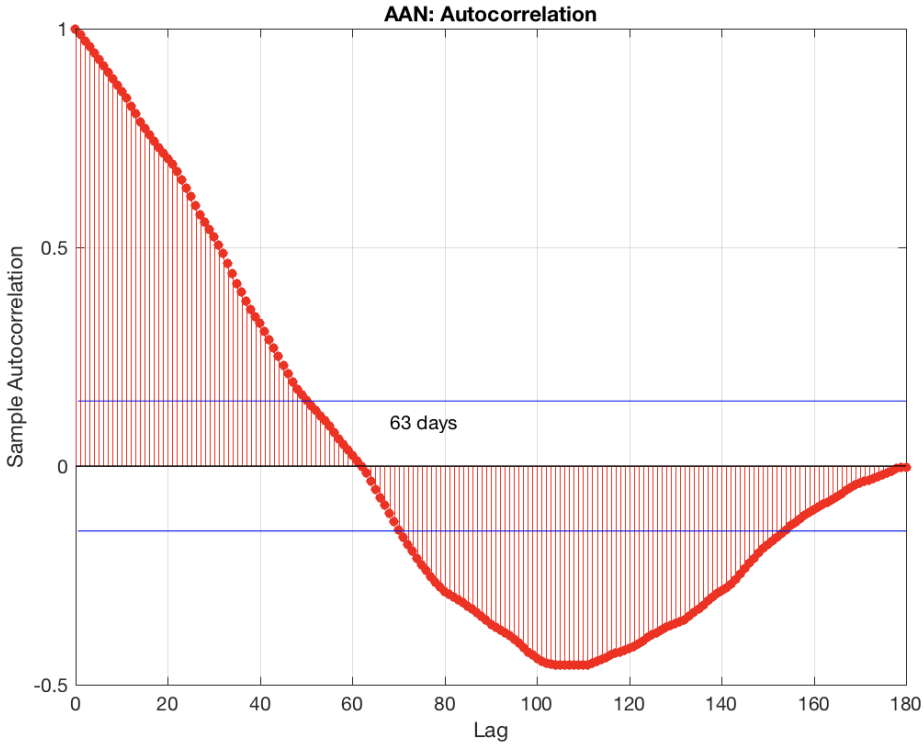


Figure 3: AAN Autocorrelation plot

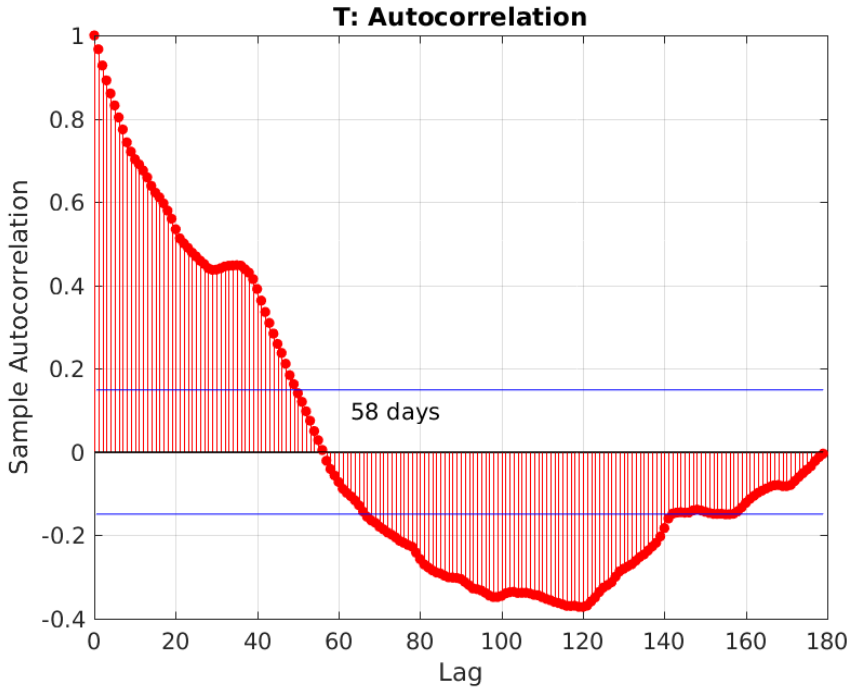


Figure 4: AT&T Autocorrelation plot

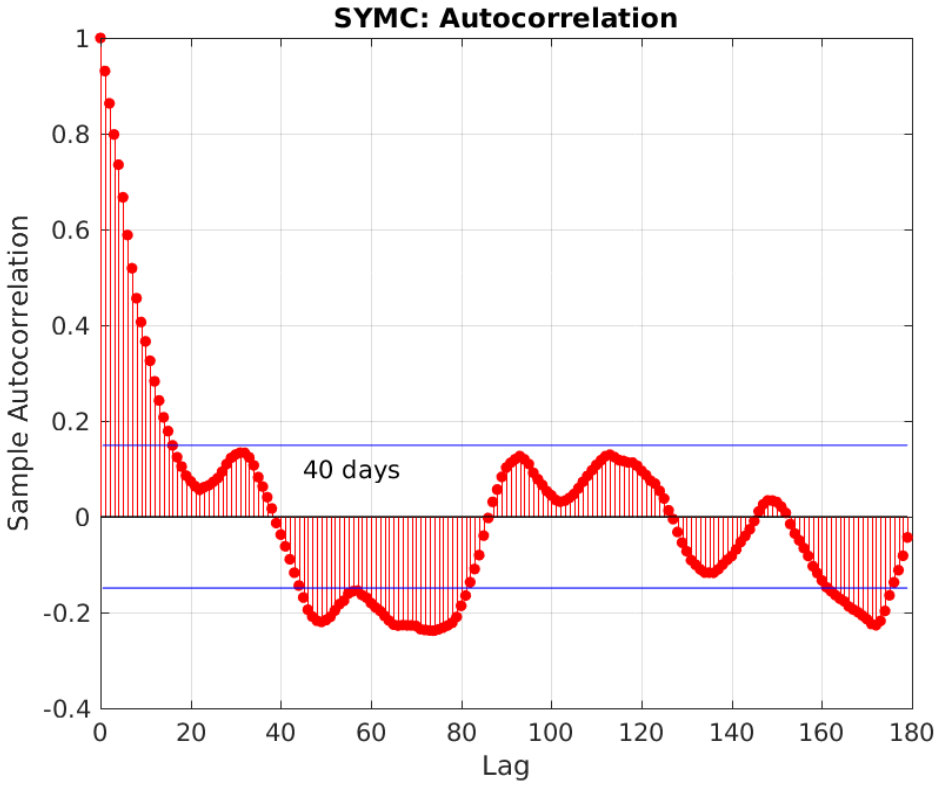


Figure 5: Symantec Autocorrelation plot

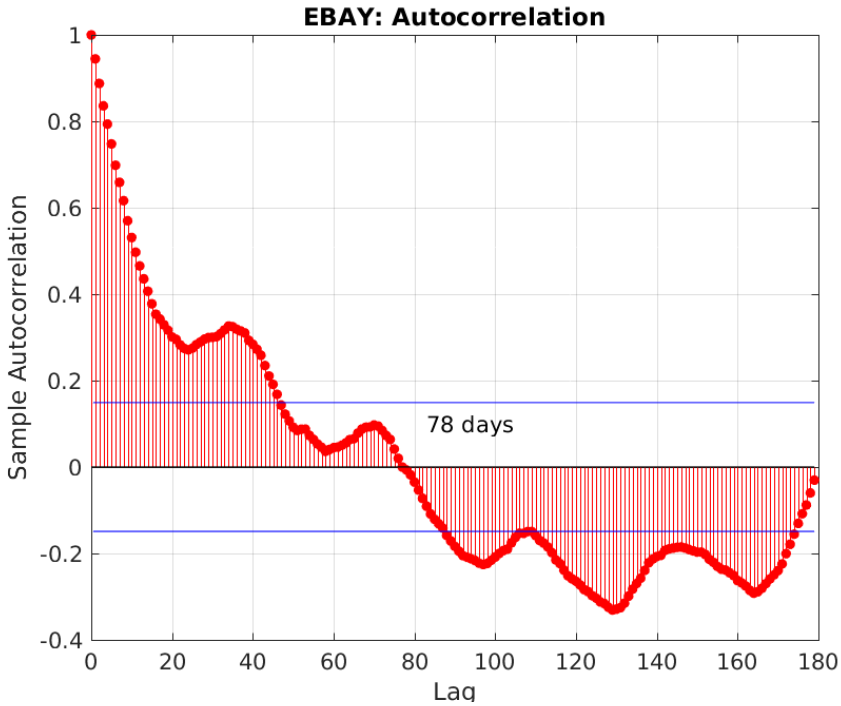


Figure 6: EBAY Autocorrelation plot

Data points are drawn within the Autocorrelation Function range to find the best period time of fitting. The range is determined at zero crossing point of sample correlation value as in Figure 5 and Figure 6. These data points is used to *choose the proper* function to avoid overfitting and underfitting.

Trendline of Stock Prices

For data analysis we used linear least squares to find the trend of the stock, which for the stocks we chose are a positive trend. Furthermore, we incorporated least square error to help us predict the price of the stock within a margin of error. In terms of stocks, it is impossible to predict any stock prices precisely; therefore, needing an estimate within a margin of error. In the aspect of residual, the residual itself will help us find the error of our least square fit. From this we can observe the difference between our fit and the actual real values to make calculated predictions in the future.

From Figure 7 and 8, If short-selling transaction is made at date 50, the huge amount of money can be lost, regardless of the decreasing linear trend.

Residual Plot is computed to investigate the uncertainty of stock and attempted to fit on prediction model.

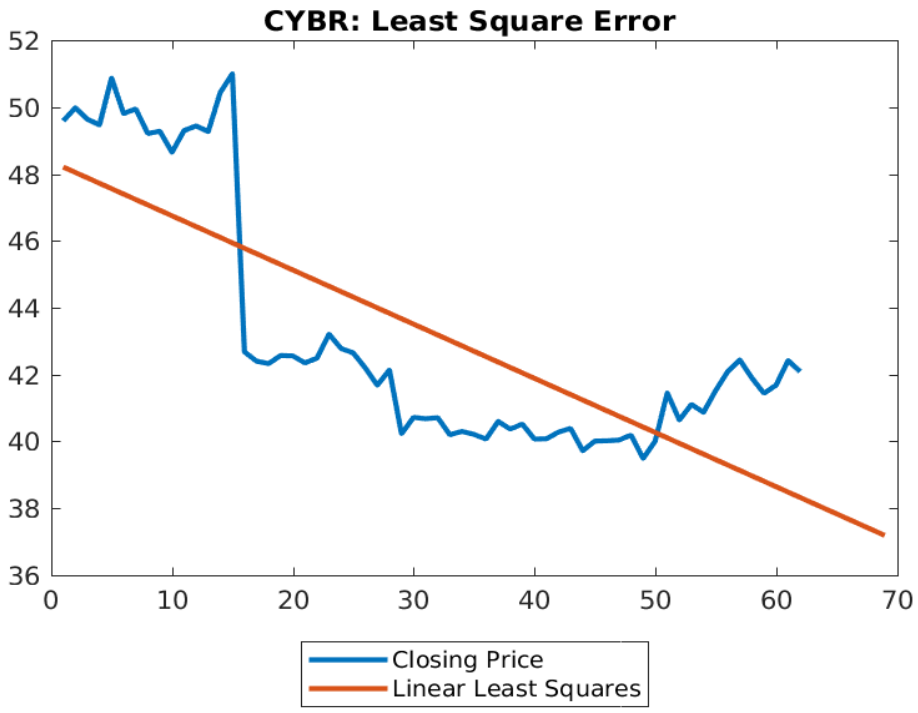


Figure 7: CyberAsk Closing Price and Linear Least Square

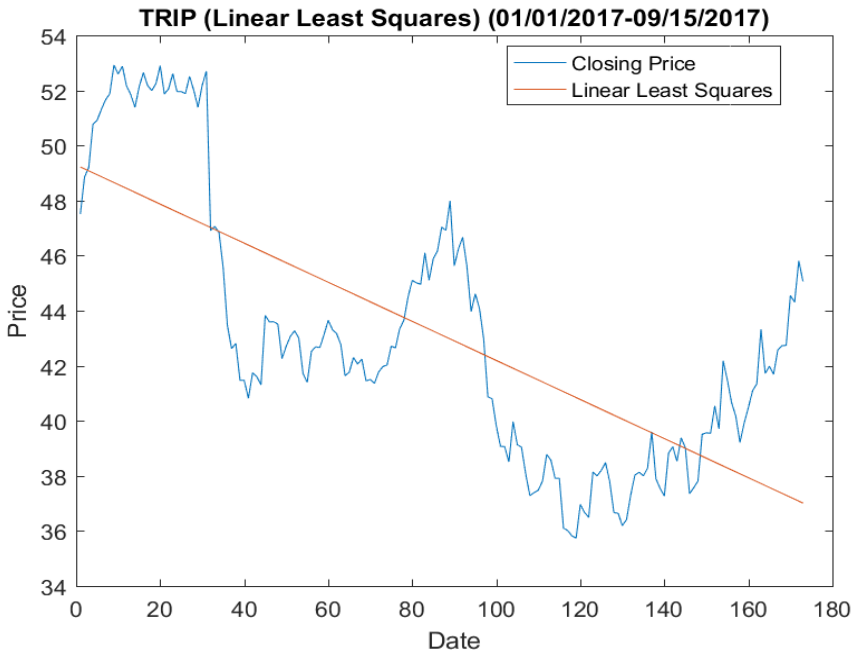


Figure 8: Trip Adviser Closing Price and Linear Least Square

Linear least square regression and residual plot also help us to see errors between the overall trend of historical data and actual stock price.

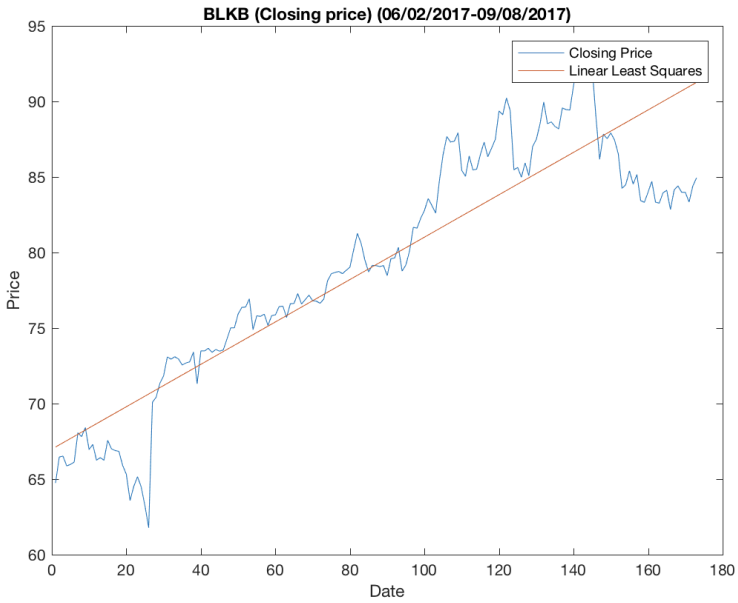


Figure 9: BLKB Closing Price

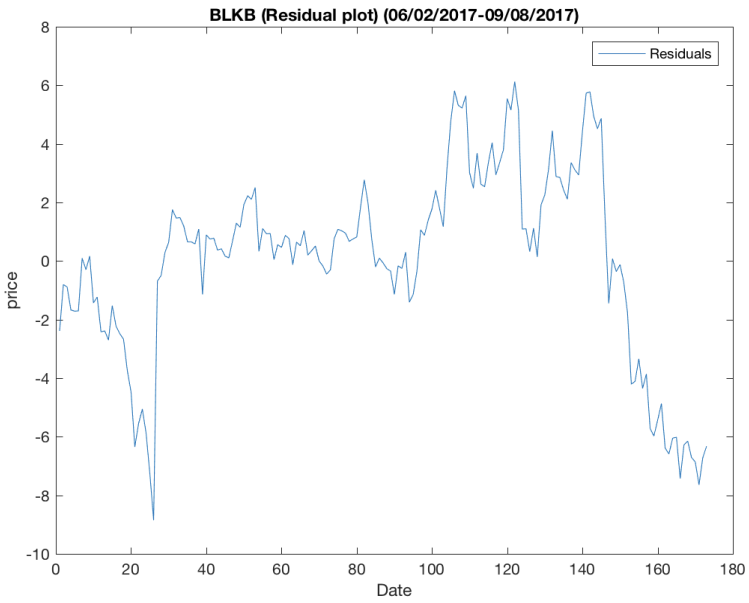


Figure 10: BLKB Residual Plot

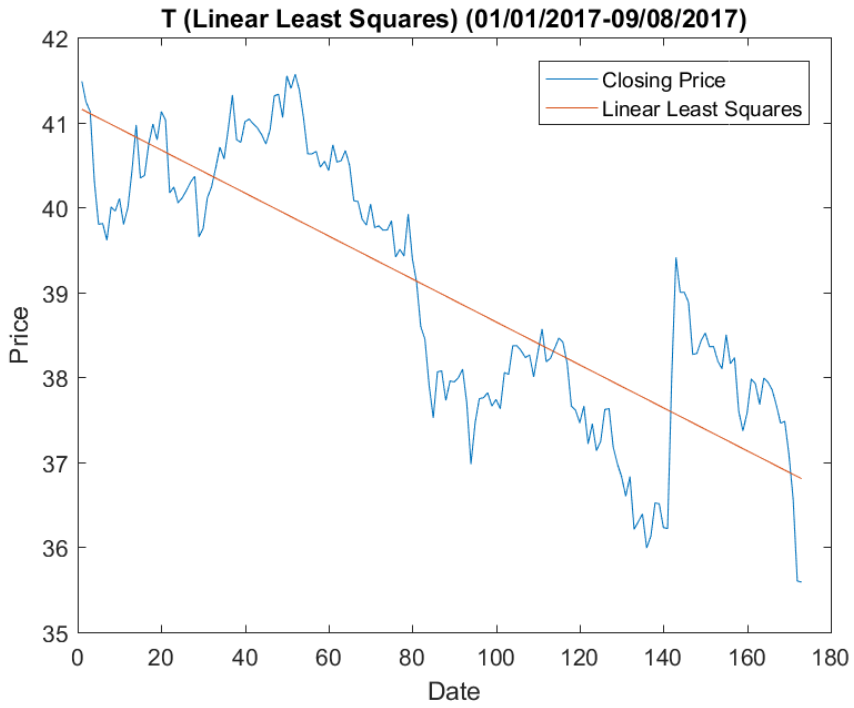


Figure 11: AT&T Closing Price

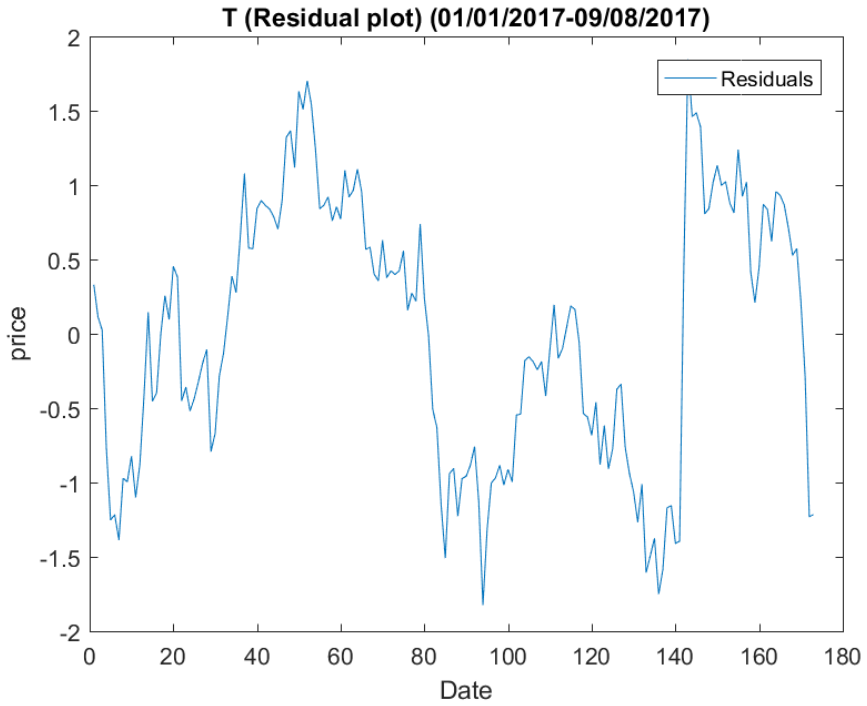


Figure 12: AT&T Residual Plot

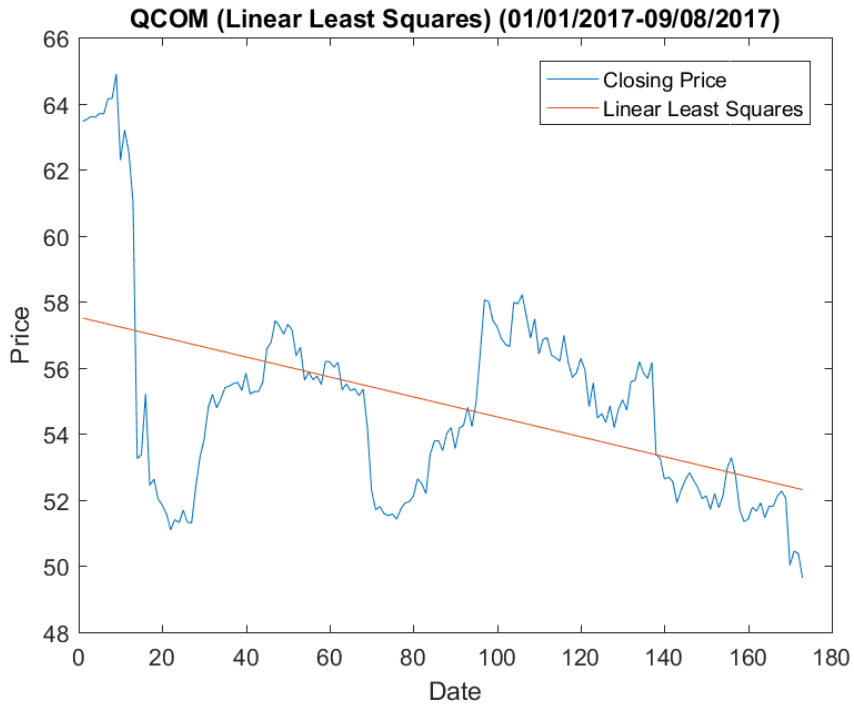


Figure 13: Qualcomm Closing Price

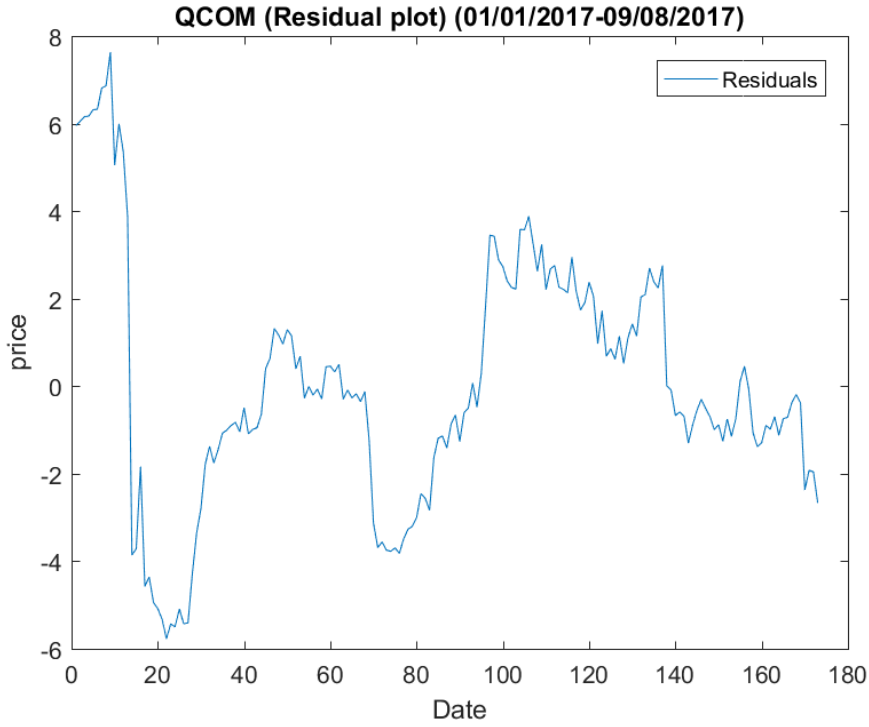


Figure 14: Qualcomm Residual Plot

Figure 9,11, and 13 shows the linear fit passing through the actual price and figures 10,12, 14 show the residual plot of the left graph. The residual plot will allow us to clearly see the error between a point on a linear fit and actual stock of the same date. Although we can see that it could roughly predict the trend but there are still huge error gaps that could ruined the prediction in extreme cases of price jumping or dropping, therefore it is necessary to use other tools to help lessen the error gaps.

Fourier Series

The nature of stocks is that it varies with time. No stock will see an increase all the time; there will be times where the stock goes and up down. Therefore, had we only use least square, our prediction of the stock price will not be accurate. We introduced fourier series to fix this issue, since Fourier series can predict sinusoidal graphs unlike least squares, which is used for a trend graph. Fourier Series is a method to break periodic functions into their sinusoidal components. The Fourier series can help us predict the the stock price that does not follow a linear trend line. By applying autocorrelation, to determine the period of time that is vital to our data, we either applied fourier 2 or fourier 3 to the predict the price of the stock. If the autocorrelation suggests that the vital data is less than 30 days, we decided to use fourier2, which has 2 sets of cos and sin. If the vital data is greater than 30 days, we used fourier3, which has 3 sets of cos and sin. This is to avoid overfitting, which can decrease the accuracy of our prediction. After obtaining the Fourier series and the linear least square, we calculated the Fourier series margin of error and used it to make a predicted guess on the stock price.

After using the method, we found that some of them can be predicted correctly such as MSFT's and Intel's stock (Figure 15 and 17); that is, the actual price is within the predicted range. Some of them are partially out of bound due to fluctuation of stock market. For example, investors are worried about Overwatch's eSport plan of ATVI, so the price dropped abruptly (Figure 16).

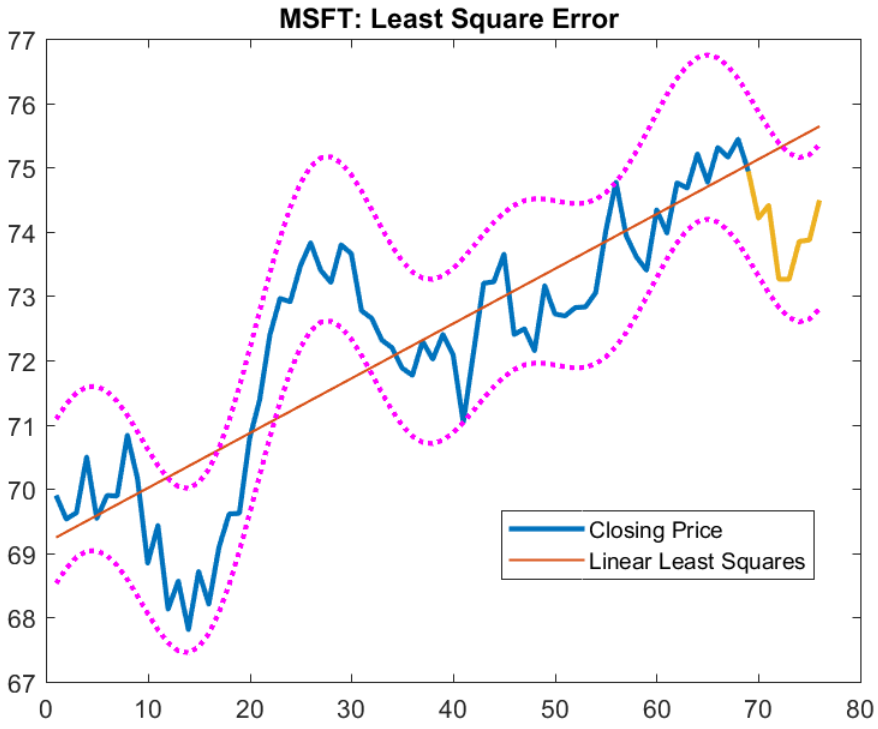


Figure 15: Microsoft Least Square Error graph

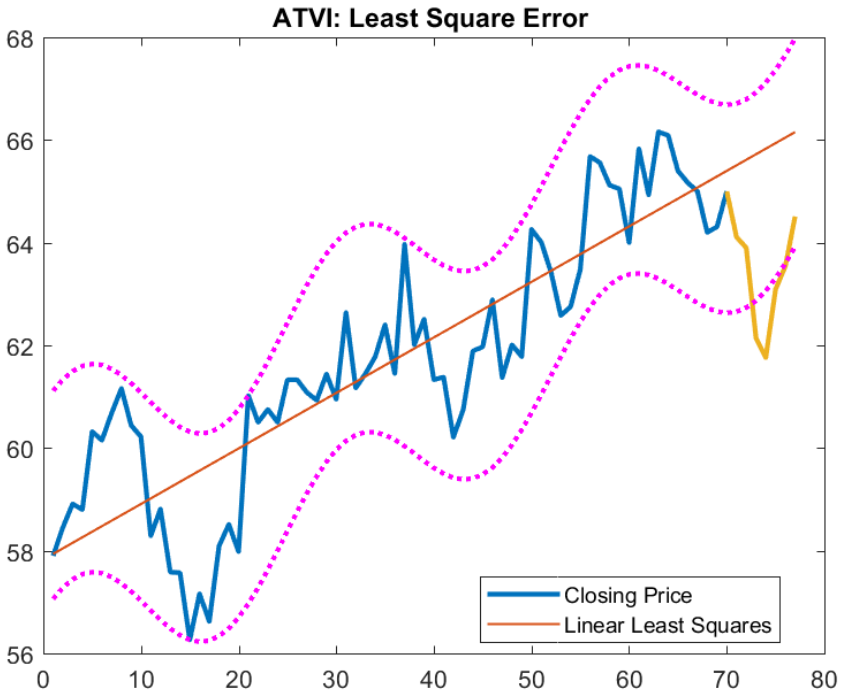


Figure 16: Activision Least Square Error graph

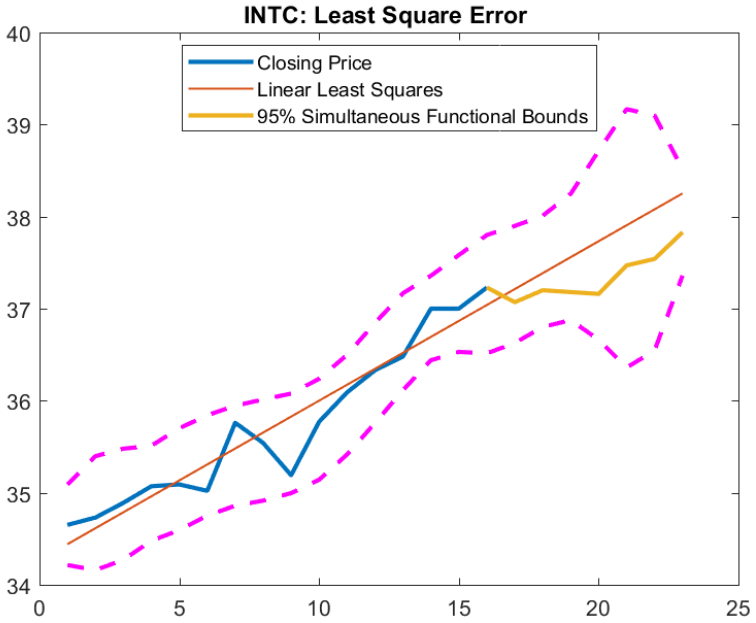


Figure 17: Intel Least Square Error

There are also some unusual events that fourier series could not predict, which resulted with problems shown in figure 18, 19, and 20. For example, Oracle's stock price dipped from its peak at ~52 dollars per share to approximately 48 dollars per share. The cause of its downfall can be traced to its incorrect forecast. The forecast was miscalculated and the result was a difference in 200 million dollars less in revenue than they had originally predicted. As according to the prediction, the price of the of the stock would've dipped so greatly that the stock price would've dropped immensely. This is one of the situations where the prediction model would be in no help because of the sudden change in the perspective of stock buyers.

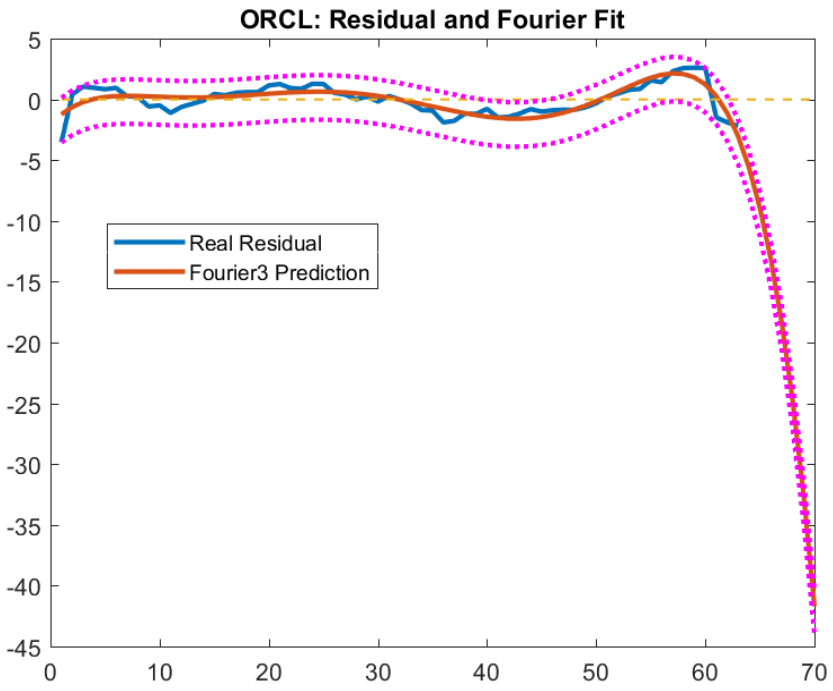


Figure 18: Oracle Residual and Fourier Fit

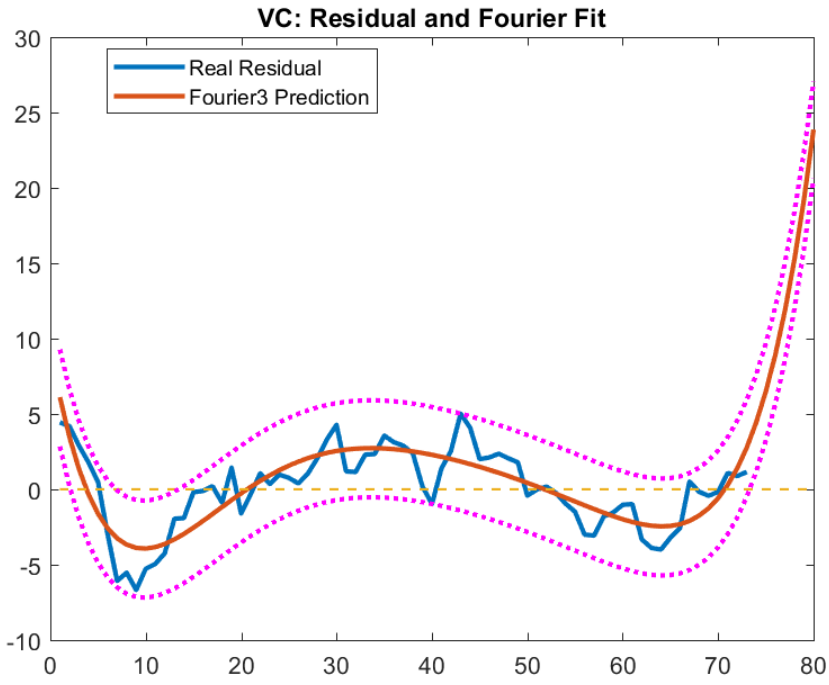


Figure 19: Visteon Corporation Residual and Fourier Fit graph

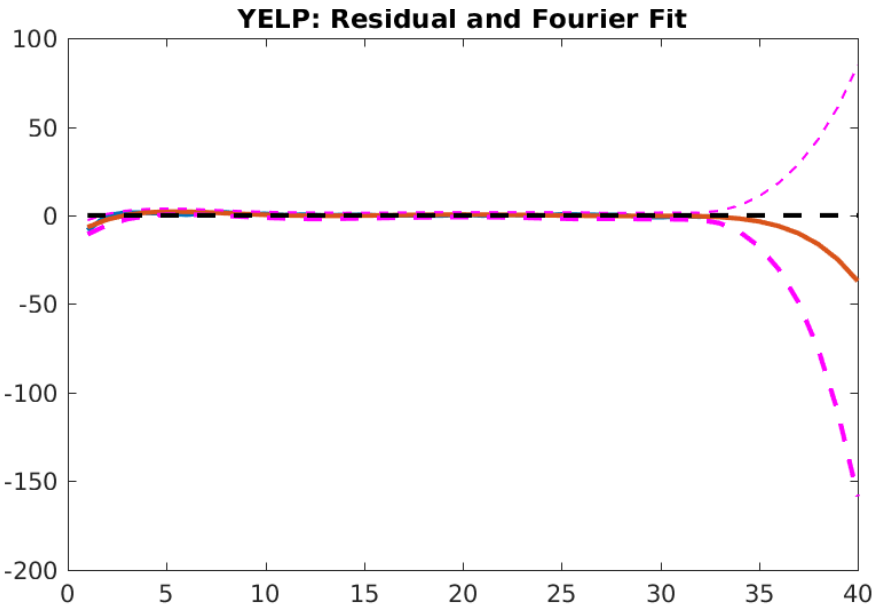


Figure 20: Yelp Residual and Fourier Fit graph

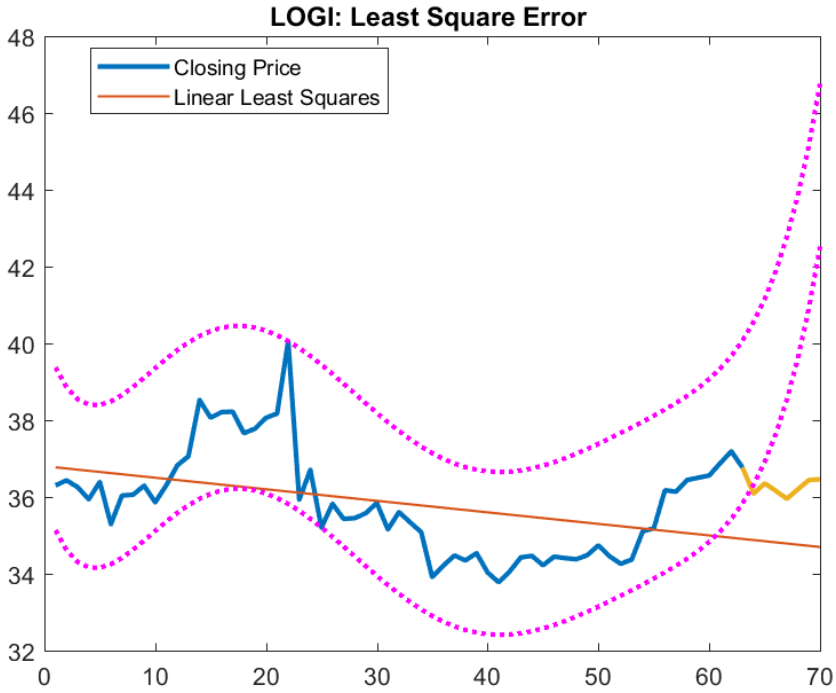


Figure 21: Logitech Least Square Error graph

Lastly and interestingly, the others' graphs are in bad shape because of Fourier fit's overfitting where the price unexpectedly goes up after learning period. We obviously cannot

use these graphs to predict the stock prices. The example is Logitech's price (Figure 21). This is a problem that we encountered and could not find out what actually caused them or how to perfectly fix it. However, we can limit how bad graphs will be. When inspecting coefficients of unexpectedly misbehaving fourier series (figure 22), we found out that they went to 10^8 and the period of the graph is more than 20 times of the observed period.

```
General model Fourier3:
fitFourier3(x) =  a0 + a1*cos(x*w) + b1*sin(x*w) +
          a2*cos(2*x*w) + b2*sin(2*x*w) + a3*cos(3*x*w) + b3*sin(3*x*w)
Coefficients (with 95% confidence bounds):
  a0 =   1.393e+08  (-1.558e+11, 1.561e+11)
  a1 =  -2.074e+08  (-2.33e+11, 2.326e+11)
  b1 =  -2.643e+07  (-2.472e+10, 2.466e+10)
  a2 =   8.113e+07  (-9.166e+10, 9.183e+10)
  b2 =   2.102e+07  (-1.966e+10, 1.97e+10)
  a3 =  -1.301e+07  (-1.492e+10, 1.49e+10)
  b3 =  -5.204e+06  (-4.897e+09, 4.886e+09)
  w =    0.003544  (-0.6573, 0.6644)
```

In addition, the predicting line (figure 22) for one period went to the coefficient values (10^8) though we would not use this model to predict for that long period.

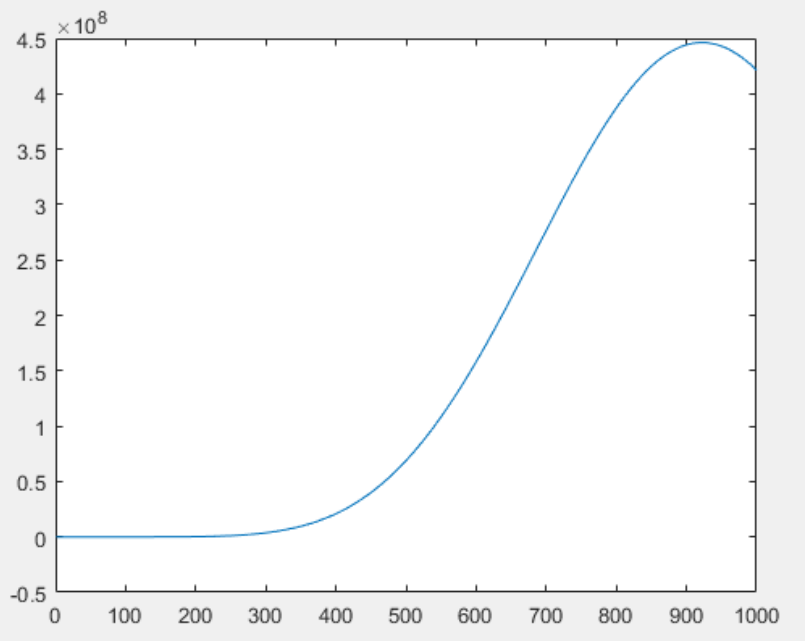


Figure 22: Unexpectedly misbehaving fourier series fit

This means fourier series overfit training data; the series regards given data as very short period, which force the series to change abruptly for very small period and result in

unreasonably huge coefficients. Hence, one of solution that we used in this research is to limit the minimum and maximum value for coefficients, which can be done in MatLab. From normal fourier series results, coefficients approximately have absolute values less than 1. More precisely, we then used upper bound and lower bound as 0.1 and -0.1 respectively. The final result after fixing bounds is shown below.

Moving Average

The volatility of each stock is a challenge to us because of its rapid changes in the price makes it harder to predict and see the outcome of our prediction. Stocks changes its price from hour to hour and definitely day to day; therefore, we can average these differences to reduce the 'noise' of these graphs to predict the future of the stock prices more accurately. Moving averages are created through constantly updated average price. Moreover, it can also help see the trend of how the stock is doing, such as if the price is increasing with time or decreasing with time.These averages are calculated through a period of time, such as five days. There are many durations of time which one can chose to calculate the moving average. Common lengths are ten days and above. However, we chose to do five days because our data set was chosen to a width of approximately a month. A short time frame will 'react' quicker to the changes of the stock price and reduce the time that the moving average takes to display a reversed trend in stock price. Moving averages is helpful and could be used for all types of investors - short term or long term- which is why it is one of the most used utility in stock trading.

The figures below 23, 24, and 25 demonstrate the use of moving averages and what it can do. As pointed out before, moving average causes a delay within the sign change in the prediction tool. Through using a moving average of 5 days to remove the noise in the stock price,we encountered a few day lag between each sign changes which can be seen by examining figure 23, 24, and 25. The major peaks are lagged and does not display the true

value of the stocks at the specific time. Even though the graph is smoother and can be calculated to predict the price more accurately, we are sacrificing the accuracy of real time stock value. This can be an issue if stock traders are not getting a real time display of what their investment is worth and may affect their earnings in a negative way. Moreover, another issue that we have to think about is the amount of days we want to average the price by to reduce noise. Too many days, the lag time will increase, and too little days then there will be the same amount of noise as the normal closing price.

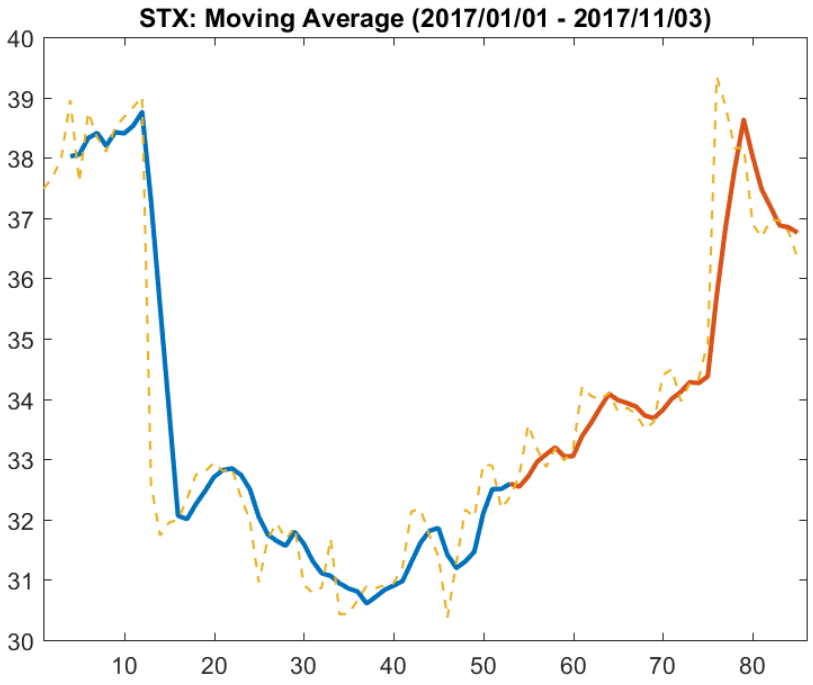


Figure 23: Seagate Technology Moving Average graph

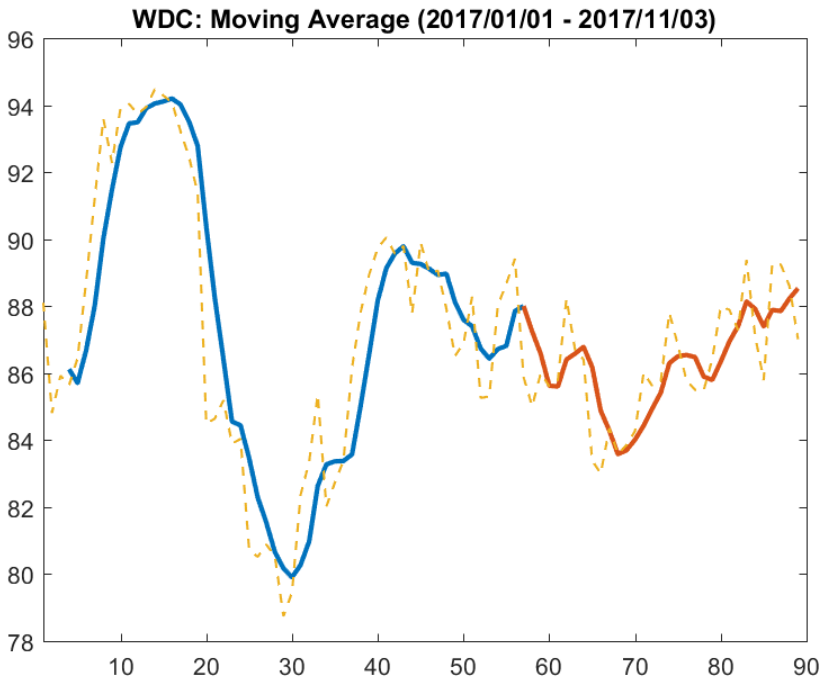


Figure 24: Western Digital Corporation Moving Average graph

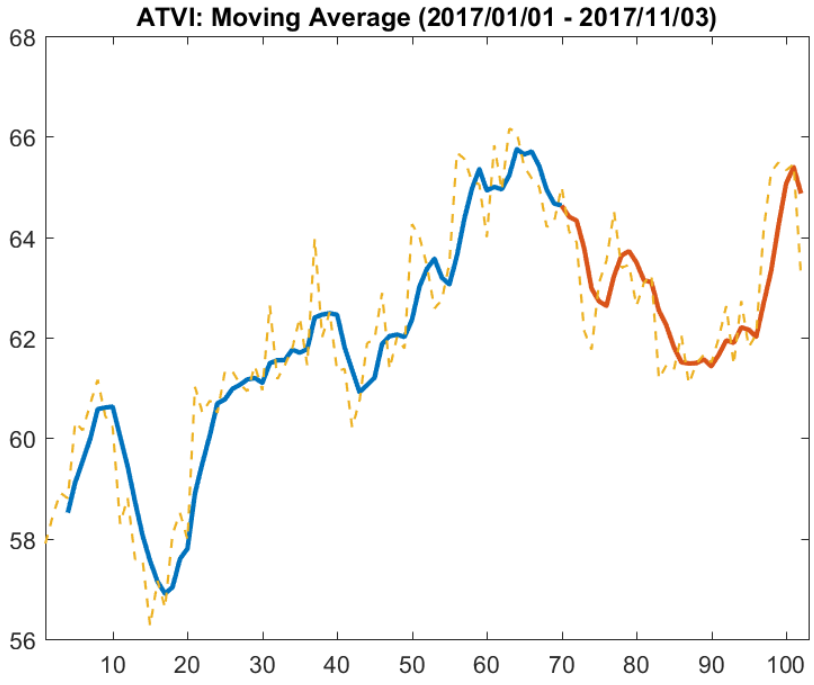


Figure 25: Activision Moving Average graph

For our scenario, we tested out these variations of moving averages between five and ten days and found that the best width, that has the least percent error and/or the most days that our model can predict, is five days, as seen in figure 26. Through calculating the percent error and counting the days of prediction it can be seen that moving average of 5 days shows the most accurate result with 3.14% while the percent error of raw data is higher. The percent error for moving average is 2.09%; it is indeed lower, however, capitalizing out of that 0.95% error yields minimal results compared to the amount of days moving average of five days can predict. By providing us with 3 more extra days of accurate prediction, we believe that it is worth that 0.95% increase in the error. However, our process is not that simple, we also faced issues were not using generates better results. Take stock 'SYMC' for example: the days of prediction for moving averages of five and ten days are zero compared to the raw data which is 32 days. This issue doesn't come up often, but is still an issue we have to fix in the future. The problem stems from the prediction with error bounds cutting off actual closing price as seen in figure 27, 28, and 29.

	Closing Price		Moving Average (5 days)		Moving Average (10 days)	
STOCK	Margin / Price percent	Days of Prediction	Margin / Price percent	Days of Prediction	Margin / Price percent	Days of Prediction
AAN	9.08%	9	4.66%	17	3.03%	1
BKFS	2.42%	11	2.04%	13	1.68%	15
BLKB	2.31%	15	1.86%	1	0.81%	2
DOX	1.14%	3	1.06%	3	0.72%	1
EBIX	2.72%	3	1.06%	0	0.85%	3
FTNT	2.54%	1	1.83%	2	0.88%	1
OTEX	1.36%	5	0.48%	2	0.26%	4
QADA	2.45%	2	1.19%	9	0.65%	6
TWOU	4.04%	1	2.35%	8	1.95%	0
ZEN	0.96%	5	0.74%	5	0.33%	4
CYBR	11.07%	8	5.16%	4	1.96%	3
EBAY	3.45%	32	2.84%	32	3.34%	32
GLW	5.83%	22	4.16%	26	2.41%	27
KYO	3.74%	9	3.15%	14	2.18%	0
QCOM	4.68%	2	3.38%	9	2.02%	0
SYMC	6.83%	32	5.41%	0	4.39%	0
T	4.06%	21	2.73%	24	1.33%	1
TRIP	9.56%	32	6.58%	32	5.31%	32
YELP	13.44%	9	9.02%	6	5.66%	3
Average	4.82%	11.7	3.14%	10.9	2.09%	7.1

Figure 26: Analysis of error bounds Table

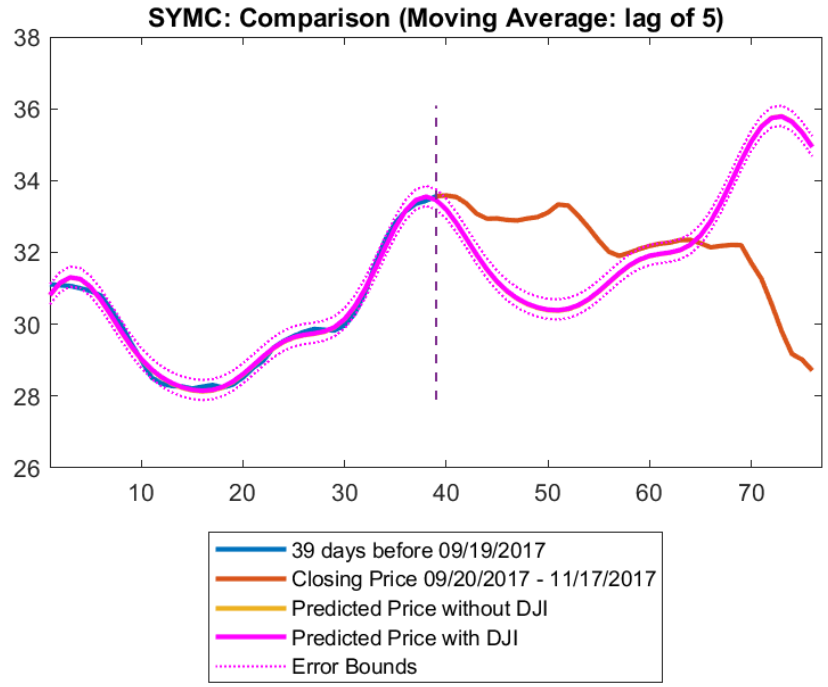


Figure 27: Prediction of Symantec stock with lag of 5

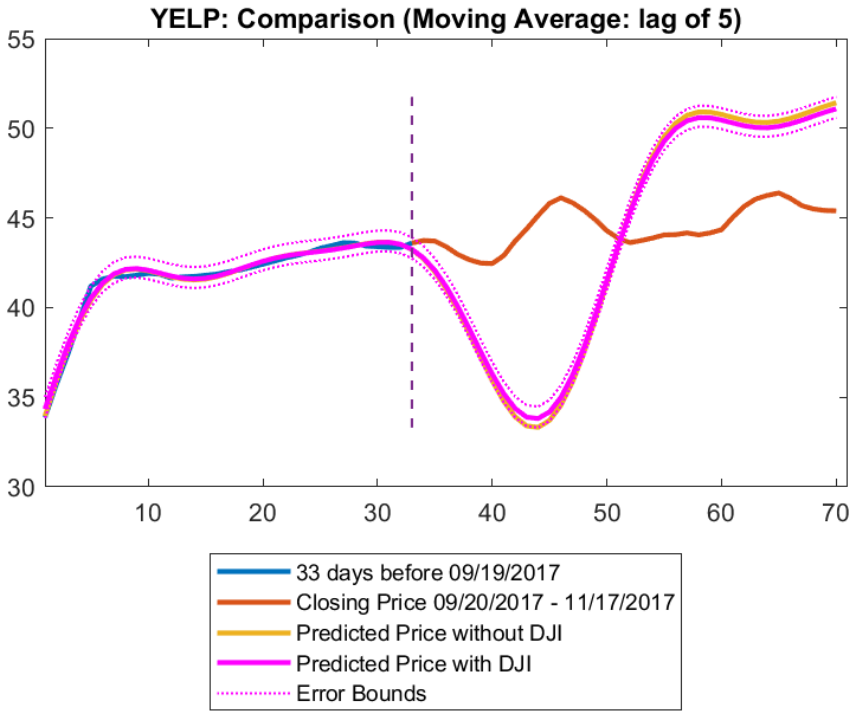


Figure 28: Prediction of Yelpwith lag of 5

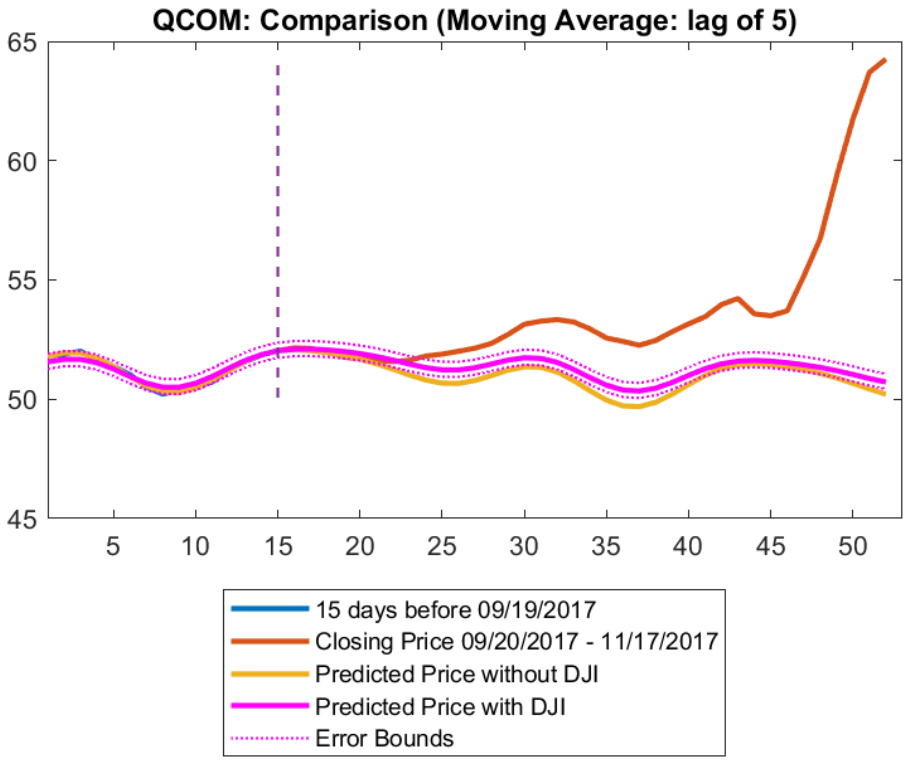


Figure 29: Prediction of Qualcomm with lag of 5

Nevertheless, the volatility still obviously exists. Rapid changes in stock price are considered to be high frequency points in its corresponding fourier transform. Hence, to get rid of high frequencies, we can apply low-pass filter, which only allows low frequencies (considered as smooth graph) to remain in the stock price. Specifically, the low-pass filter used for this module is called Butterworth filter. In figure 30, Western Digital Corp price stock gets smoothed.

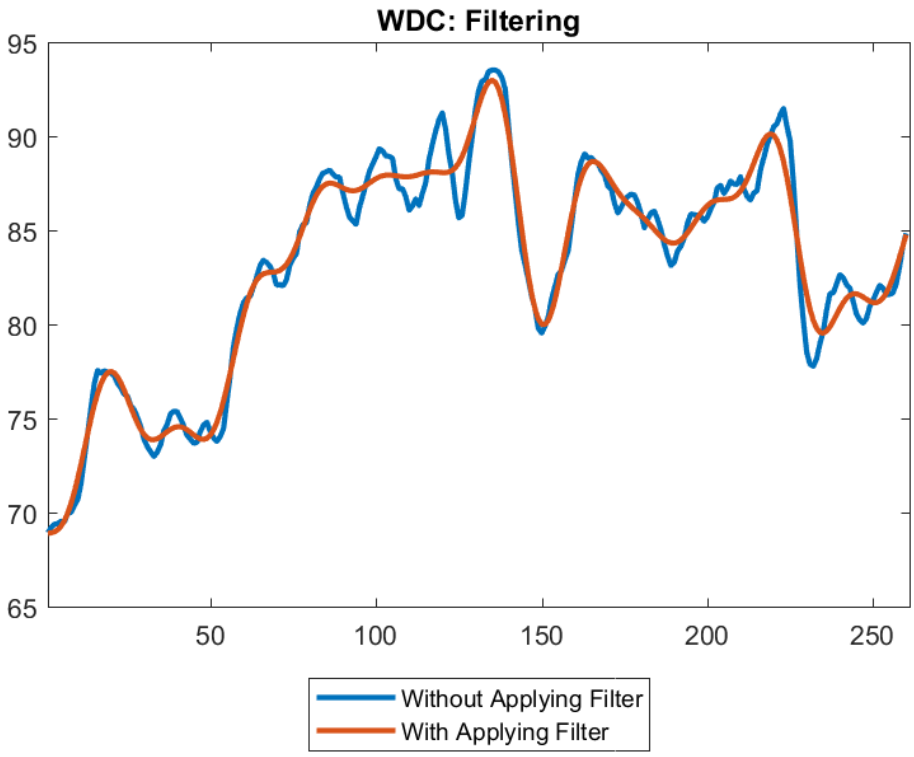


Figure 30: Western Digital Corp Price Comparison of with and without low-pass filter

However, it is not exactly right to state that filtering process makes predictions become more precise; there occurs all three possible outcomes in comparison of predictions before and after applying the filter as shown in figure from 31 to 33. It is worth noting that half of the sample stocks provide no significant change in comparison, while half of the remaining are better and the others are worse.

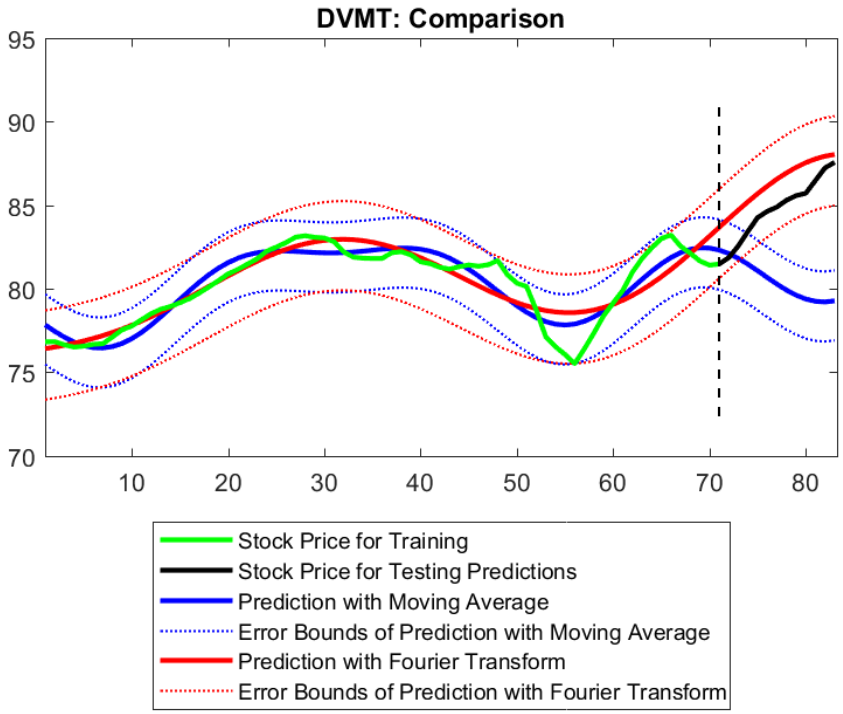


Figure 29: Dell Technologies Inc Prediction Comparison

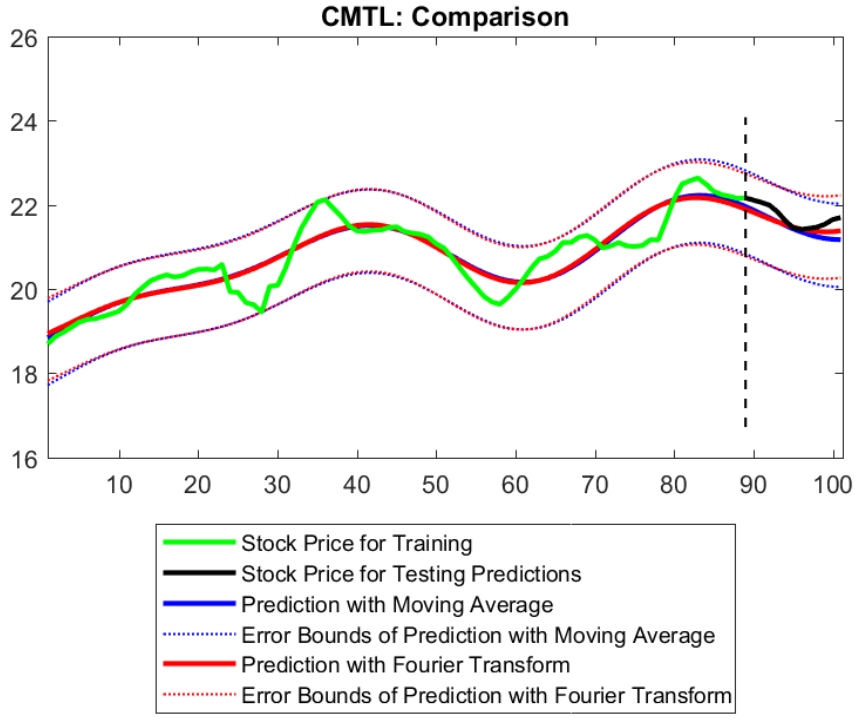


Figure 30: Comtech Telecomm. Corp. Prediction Comparison

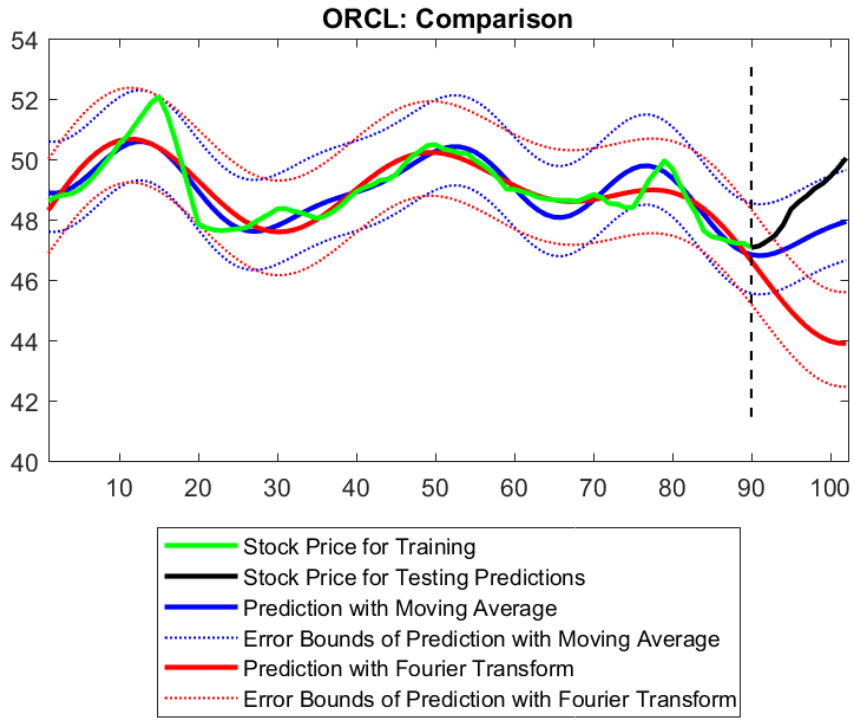


Figure 31: Oracle Corporation Prediction Comparison

NASDAQ and DOW JONES index

In the first quarter, we predict the stock price by finding the autocorrelation and fit the significant days into fourier series. This quarter we did the same thing but also applied it to NASDAQ and DOW JONES index (figure 30) and then found the correlation between the model function of the actual stock and these indexes. After that, we would get the correlation number range from 0 to 1, the higher the correlation number means that the indexes and stock share strong positive relationship. Then we could use the correlation number to put into the formula ,C * index + (1-c) * predicted stock, where C is the correction number. This is the formula that we use to plot the final model function.

	Closing Price						Closing Price with Moving Average (5 days)					
	Stock Only		With NASDAQ		With Dow Jones		Stock Only		With NASDAQ		With Dow Jones	
STOCK	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days
ATVI	7.201	8*	9.502	12*	12.67	37	4.864	6	5.625	11*	9.685	37*
CMTL	6.609	11*	6.657	12*	9.038	12*	4.338	8*	4.689	13*	5.090	13*
INTC	3.456	4	4.549	0	6.973	6	2.064	2	4.854	0	5.241	6
LOGI	11.61	37*	13.10	37*	12.44	37*	4.491	8	5.885	9*	7.258	37*
MSFT	3.769	22	5.955	21	7.244	19	1.681	12*	3.477	21	4.713	19
MSI	4.713	30*	5.311	21*	4.545	30*	2.529	13*	3.711	13*	2.613	12*
ORCL	10.09	2*	11.69	6*	10.11	3	5.656	5	8.223	9*	5.758	6
STX	19.63	4	20.60	6	24.05	12	13.55	2	13.97	3	19.29	24
WDC	8.319	12*	9.530	37	11.08	37*	4.716	2	4.703	1	10.01	3*

Figure 34: Table data of comparing NASDAQ and Dow Jones model

After we experimented using the moving average and got the results, we decided to add NASDAQ and DOW JONES index into the model function. This chart indicates that with moving average and NASDAQ/DOWJONES index, the gap between error boundaries were lessen compared to the one without moving average and NASDAQ/DOWJONES index. However, even though the indexes did not help lessen the boundaries error in the moving average section, the number of days for actual closing price inside the boundaries increased significantly for most of the stocks. As a result, it can be concluded that the indexes produced positive impact for this set of stock price. These graphs of ATVI (figure 31), CMTL(figure 32), LOGI (figure 33)and also support the claim that NASDAQ and DOW JONES were useful in predicting the prices, as we can see from the pictures that predicted price with DJI and NDXX indexes are closer to the actual prices than the prediction without the indexes.

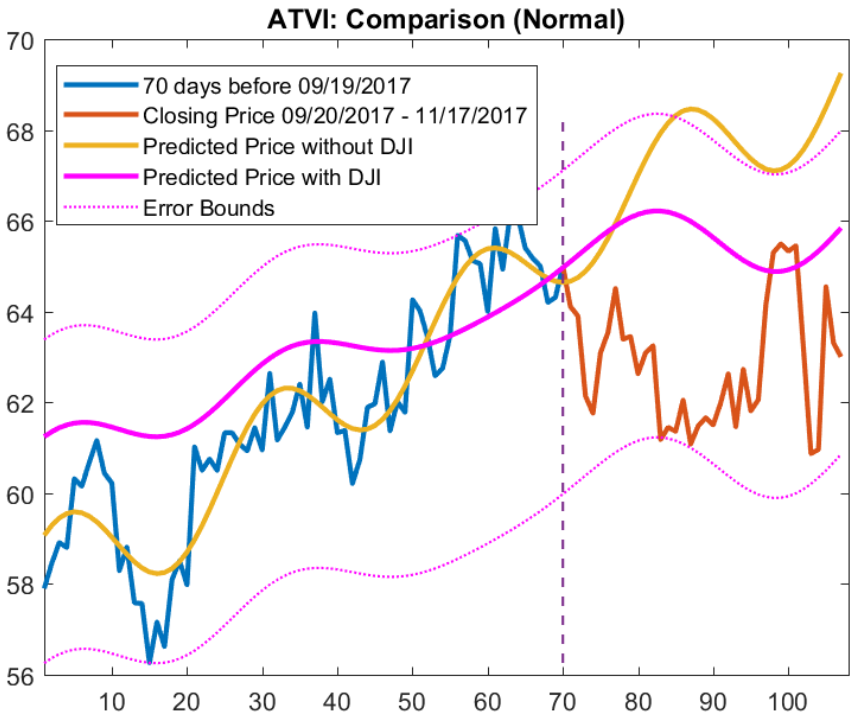


Figure 35: Activision stock prediction using Dow Jones model

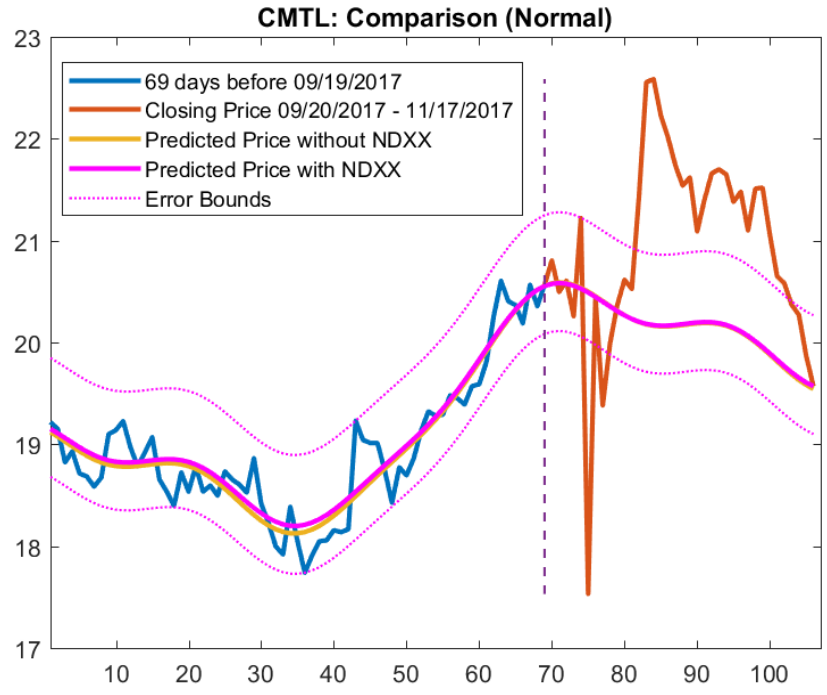


Figure 36: Comtech Telecommunications stock prediction using Dow Jones model

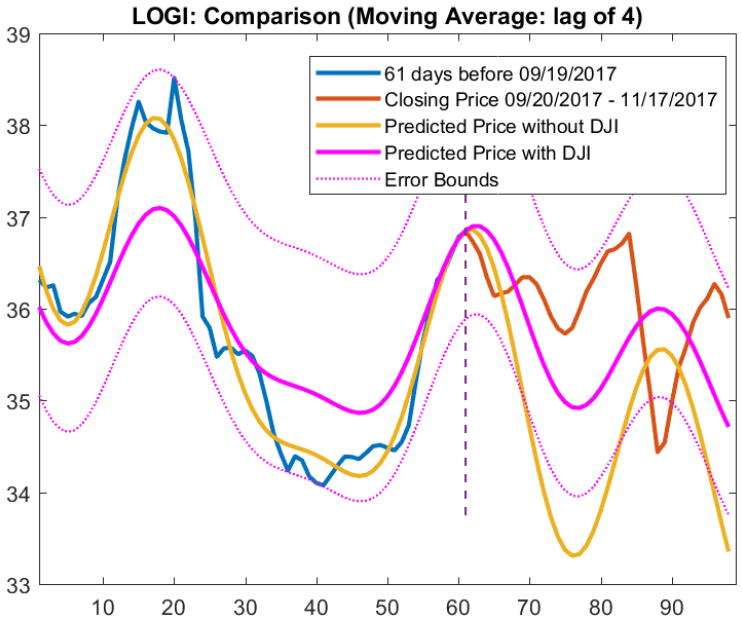


Figure 37: Logitech stock prediction using Dow Jones model

However, in some cases the indexes and moving average may not produced the expected results. Figure 34 suggests that there is no clear improvement when the indexes were included in the function for the closing price with moving average section, despite the data indicated improvement between the closing price and closing price with moving average section. This mean there is no effect when the indexes were added to the equation. This analysis conveys that there are still errors that limit the consistency of our function model.

	Closing Price						Closing Price with Moving Average (5 days)					
	Without NASDAQ		With NASDAQ		With DOW JONES		Without NASDAQ		With NASDAQ		With DOW JONES	
STOCK	Margin – Price percent	Days of Prediction	Margin – Price percent	Days of Prediction	Margin – Price percent	Days of Prediction	Margin – Price percent	Days of Prediction	Margin – Price percent	Days of Prediction	Margin – Price percent	Days of Prediction
CYBR	11.07%	9	23.29%	38	25.44%	38	5.16%	4	18.43%	38	20.57%	38
EBAY	3.45%	10	9.18%	8	8.10%	7	2.84%	10	6.43%	9	5.28%	7
GLW	5.83%	21	8.31%	21	10.55%	22	4.16%	11	5.61%	11	9.02%	21
KYO	3.74%	9	6.74%	12	6.74%	12	3.15%	14	5.73%	14	5.13%	13
QCOM	4.68%	3	4.26%	1	4.35%	2	3.38%	9	1.65%	2	1.26%	2
SYMC	6.83%	32	16.73%	30	14.96%	30	5.41%	32	10.45%	30	1.99%	30
T	4.06%	4	7.38%	6	8.08%	14	2.73%	3	2.96%	4	4.59%	16
TRIP	9.56%	9	13.80%	11	11.23%	4	6.58%	13	4.97%	11	7.28%	2
YELP	13.44%	9	25.15%	6	22.37%	5	9.02%	4	5.04%	8	3.45%	4
Average	6.96%	11.78	12.76%	14.78	12.42%	14.89	4.71%	11.11	6.81%	14.11	6.51%	14.11

Figure 38: Accurate prediction day of models comparison table of portfolio 1

We also tried changing the range of the date of the prediction from January 1, 2017 - September 19, 2017 to January 1, 2017 - October 15, 2017 to see if there are any improvements for the predicted days. However, during mid october to november, it was a new dividend opening date which caused the prices to be unstable. This directly impacts data prediction negatively.

	Closing Price						Closing Price with Moving Average (5 days)					
	Stock Only		With NASDAQ		With Dow Jones		Stock Only		With NASDAQ		With Dow Jones	
STOCK	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days	Percent Error (%)	Predicted Days
ATVI	7.395	2*	6.758	29*	7.895	29*	3.120	21*	3.855	20*	6.170	29*
CMTL	22.29	29	23.63	10	25.85	29*	9.973	4	11.12	4	14.59	12
INTC	3.026	2	9.728	0	10.17	4	1.415	1	8.536	2	8.384	3
LOGI	11.84	23*	13.29	29*	11.98	29*	4.695	17*	6.249	29*	5.182	24*
MSFT	5.110	5	7.852	0	6.999	4	2.426	2	4.043	12	4.198	12
MSI	3.979	3	5.153	5	4.094	5	2.106	4	3.930	4	3.403	7
ORCL	7.563	29*	8.037	29	8.831	29	6.053	12	6.667	29*	8.352	29
STX	7.492	6	10.54	3	9.395	4	6.361	7	6.994	3	8.783	4
WDC	7.820	2	7.877	3	7.209	3*	3.598	1	3.564	3	6.149	3
Avg	8.502	11.2	10.32	12	10.27	15.11	4.416	8	6.106	11.78	7.246	13.67

Figure 39: Accurate prediction day of models comparison table of portfolio 2

Figure 38 features the same set of stock from figure 39, but the period that actual price falls inside the error boundaries dropped significantly after the range was changed. This indicates that the model function works best when the stock prices are stable and not fluctuating.

Volatility

Volatility is the rapid change in the price of the stock. Through this, we came up with a hypothesis of using volatility to aid our model in the the effectiveness of the duration of days our model can accurately predict the stocks. The method that we used included taking a given stock's closing price and subtract it with its opening price to find the percentage mean of difference. Then we used the percentage mean difference to find the correlation with numbers of days our models can predict. If they correlate, it means that we can use volatility to estimate the days that our model can accurately predict the price of a given stock.

Our hypothesis states that volatility and accurately predicted days will share a negative relationship, meaning as the volatility of a stock is high, number of accurately predicted days will be low. In order to test our hypothesis, we decided to find the correlation between predicted days and volatility of our 20 stock using previous models of NASDAQ and DOW JONES. In theory, if our hypothesis is correct and prove that volatility could help estimate the number of days that we can accurately predict the stock prices, the correlation must be negative because we believe that as the volatility decreases, number of predicted days within the error boundaries must increase. The first attempt was to perform Pearson correlation of percent mean difference between closing price and opening price and accurately predicted days.

	DOW JONES			NASDAQ			Percent Mean of difference between closing and opening prices
Stock Name	Margin(Percent)	Day(s)	Day(s) within +- 5% error	Margin Percent	Day(s)	Day(s) within +- 5% error	
AAN	16.55	6	10	17.49	13	13	0.0041%
ATVI	6.26	3	5	5.86	2	5	0.0034%
CMTL	13.25	13	13	13.09	13	13	0.0044%
CSCO	6.94	4	5	8.23	2	5	0.0019%
CYBR	8.23	1	10	9.72	8	10	0.0035%
DOX	3.44	9	7	2.94	6	6	0.0018%
DVMT	7.58	4	6	7.67	5	5	0.0025%
EA	20.79	13	10	18.82	13	6	0.0030%
EBAY	7.01	10	10	7.6	10	10	0.0029%
EBIX	11.18	5	10	13.29	6	10	0.0029%
FTNT	6.77	3	8	7.97	3	6	0.0030%
INTC	10.57	7	7	12.2	7	13	0.0020%
LOGI	16.34	10	13	11.49	10	13	0.0024%
MCHP	3.23	8	7	3.17	8	9	0.0030%
ORCL	9.43	13	13	8.07	13	10	0.0020%
QCOM	6.33	0	2	1.77	4	5	0.0029%
SYMC	1.45	13	13	2.3	13	13	0.0032%
T	5.03	13	13	4.89	13	13	0.0023%
TWOU	25.32	13	4	26.41	13	3	0.0040%
WDC	13.71	13	13	9.58	13	13	0.0041%
Correlation with	**closing-opening**	**0.122**	**0.143**		**0.395**	**0.153**	

Figure 40: Correlation of volatility from closing and opening price and model predicted day

The results from figure 40 proves that our hypothesis is incorrect because the correlations between volatility and predicted days of 20 stocks are both positive, 0.153 for NASDAQ and 0.143 for DOW JONES model. This means that the two variables have a weak positive correlation from the given data instead of negative like our hypothesis stated.

As a result, we tried to change from difference of opening and closing prices to difference of maximum and minimum prices to see if there are any changes in the correlation.

	DOW JONES			NASDAQ			Percent Mean of difference between closing and opening prices	Percent Mean of difference between minimumand maximum prices
Stock Name	Margin(Percent)	Day(s)	Day(s) within +- 5% error	Margin Percent	Day(s)	Day(s) within +- 5% error		
AAN	16.55	6	10	17.49	13	13	0.0041%	0.0079%
ATVI	6.26	3	5	5.86	2	5	0.0034%	0.0071%
CMTL	13.25	13	13	13.09	13	13	0.0044%	0.0101%
CSCO	6.94	4	5	8.23	2	5	0.0019%	0.0041%
CYBR	8.23	1	10	9.72	8	10	0.0035%	0.0079%
DOX	3.44	9	7	2.94	6	6	0.0018%	0.0044%
DVMT	7.58	4	6	7.67	5	5	0.0025%	0.0052%
EA	20.79	13	10	18.82	13	6	0.0030%	0.0066%
EBAY	7.01	10	10	7.6	10	10	0.0029%	0.0059%
EBIX	11.18	5	10	13.29	6	10	0.0029%	0.0061%
FTNT	6.77	3	8	7.97	3	6	0.0030%	0.0067%
INTC	10.57	7	7	12.2	7	13	0.0020%	0.0047%
LOGI	16.34	10	13	11.49	10	13	0.0024%	0.0044%
MCHP	3.23	8	7	3.17	8	9	0.0030%	0.0060%
ORCL	9.43	13	13	8.07	13	10	0.0020%	0.0042%
QCOM	6.33	0	2	1.77	4	5	0.0029%	0.0055%
SYMC	1.45	13	13	2.3	13	13	0.0032%	0.0068%
T	5.03	13	13	4.89	13	13	0.0023%	0.0046%
TWOU	25.32	13	4	26.41	13	3	0.0040%	0.0081%
WDC	13.71	13	13	9.58	13	13	0.0041%	0.0088%
Correlation with	**closing-opening**	**0.122**	**0.143**		**0.395**	**0.153**		
	max-min	**0.14**	**0.179**		**0.362**	**0.155**		

Figure 41: Correlation of difference volatility methods and model predicted day Table

Figure 41 shows that there's a minor change in correlation from the result of using difference maximum and minimum, which the correlations still remain positive. The correlations between new volatility and predicted days are 0.155 for NASDAQ and 0.179 for DOW JONES model. This confirms that our hypothesis is still wrong, and we have sufficient evidence to conclude that difference of opening and closing prices and difference of maximum and minimum prices could not be used to estimate the number of days that could predict stock prices accurately.

Using Fourier transform to detect volatility of stock

In most of the stock analysis, stock price is often virtualized on a time domain graph showing the price is changing over time. However, stock price is one of financial time series which has unpredictable signals affected by external market and transient events. To overcome this problem, a frequency-domain has been used to analyze the bands and range of frequency, which usually classified as low and high, and period shifts. Sudden change of stock price directly increases the magnitude in high frequency band. As stated in previous section, we have a hypothesis that the amount of high frequency has an inverse correlation with the number of days that could predict stock prices accurately of our models.

Fourier transform is a famous function that alters the time domain to frequency domain, but still, containing all of original data, and can be reversely transform by using inverse Fourier transform. Fourier transform assumes that the input signal is stationary in some of frequency domain. However, stock price is known as a non-stationary time series. In this project, we use Fast Fourier transform(FFT) which yields real numbers alongs with imaginary numbers. The imaginary numbers can be transformed to real number value using Hilbert transform, but, in this project, all of imaginary part will be ignored.

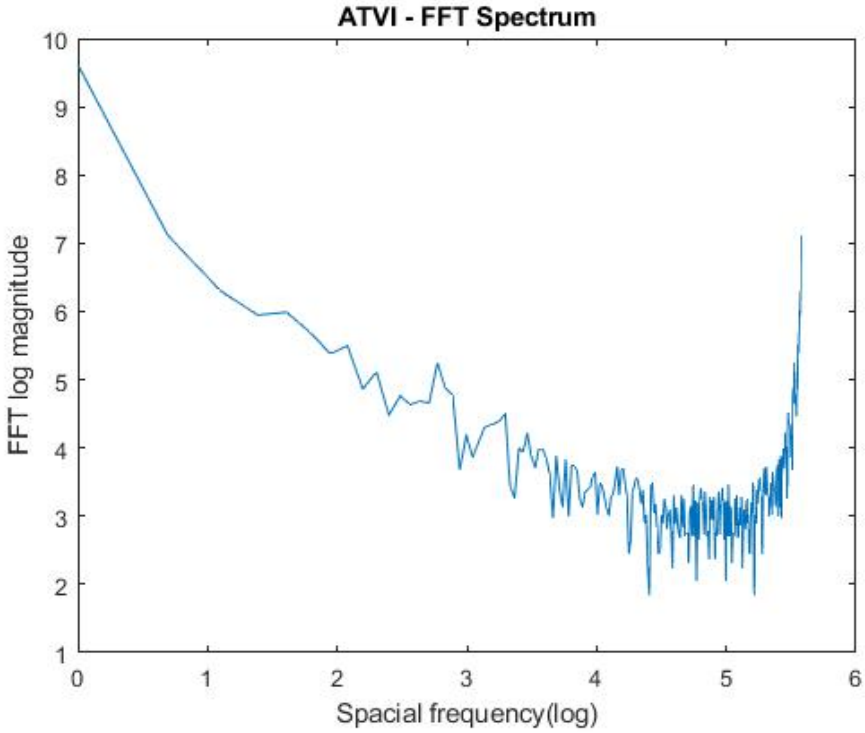


Figure 42: Spectrum of CMTL stocks in frequency domain

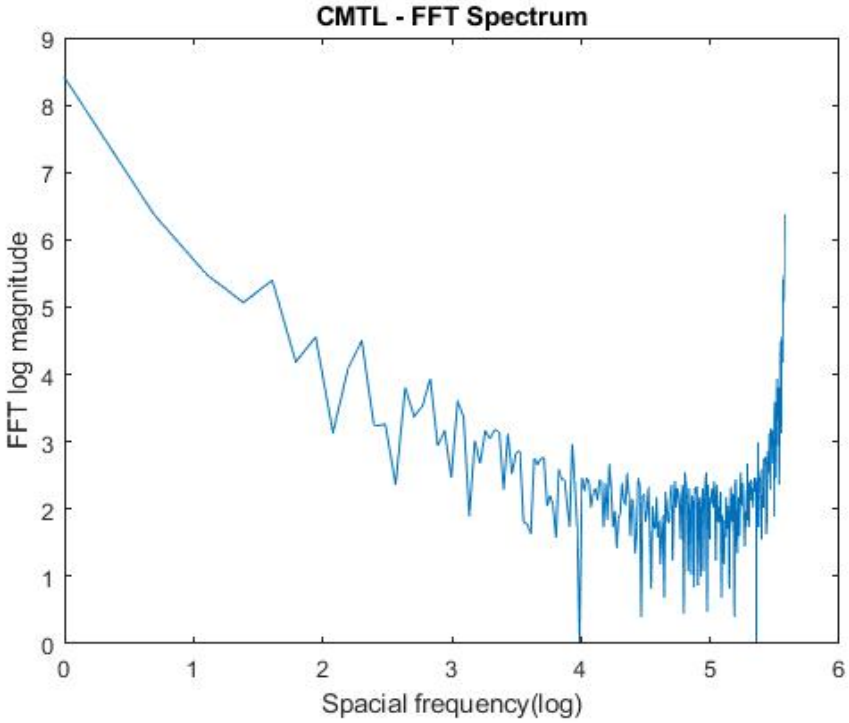


Figure 43: Spectrum of CMTL stocks in frequency domain

Figures 42 and 43 Fast Fourier Transform graphs showing Comtech Telecomm Corp. stock price in frequency domain with log scale and log magnitude. We have examine the graph of every stocks and choose 5(in log scale) is a cutting point for high frequency. Then, we calculate the summation of log magnitude in high frequency band, and use it as a new measurement of volatility. Similar to previous section, we conduct a correlation between summation of log magnitude in high frequency band and the number of days that could predict stock prices accurately of our models in a table below.

	DOW JONES			NASDAQ			Percent Mean of difference between closing and opening prices	Percent Mean of difference between minimumand maximum prices	Sum magnitude of log high FFT domain
Stock Name	Margin(Percent)	Day(s)	Day(s) within +- 5% error	Margin Percent	Day(s)	Day(s) within +- 5% error			
AAN	16.55	6	10	17.49	13	13	0.0041%	0.0079%	391.13
ATVI	6.26	3	5	5.86	2	5	0.0034%	0.0071%	493.91
CMTL	13.25	13	13	13.09	13	13	0.0044%	0.0101%	385.74
CSCO	6.94	4	5	8.23	2	5	0.0019%	0.0041%	402.61
CYBR	8.23	1	10	9.72	8	10	0.0035%	0.0079%	353.53
DOX	3.44	9	7	2.94	6	6	0.0018%	0.0044%	398.80
DVMT	7.58	4	6	7.67	5	5	0.0025%	0.0052%	495.59
EA	20.79	13	10	18.82	13	6	0.0030%	0.0066%	503.87
EBAY	7.01	10	10	7.6	10	10	0.0029%	0.0059%	375.46
EBIX	11.18	5	10	13.29	6	10	0.0029%	0.0061%	467.41
FTNT	6.77	3	8	7.97	3	6	0.0030%	0.0067%	424.52
INTC	10.57	7	7	12.2	7	13	0.0020%	0.0047%	373.20
LOGI	16.34	10	13	11.49	10	13	0.0024%	0.0044%	406.76
MCHP	3.23	8	7	3.17	8	9	0.0030%	0.0060%	496.67
ORCL	9.43	13	13	8.07	13	10	0.0020%	0.0042%	402.90
QCOM	6.33	0	2	1.77	4	5	0.0029%	0.0055%	384.37
SYMC	1.45	13	13	2.3	13	13	0.0032%	0.0068%	327.28
T	5.03	13	13	4.89	13	13	0.0023%	0.0046%	270.94
TWOU	25.32	13	4	26.41	13	3	0.0040%	0.0081%	518.15
WDC	13.71	13	13	9.58	13	13	0.0041%	0.0088%	447.28
Correlation with	closing-opening	0.122	0.143		0.395	0.153			
	max-min	0.14	0.179		0.362	0.155			
	high FFT	-0.083	-0.43		-0.209	-0.589			

Figure 44: Correlation of difference volatility methods and model predicted day Table

Figure 44 displays that the pearson correlation of log magnitude in high frequency band and the number of days that could predict stock prices with Nasdaq model accurately is -0.589, which is moderate correlation strength, within 5% error boundary. The estimated duration that model perform accurately can be calculated by a following linear equation.

$$\widehat{days} = 22.536 - 0.0324 \text{ (Sum magnitude of log high FFT)}$$

For Dow jones model, the correlation is -0.430, which is weaker but still considered as moderate correlation strength. However, the correlations for adaptive boundary are inadequate to be used, in both Nasdaq and Dow jones models.

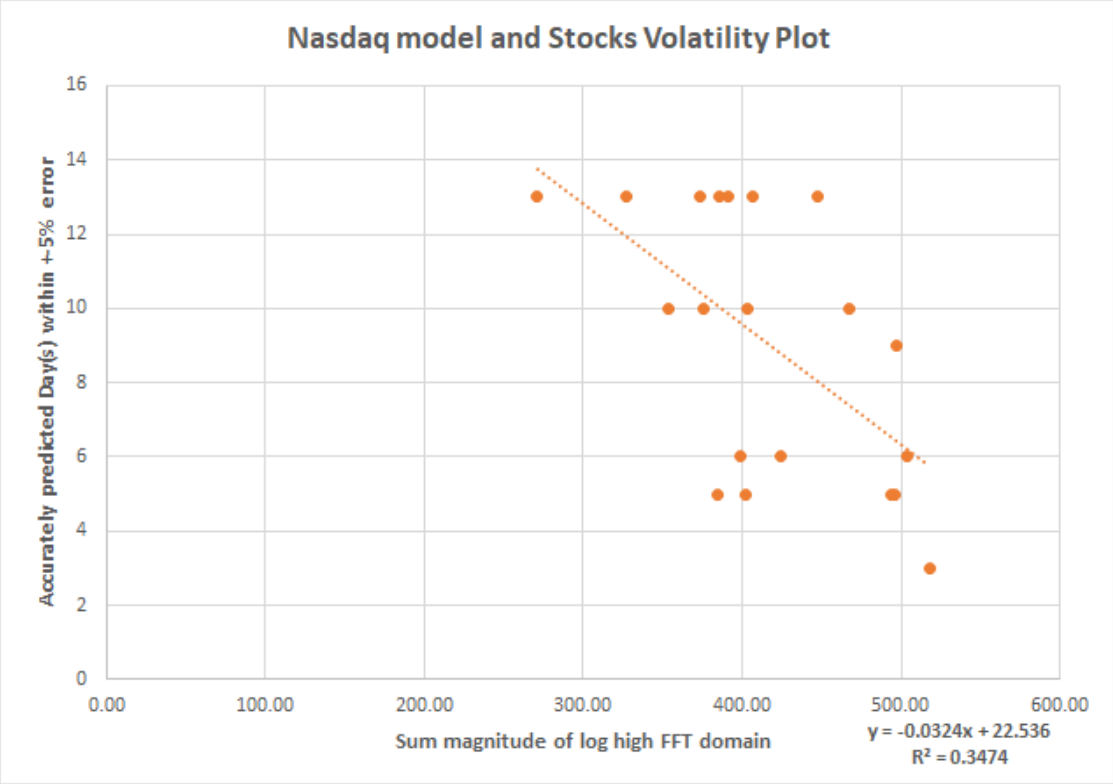


Figure 45: Fitting trend of Volatility and accurately predicted days

Recommendations for Future IQPs

As we proceed into this IQP project, we have encountered two major issues. The first issue has to do with fourier series and overfitting.

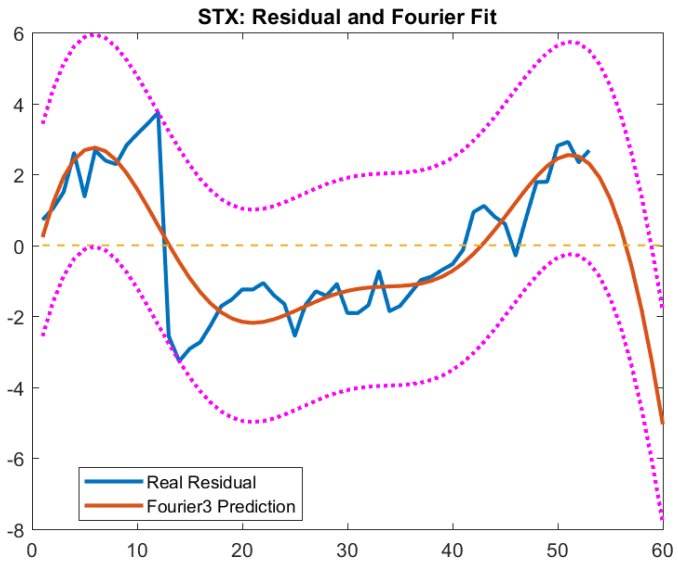


Figure 46: Misbehavior of Fourier Series fit

The fourier fit of STX here shows an unusual pattern that might be the result of overfitting. We couldn't find the source of this error or how to completely fix it yet. It might be an interesting topic for further topic for the new IQP project. Another issue has to with the consistency of our function model. The graph and data analysis shows different result from each set of stock. For example, the NASDAQ and DOW JONES index were able to improve one set of stock but produce an opposite result for another set, which we are still working to improve the error. Future IQPs, could do a further study on this issue and find out the source of the problem. We believe that improving the consistency could greatly enhance the performance of the function model.

Conclusion

Until this point, we are working toward our goal with satisfying results although there is still a problem with consistency. The methods that we learned like linear least square, fourier series, volatility, fourier transform or moving average seemed to be really helpful. The experiment to include NASDAQ and DOW JONES index also seemed to be working really well despite the fact that further study should be done between the correlation of these two indexes and stock prices. At the end of term C we can conclude that the best tool is to use the combination of fourier series, moving average and DOW JONES index. We were able to obtain the average boundaries error of 7.73% and the average days inside boundaries is 17.44 days from stock portfolio 4.

Appendix A: Relative Strength Indicator (RSI):

Stock price by itself is challenging to determine the transaction choice. Therefore, RSI was developed for measuring the speed and change of price movements; in other words, momentum oscillator of the price. The RSI value is vary between 0 and 100. The RSI can be simply considered overbought when the value raises above 70 ,and oversold when below 30. RSI value around 50 should not be used in any circumstances. Selling or Buying Signals in any transaction options can be inferred by looking for divergences and failure swings.

Appendix B: Matlab Codes

To execute the following Matlab codes, stocks files need to be downloaded as CSV format, and located in the same directory with Matlab codes.

a) Prediction with Fourier Series and NASDAQ correlation

```
file = dir('*.csv');
[refRawData, refPredictedConstant] = NASDAQData();
nBefore = 265 - 253; % To 01/02/2018
% assume that the first file is NASDAQ-100
for I = 1:length(file)
    %% Initialize
    stockName = file(I).name;   stockName = stockName(1:end-4);
    if(strcmp(stockName,'^IXIC'))
        continue;
    end
    temp = csvread(file(I).name, 1, 1);
    x = linspace(1, length(temp), length(temp))';
    y = temp(1:end,5);
    m = mean2(abs((temp(1:253,1)-temp(1:253,4))/temp(1:253,1)));
    hl = mean2(abs((temp(1:253,2)-temp(1:253,3))/temp(1:253,4)));
    % Normal
    % do nothing
    % Moving Average
    lag = 5;
    y = tsmovavg(y', 's', lag);
    y = y(lag: end)';
    yRawData = y;
    n = size(y);    n = n(1);   nPredicted = n - nBefore;


    %% find autocorrelation
    figureAutocorr = figure;
    title(strcat(stockName, ': Autocorrelation'));
    yCorr = autocorr(y(1: nPredicted), nPredicted - 1);
    autocorr(y(1: nPredicted), nPredicted - 1);
    % finding the intercept
    nst = 0;
    for i = 1: nPredicted - 1
        if(yCorr(i) * yCorr(i + 1) < 0)
            nst = nPredicted - i;
            break;
        end
    end

    % edit x and y to the range
```

```
    y = y(nst: end);
    n = length(y);
    nPredicted = n - nBefore;
    x = linspace(1, n, n)';

    text(nPredicted+5,0.1,strcat(num2str(nPredicted),' days'));
 %   saveas(figureAutocorr, strcat(stockName,'_autocorr.png'));

    % normalize y
    yNormal = y(nPredicted);
    y = y / yNormal;

    %% find fit
    coef = polyfit(x(1:nPredicted), y(1: nPredicted), 1);
    yFit = coef(1) * x + coef(2);   % length of n
    yRes = y - yFit;    % length of n
        %% find fourier
    fitFourier = fit(x(1: nPredicted), yRes(1: nPredicted), 'fourier3',...
        'Upper', [0.1 0.1 0.1 0.1 0.1 0.1 0.1 +Inf], 'Lower', [-0.1 -0.1 -0.1
-0.1 -0.1 -0.1 -0.1 -Inf]);
    yFourier = fitFourier(x);
    yPredicted = yFit + yFourier;

    %% find correlation

    corr = crosscorr(yRawData(nst: end - nBefore), refRawData(nst:end -
nBefore), 1);
    corr = abs(corr(2));

    nnn = 0;
    refPredicted = refPredictedConstant;
    if(length(refPredicted) < length(yPredicted))
        nnn = length(yPredicted) - length(refPredicted);
        yPredicted = yPredicted(nnn + 1: end);
    else
        nnn = length(refPredicted) - length(yPredicted);
        refPredicted = refPredicted(nnn + 1: end);
    end


    allPredicted = (corr * refPredicted + (1-corr) * yPredicted) * yNormal;

    %plot(x(end - nBefore: end), allPredicted, x(end - nBefore: end),
yRawData(end - nBefore: end));
    nn = length(allPredicted);
    x = linspace(1, nn, nn)';
    yRawData = yRawData(end - nn + 1: end);
```

```
    mxYAll = max(yRawData(1: end - nBefore) - allPredicted(1: end -
nBefore));
    mnYAll = min(yRawData(1: end - nBefore) - allPredicted(1: end -
nBefore));
    marginPercent = (mxYAll - mnYAll) / min(yRawData);

    fprintf('%s\t%f\t%f\t%f\n',stockName,marginPercent*100,m,hl);

    mxY = max(max(yRawData));
    mnY = min(min(yRawData));
    x0920 = (nn - nBefore) * ones(30, 1);
    y0920 = linspace(mnY, mxY, 30);


    mxY = max(max(allPredicted + mxYAll));
    mnY = min(min(allPredicted + mnYAll));
    x0920 = (nn - nBefore) * ones(30, 1);
    y0920 = linspace(mnY, mxY, 30);
    figurePredict = figure;
    p = plot(x(1:x0920(1)), allPredicted(1:x0920(1)),...
        x(x0920(1):end), allPredicted(x0920(1):end), 'm',...
        x, allPredicted + mxYAll, ':m',...
        x, allPredicted + mnYAll, ':m',...
        x0920, y0920, '--',...
        'LineWidth', 2);
    p(3).LineWidth = 1;
    p(4).LineWidth = 1;
    p(5).LineWidth = 1;
    xlim([1 nn+1]);
    title(strcat(stockName, ': Prediction (Moving Average: lag of 5)'));
    ll = strcat(num2str(nn - nBefore),' days before 01/02/2018');
    legend(ll,'Predicted price 01/02/2018 - 01/19/2018' ,'Error Bounds',
'Location', 'southoutside')
  %  saveas(figurePredict, strcat(stockName, '_movavg5_predict.png'));

        mxY = max(max(max(yRawData), max(allPredicted + mxYAll)));
    mnY = min(min(min(yRawData), min(allPredicted + mnYAll)));
    x0920 = (nn - nBefore) * ones(30, 1);
    y0920 = linspace(mnY, mxY, 30);
    figureTogether = figure;
    p = plot(x(1:x0920(1)), yRawData(1:x0920(1)),...
        x(x0920(1):end), yRawData(x0920(1):end),...
        x,yPredicted* yNormal,...
        x, allPredicted,'m',...
        x, allPredicted + mxYAll, ':m',...
        x, allPredicted + mnYAll, ':m',...
        x, allPredicted - 0.05*allPredicted , ':b',...
```

```
        x, allPredicted + 0.05*allPredicted, ':b',...
        x0920, y0920, '--', ...
        'LineWidth', 2);

    p(5).LineWidth = 1;
    p(6).LineWidth = 1;
    p(7).LineWidth = 1;
    p(8).LineWidth = 1;
    p(9).LineWidth = 1;
    xlim([1 nn+1]);
    title(strcat(stockName, ': Comparison (Moving Average: lag of 5)'));
    ll = strcat(num2str(nn - nBefore),' days before 01/02/2018');
    legend(ll,'Closing Price 01/02/2018 - 01/19/2018' ,'Predicted Price
without NASDAQ','Predicted Price with NASDAQ','Error Bounds', 'Location',
'southoutside')
    saveas(figureTogether, strcat(stockName,
'_movavg5_together_nasdaq.png'));
end
close all;
```

```
function [yAllData, yPredicted] = NASDAQData()


    file = dir('^IXIC.csv');

    if length(file) <= 0
        fprintf('Please download NASDAQ-100 \n');
        return;
    end


    %% Initialize

    dayBeforePredicted = 265 - 253; % To 01/02/2018

    temp = csvread(file(1).name, 1, 1);
    x = linspace(1, length(temp), length(temp))';
    y = temp(1:end,5);
    yAllData = y;
    % Normal
    % do nothing

    % Moving Average
    lag = 5;
    y = tsmovavg(y', 's', lag);
```

```
    y = y(lag: end)';

    n = size(y);    n = n(1);   nPredicted = n - dayBeforePredicted;


    %% find autocorrelation
    figureAutocorr = figure;
    autocorr(y(1: nPredicted), nPredicted - 1);
    title(strcat('INDEXNASDAQ', ': Autocorrelation'));
    yCorr = autocorr(y(1: nPredicted), nPredicted - 1);

    % finding the intercept
    nst = 0;
    for i = 1: nPredicted - 1
        if(yCorr(i) * yCorr(i + 1) < 0)
            nst = nPredicted - i;
            break;
        end
    end

    % edit x and y to the range
    y = y(nst: end);
    n = length(y);
    nPredicted = n - dayBeforePredicted;
    x = linspace(1, n, n)';
    % normalize y
    y = y / y(nPredicted);

    text(nPredicted+5,0.1,strcat(num2str(nPredicted),' days'));
    saveas(figureAutocorr, strcat('NASDAQ_autocorr.png'));

    %% find fit
    coef = polyfit(x(1:nPredicted), y(1: nPredicted), 1);
    yFit = coef(1) * x + coef(2);   % length of n
    yRes = y - yFit;    % length of n


    %% find fourier
    fitFourier = fit(x(1: nPredicted), yRes(1: nPredicted), 'fourier3',...
        'Upper', [0.1 0.1 0.1 0.1 0.1 0.1 0.1 +Inf], 'Lower', [-0.1 -0.1 -0.1
-0.1 -0.1 -0.1 -0.1 -Inf]);
    yFourier = fitFourier(x);
    yPredicted = yFit + yFourier;
    mxRes = max(max(yRes(1: nPredicted) - yFourier(1: nPredicted)));
    mnRes = min(min(yRes(1: nPredicted) - yFourier(1: nPredicted)));

    % plot fit
```

```
    figureFit = figure;
    p = plot(x(1: nPredicted), y(1: nPredicted),...
        x(nPredicted: end), y(nPredicted: end),...
         x, yFit + yFourier ,...
        x, yPredicted + mxRes , ':m',...
        x, yPredicted + mnRes , ':m',...
        'LineWidth', 2);
    title(strcat('NASDAQ : Least Square Error '));
    legend('Closing Price 01/01/2017 - 12/29/2017','Closing Price 01/02/2018
- 01/19/2018' ,'Linear Least Squares','Predicted Price','Error Bounds',
'Location', 'southoutside')
    saveas(figureFit, strcat('NASDAQ_fit_n.png'));
    %close all;

end
```

b) Prediction with Fourier Series and DOW JONES correlation

```
file = dir('*.csv');

[refRawData, refPredictedConstant] = DJIData();

nBefore = 265 - 253; % To 01/02/2018

for I = 1:length(file)
    %% Initialize
    stockName = file(I).name;   stockName = stockName(1:end-4);
    if(strcmp(stockName,'^DJI'))
        continue;
    end
    temp = csvread(file(I).name, 1, 1);
    x = linspace(1, length(temp), length(temp))';
    y = temp(1:end,5);
    m = mean2(abs((temp(1:253,1)-temp(1:253,4))/temp(1:253,1)));
    % Normal
    % do nothing

    % Moving Average
    lag = 5;
    y = tsmovavg(y', 's', lag);
    y = y(lag: end)';

    yRawData = y;
    n = size(y);    n = n(1);   nPredicted = n - nBefore;
```

```
    %% find autocorrelation
    figureAutocorr = figure;
    title(strcat(stockName, ': Autocorrelation'));
    yCorr = autocorr(y(1: nPredicted), nPredicted - 1);
    autocorr(y(1: nPredicted), nPredicted - 1);
    % finding the intercept
    nst = 0;
    for i = 1: nPredicted - 1
        if(yCorr(i) * yCorr(i + 1) < 0)
            nst = nPredicted - i;
            break;
        end
    end

    % edit x and y to the range
    y = y(nst: end);
    n = length(y);
    nPredicted = n - nBefore;
    x = linspace(1, n, n)';

    text(nPredicted+5,0.1,strcat(num2str(nPredicted),' days'));
%   saveas(figureAutocorr, strcat(stockName,'_autocorr.png'));

    % normalize y
    yNormal = y(nPredicted);
    y = y / yNormal;


    %% find fit
    coef = polyfit(x(1:nPredicted), y(1: nPredicted), 1);
    yFit = coef(1) * x + coef(2);   % length of n
    yRes = y - yFit;    % length of n


    %% find fourier
    fitFourier = fit(x(1: nPredicted), yRes(1: nPredicted), 'fourier3',...
        'Upper', [0.1 0.1 0.1 0.1 0.1 0.1 0.1 +Inf], 'Lower', [-0.1 -0.1 -0.1
-0.1 -0.1 -0.1 -0.1 -Inf]);
    yFourier = fitFourier(x);
    yPredicted = yFit + yFourier;

    %% find correlation

    corr = crosscorr(yRawData(nst: end - nBefore), refRawData(nst:end -
nBefore), 1);
```

```
    corr = abs(corr(2));

    nnn = 0;
    refPredicted = refPredictedConstant;
    if(length(refPredicted) < length(yPredicted))
        nnn = length(yPredicted) - length(refPredicted);
        yPredicted = yPredicted(nnn + 1: end);
    else
        nnn = length(refPredicted) - length(yPredicted);
        refPredicted = refPredicted(nnn + 1: end);
    end


    allPredicted = (corr * refPredicted + (1-corr) * yPredicted) * yNormal;

    %plot(x(end - nBefore: end), allPredicted, x(end - nBefore: end),
yRawData(end - nBefore: end));
    nn = length(allPredicted);
    x = linspace(1, nn, nn)';
    yRawData = yRawData(end - nn + 1: end);
    mxYAll = max(yRawData(1: end - nBefore) - allPredicted(1: end -
nBefore));
    mnYAll = min(yRawData(1: end - nBefore) - allPredicted(1: end -
nBefore));
    marginPercent = (mxYAll - mnYAll) / min(yRawData);

    fprintf('%s\t%f\t%f\n',stockName,marginPercent*100,m);

    mxY = max(max(yRawData));
    mnY = min(min(yRawData));
    x0920 = (nn - nBefore) * ones(30, 1);
    y0920 = linspace(mnY, mxY, 30);


    mxY = max(max(allPredicted + mxYAll));
    mnY = min(min(allPredicted + mnYAll));
    x0920 = (nn - nBefore) * ones(30, 1);
    y0920 = linspace(mnY, mxY, 30);
    figurePredict = figure;
    p = plot(x(1:x0920(1)), allPredicted(1:x0920(1)),...
        x(x0920(1):end), allPredicted(x0920(1):end), 'm',...
        x, allPredicted + mxYAll, ':m',...
        x, allPredicted + mnYAll, ':m',...
        x0920, y0920, '--',...
        'LineWidth', 2);
    p(3).LineWidth = 1;
    p(4).LineWidth = 1;
```

```
    p(5).LineWidth = 1;
    xlim([1 nn+1]);
    title(strcat(stockName, ': Prediction (Moving Average: lag of 5)'));
    ll = strcat(num2str(nn - nBefore),' days before 01/02/2018');
    legend(ll,'Predicted price 01/02/2018 - 01/19/2018' ,'Error Bounds',
'Location', 'southoutside');
   % saveas(figurePredict, strcat(stockName, '_movavg5_predict.png'));



    mxY = max(max(max(yRawData), max(allPredicted + mxYAll)));
    mnY = min(min(min(yRawData), min(allPredicted + mnYAll)));
    x0920 = (nn - nBefore) * ones(30, 1);
    y0920 = linspace(mnY, mxY, 30);
    figureTogether = figure;
      p = plot(x(1:x0920(1)), yRawData(1:x0920(1)),...
        x(x0920(1):end), yRawData(x0920(1):end),...
        x,yPredicted* yNormal,...
        x, allPredicted,'m',...
        x, allPredicted + mxYAll, ':m',...
        x, allPredicted + mnYAll, ':m',...
        x, allPredicted - 0.05*allPredicted , ':b',...
        x, allPredicted + 0.05*allPredicted, ':b',...
        x0920, y0920, '--', ...
        'LineWidth', 2);

    p(5).LineWidth = 1;
    p(6).LineWidth = 1;
    p(7).LineWidth = 1;
    p(8).LineWidth = 1;
    p(9).LineWidth = 1;
    xlim([1 nn+1]);
    title(strcat(stockName, ': Comparison (Moving Average: lag of 5)'));
    ll = strcat(num2str(nn - nBefore),' days before 01/02/2018');
    legend(ll,'Closing Price 01/02/2018 - 01/19/2018' ,'Predicted Price
without DJI','Predicted Price with DJI','Error Bounds', 'Location',
'southoutside')
    saveas(figureTogether, strcat(stockName, '_movavg5_dow.png'));

end

close all;
```

```
function [yAllData, yPredicted] = NASDAQData()
```

```
file = dir('^DJI.csv');

if length(file) <= 0
    fprintf('Please download DJI \n');
    return;
end


%% Initialize

dayBeforePredicted = 265 - 253; % To 01/02/2018

temp = csvread(file(1).name, 1, 1);
x = linspace(1, length(temp), length(temp))';
y = temp(1:end,5);
yAllData = y;
% Normal
% do nothing

% Moving Average
lag = 5;
y = tsmovavg(y', 's', lag);
y = y(lag: end)';

n = size(y);    n = n(1);   nPredicted = n - dayBeforePredicted;


%% find autocorrelation
figureAutocorr = figure;
autocorr(y(1: nPredicted), nPredicted - 1);
title(strcat('DJI', ': Autocorrelation'));
yCorr = autocorr(y(1: nPredicted), nPredicted - 1);

% finding the intercept
nst = 0;
for i = 1: nPredicted - 1
    if(yCorr(i) * yCorr(i + 1) < 0)
        nst = nPredicted - i;
        break;
    end
end

% edit x and y to the range
y = y(nst: end);
n = length(y);
```

```
    nPredicted = n - dayBeforePredicted;
    x = linspace(1, n, n)';
    % normalize y
    y = y / y(nPredicted);

    text(nPredicted+5,0.1,strcat(num2str(nPredicted),' days'));
    saveas(figureAutocorr, strcat('DJI_autocorr.png'));

    %% find fit
    coef = polyfit(x(1:nPredicted), y(1: nPredicted), 1);
    yFit = coef(1) * x + coef(2);   % length of n
    yRes = y - yFit;    % length of n


    %% find fourier
    fitFourier = fit(x(1: nPredicted), yRes(1: nPredicted), 'fourier3',...
        'Upper', [0.1 0.1 0.1 0.1 0.1 0.1 0.1 +Inf], 'Lower', [-0.1 -0.1 -0.1
-0.1 -0.1 -0.1 -0.1 -Inf]);
    yFourier = fitFourier(x);
    yPredicted = yFit + yFourier;
    mxRes = max(max(yRes(1: nPredicted) - yFourier(1: nPredicted)));
    mnRes = min(min(yRes(1: nPredicted) - yFourier(1: nPredicted)));

    % plot fit
    figureFit = figure;
    p = plot(x(1: nPredicted), y(1: nPredicted),...
        x(nPredicted: end), y(nPredicted: end),...
         x, yFit + yFourier ,...
        x, yPredicted + mxRes , ':m',...
        x, yPredicted + mnRes , ':m',...
        'LineWidth', 2);
    title(strcat('NASDAQ : Least Square Error '));
    legend('Closing Price 01/01/2017 - 12/29/2017','Closing Price 01/02/2018
- 01/19/2018' ,'Linear Least Squares','Predicted Price','Error Bounds',
'Location', 'southoutside')
    saveas(figureFit, strcat('DJI_fit_n.png'));
    %close all;

end
```

c) Prediction from moving average with and without Fourier Series Smoothing

```
%{
	The purpose of this program is to compare the prediction results with
	and without applying Fast Fourier Transform (FFT) to moving average of
	5 days
%}

% find all data files
files = dir('*.csv');

% the number of days predicted
nBefore = 265 - 253; % To 01/02/2018

% iterate the files
for I = 1:length(files)
	%% Initialize
	stockName = files(I).name;   stockName = stockName(1:end-4);
	if(strcmp(stockName,'^IXIC')) % ignore index stock
		continue;
	end
	temp = csvread(files(I).name, 1, 1);
	stockRawData = temp(1:end,5);  % data


	%% Prepare Data
	filtering = figure;
	% Moving Average
	lag = 5;
	yMovAvg = tsmovavg(stockRawData', 's', lag);
	yMovAvg = yMovAvg(lag: end)';
	rawMovAvg = yMovAvg;


	%{
	% Fast Fourier Transform
	fft_values = fft(yMovAvg);
	% get rid of values that have high frequencies
	fft_values(exp(5): end) = 0;
	yFFT = ifft(fft_values);
	%}

	 % Butterworth filter of order 6
	fc = exp(5); % Cut off frequency
	fs = 264 * 6; % Sampling rate
	%[b, a] = butter(6, fc / (fs / 2));
	[b, a] = butter(6, 0.1);
```

```
yFFT = filtfilt(b, a, yMovAvg);
%yFFT = [yFFT(2:end); yMovAvg(end)];

% observe graph
n = length(yFFT);
x = linspace(1,n,n)';
plot(x, yMovAvg, x, yFFT(1:end), 'LineWidth', 2);
legend('Without Applying Filter', 'With Applying Filter',...
       'Location', 'southoutside');
xlim([1 n + 1]);
title(strcat(stockName, ': Filtering'));
saveas(filtering, strcat(stockName, '_filtering.png'));

%{
% wavelet denoising
yFFT = cmddenoise(yMovAvg, 'db3', 3);
yFFT = yFFT';
%}


% find the number of days to use as predicting base
n = size(yFFT);    n = n(1);
nPred = n - nBefore;


%% Find Autocorrelation
% -> moving average
corrMovAvg = autocorr(yMovAvg(1: nPred), nPred - 1);
% finding the intercept
n = 0;
for i = 1: nPred - 1
       if(corrMovAvg(i) * corrMovAvg(i + 1) < 0)
              n = nPred - i;
              break;
       end
end
% edit range to fit better
yMovAvg = yMovAvg(n: end);
n = length(yMovAvg);
nPredMovAvg = n - nBefore;
xMovAvg = linspace(1, n, n)';
% normalize data
normValMovAvg = yMovAvg(nPredMovAvg);
yMovAvg = yMovAvg / normValMovAvg;

% -> fast fourier transform
corrFFT = autocorr(yFFT(1: nPred), nPred - 1);
```

```
	% finding the intercept
	n = 0;
	for i = 1: nPred - 1
		if(corrFFT(i) * corrFFT(i + 1) < 0)
			n = nPred - i;
			break;
		end
	end
	% edit range to fit better
	yFFT = yFFT(n: end);
	n = length(yFFT);
	nPredFFT = n - nBefore;
	xFFT = linspace(1, n, n)';
	% normalize data
	normValFFT = yFFT(nPredFFT);
	yFFT = yFFT / normValFFT;


	%% Find Linear Fit for Data
	% -> moving average
	coef = polyfit(xMovAvg(1: nPredMovAvg), yMovAvg(1: nPredMovAvg), 1); %
predicting data
	fitMovAvg = coef(1) * xMovAvg + coef(2); % whole data
	resMovAvg = yMovAvg - fitMovAvg; % whole data

	% -> fast fourier transform
	coef = polyfit(xFFT(1: nPredFFT), yFFT(1: nPredFFT), 1); % predicting data
	fitFFT = coef(1) * xFFT + coef(2); % whole data
	resFFT = yFFT - fitFFT; % whole data


	%% Find Fourier Fit for Residual
	% -> moving average
	fourier = fit(xMovAvg(1: nPredMovAvg), resMovAvg(1: nPredMovAvg), 'fourier3',...
			'Upper', [0.1 0.1 0.1 0.1 0.1 0.1 0.1 +Inf],...
			'Lower', [-0.1 -0.1 -0.1 -0.1 -0.1 -0.1 -0.1 -Inf]);
	fourMovAvg = fourier(xMovAvg);
	predMovAvg = (fitMovAvg + fourMovAvg) * normValMovAvg;

	% -> fast fourier transform
	fourier = fit(xFFT(1: nPredFFT), resFFT(1: nPredFFT), 'fourier3',...
			'Upper', [0.1 0.1 0.1 0.1 0.1 0.1 0.1 +Inf],...
			'Lower', [-0.1 -0.1 -0.1 -0.1 -0.1 -0.1 -0.1 -Inf]);
	fourFFT = fourier(xFFT);
	predFFT = (fitFFT + fourFFT) * normValFFT;

	% Note: now predMovAvg and predFFT are in normal scale
```

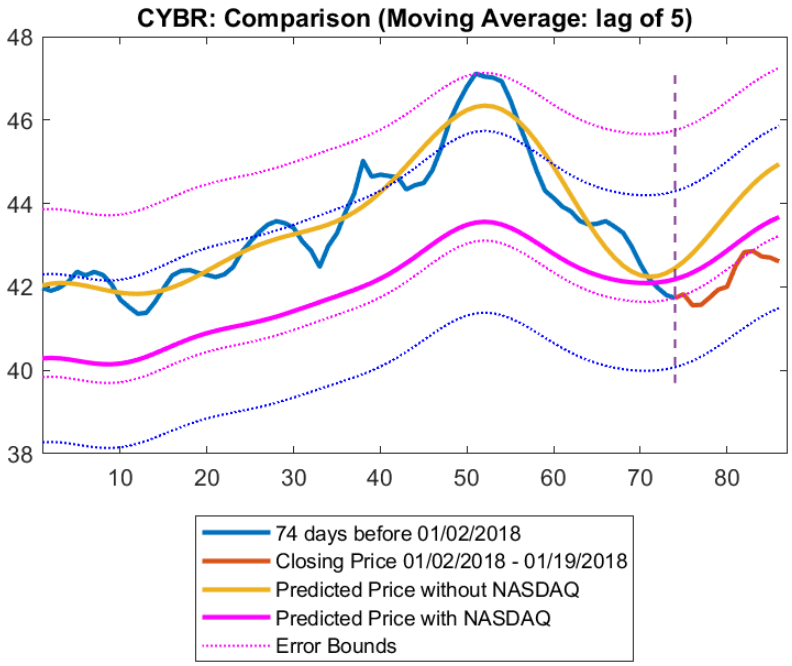
```
	%% Find Error Bounds
	% -> moving average
	n = length(predMovAvg);
	rawTemp = rawMovAvg(end - n + 1: end);
	plusDiffMovAvg = max(rawTemp(1: end - nBefore) - predMovAvg(1: end -
nBefore));
	minusDiffMovAvg = min(rawTemp(1: end - nBefore) - predMovAvg(1: end -
nBefore));

	% -> fast fourier transform
	n = length(predFFT);
	rawTemp = rawMovAvg(end - n + 1: end);
	plusDiffFFT = max(rawTemp(1: end - nBefore) - predFFT(1: end - nBefore));
	minusDiffFFT = min(rawTemp(1: end - nBefore) - predFFT(1: end - nBefore));


	%% Plot
	comparison = figure;
	minN = min(length(predMovAvg), length(predFFT));
	% predicted price lengths change
	x = linspace(1, minN, minN);
	predMovAvg = predMovAvg(end - minN + 1: end);
	predFFT = predFFT(end - minN + 1: end);
	rawTemp = rawMovAvg(end - minN + 1: end);
	minY = min(min([predMovAvg + minusDiffMovAvg, predFFT + minusDiffFFT,
rawTemp])) - 1;
	maxY = max(max([predMovAvg + plusDiffMovAvg, predFFT + plusDiffFFT,
rawTemp])) + 1;

	plotAll = plot(x, predMovAvg, 'b',...
		x, predMovAvg + plusDiffMovAvg, 'b:',...
		x, predMovAvg + minusDiffMovAvg, 'b:',...
		x, predFFT, 'r',...
		x, predFFT + plusDiffFFT, 'r:',...
		x, predFFT + minusDiffFFT, 'r:',...
		x(1: end - nBefore), rawTemp(1: end - nBefore), 'g',... % predicting data
		x(end - nBefore: end), rawTemp(end - nBefore: end), 'k',... % predicteddata
		ones(20, 1) * (length(x) - nBefore), linspace(minY, maxY, 20), 'k--',...
		'LineWidth', 2);
	plotAll(2).LineWidth = 1;
	plotAll(3).LineWidth = 1;
	plotAll(5).LineWidth = 1;
	plotAll(6).LineWidth = 1;
	plotAll(9).LineWidth = 1;
	xlim([1 length(x) + 0.25]);
```

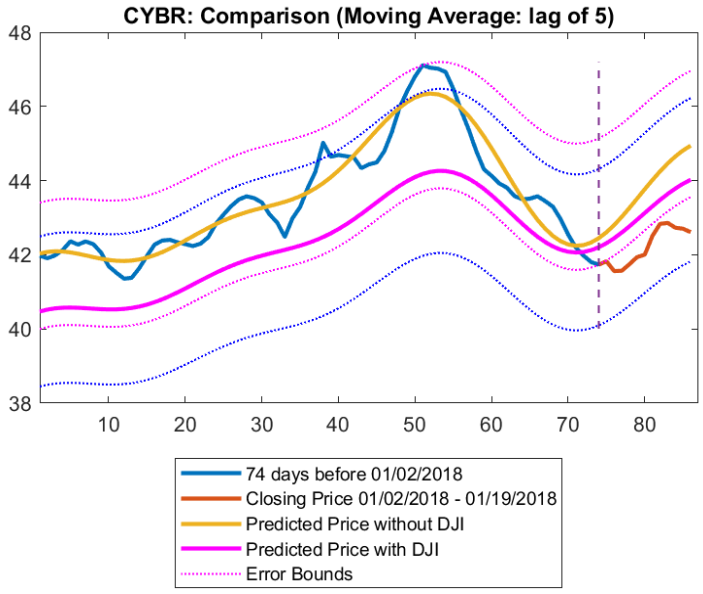
```
        legend([plotAll(7) plotAll(8) plotAll(1) plotAll(2) plotAll(4) plotAll(5)],...
               'Stock Price for Training',...
               'Stock Price for Testing Predictions',...
               'Prediction with Moving Average',...
               'Error Bounds of Prediction with Moving Average',...
               'Prediction with Fourier Transform',...
               'Error Bounds of Prediction with Fourier Transform',...
               'Location', 'southoutside');
        title(strcat(stockName, ': Comparison'));
        saveas(comparison, strcat(stockName, '_comparison.png'));

end

close all;
```

Appendix C: Stocks Price Prediction using Nasdaq and Dow Jones correlation model

Portfolio 1:

Cyberark Software Ltd stock

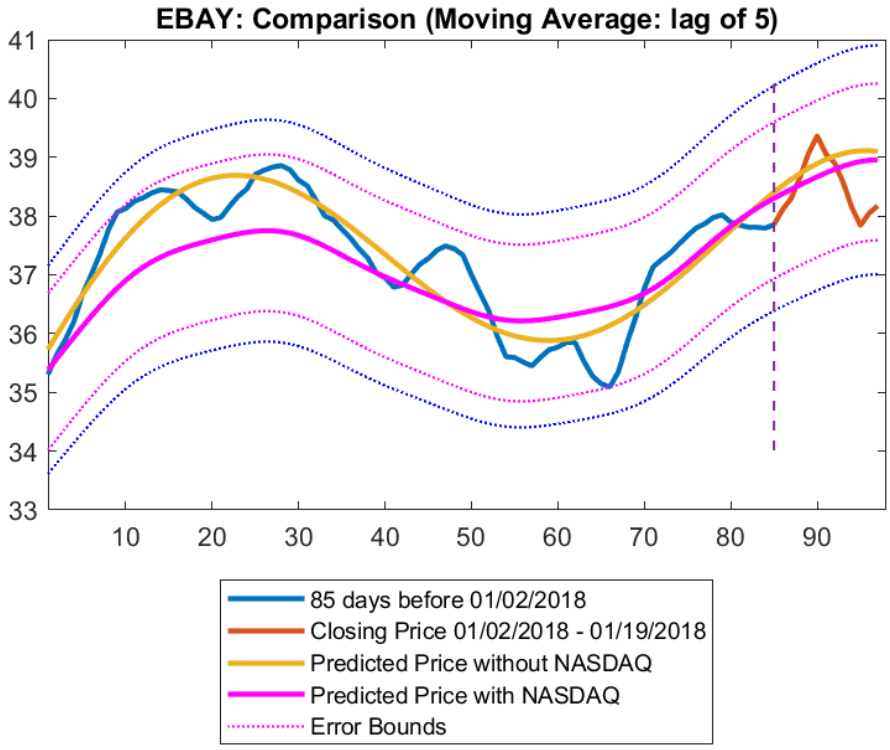


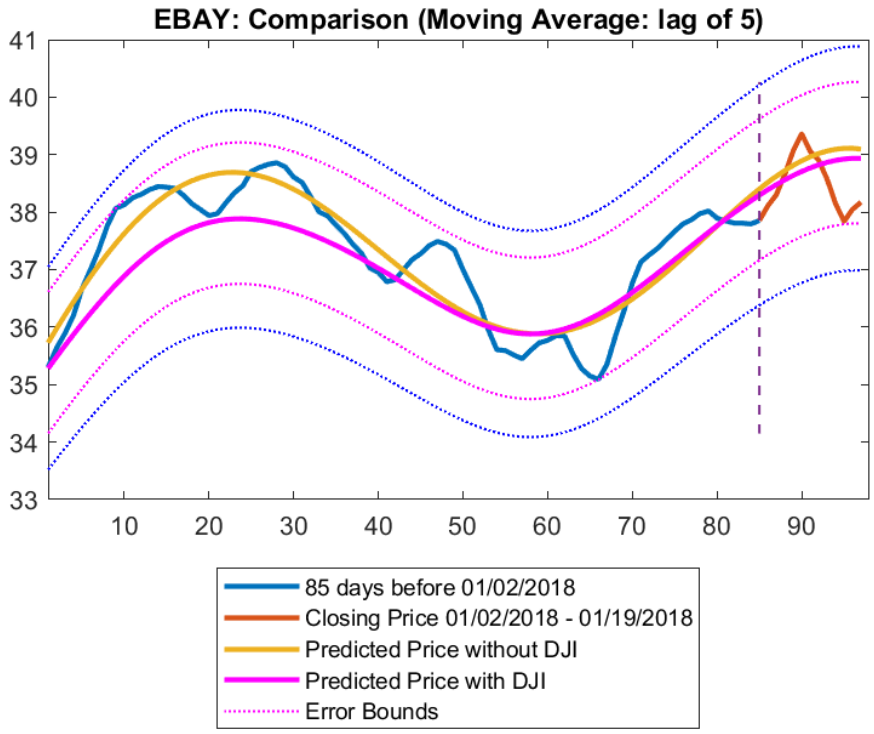


Cyberark Software Ltd stock price prediction is fitted by using closing price of 74 days before Jan 2, 2018. The predicted price using Nasdaq correlation is slightly higher than the actual

price, and more accurate comparing the the model without Nasdaq correlation. For the Dow jones model, the prediction does not differ from the Nasdaq model. Both of their price boundaries are able to captures all of testing data.

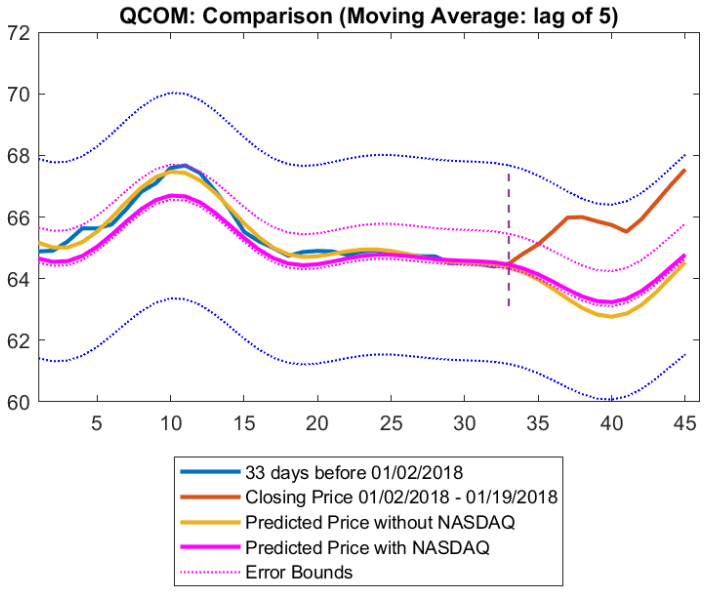
eBay Inc

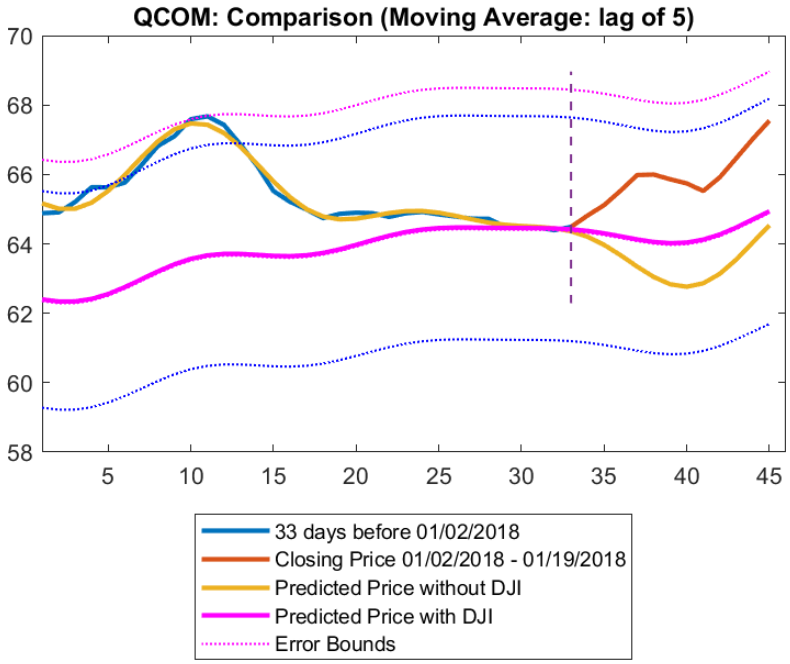




eBay Inc. stock price prediction is fitted by using closing price of 85 days before Jan 2, 2018. The predicted price using Nasdaq correlation is at about the average price of testing data. For the Dow jones model, the prediction does not differ from the Nasdaq model. Both of their price boundaries are able to captures all of testing data.

QUALCOMM, Inc.

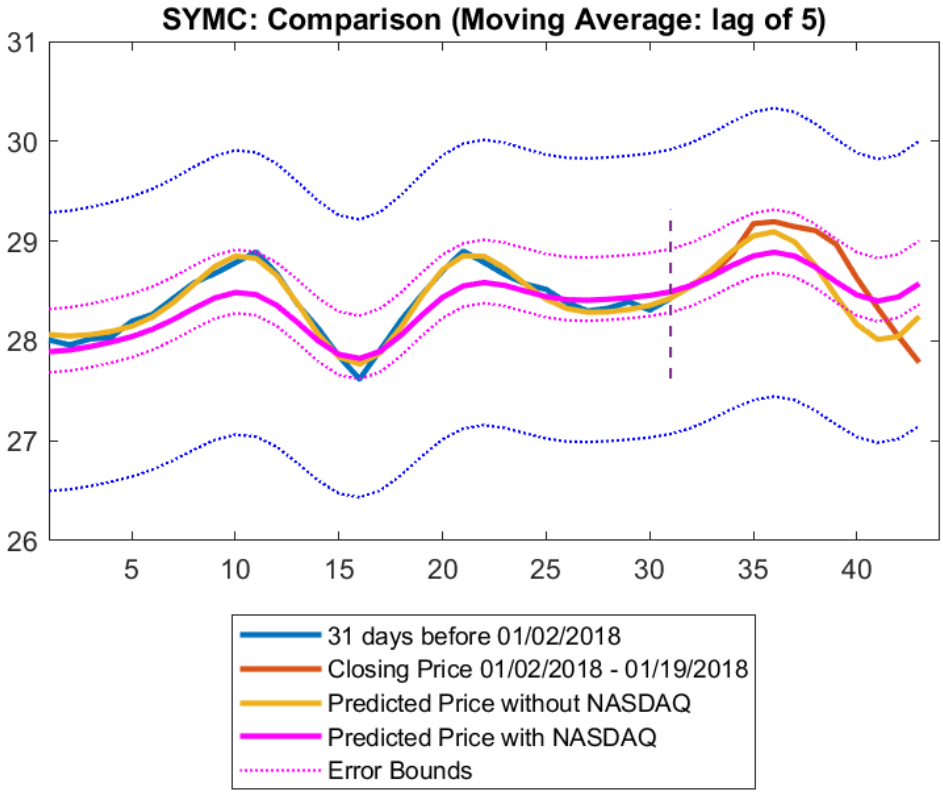


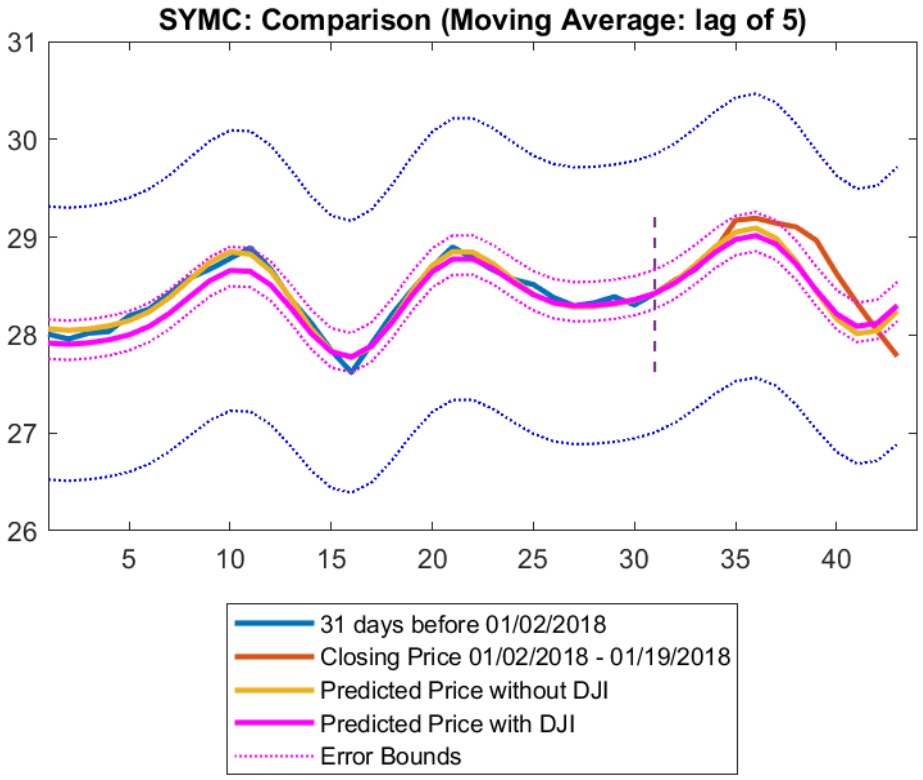


QUALCOMM, Inc. stock price prediction is fitted by using closing price of 33 days before Jan 2, 2018. The model with Nasdaq correlation and without fails to predict the stock price. The closing price after Jan 1, 2018 is increase without prior trend. For the Dow jones model,

the prediction is also unsuccessful to predict the stock price, but it is closed to the actual price than the Nasdaq model.

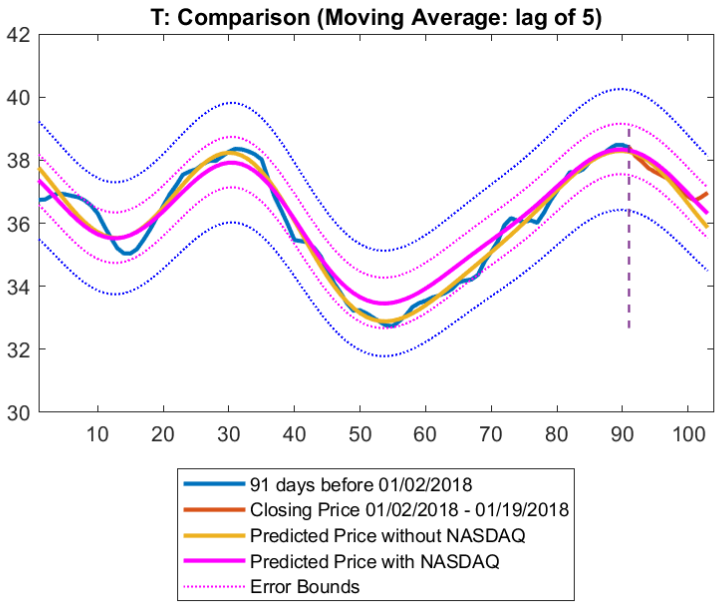
Symantec Corporation

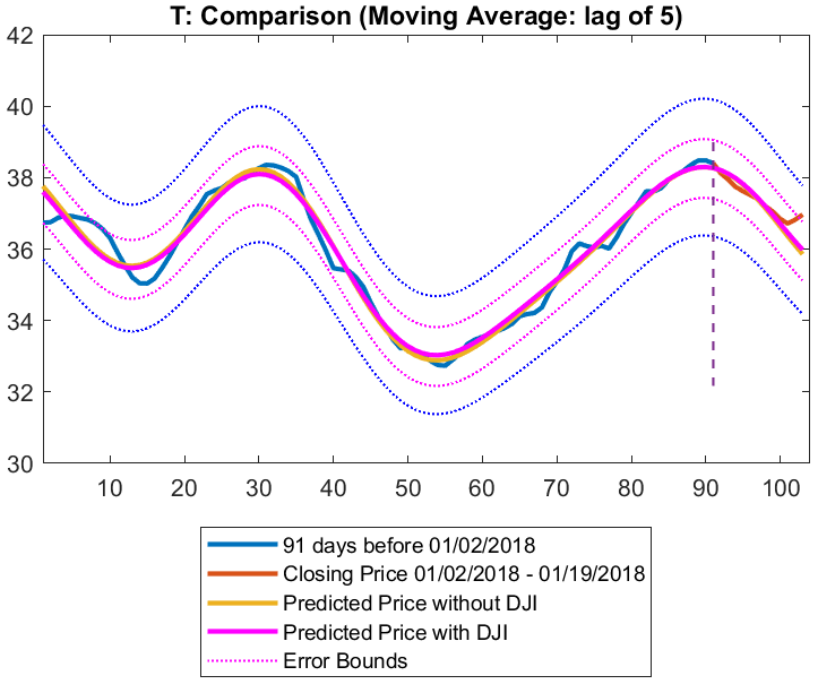




Symantec Corporation stock price prediction is fitted by using closing price of 31 days before Jan 2, 2018. The predicted price using Nasdaq correlation is very precise to the actual price, but does not much better from the model without Nasdaq correlation. For the Dow jones model, the prediction is more precise than Nasdaq model. However, Both of their price boundaries are able to captures all of testing data well.

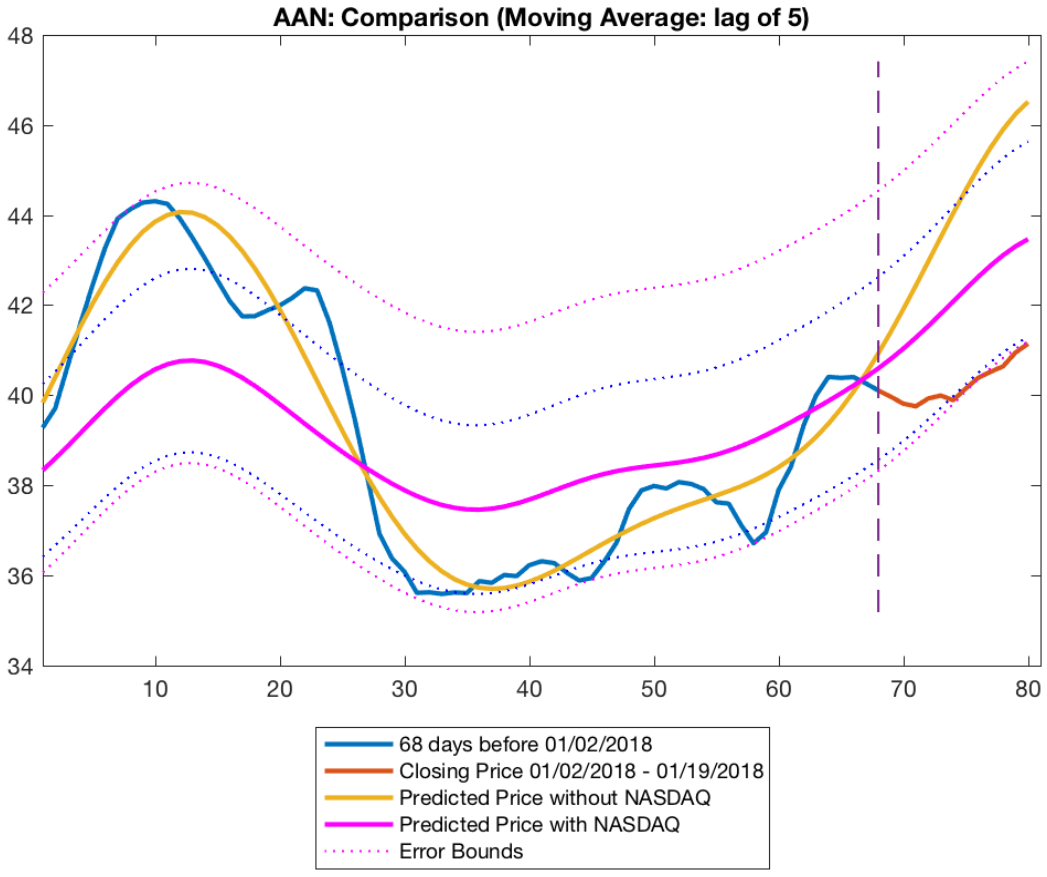
AT&T Inc.

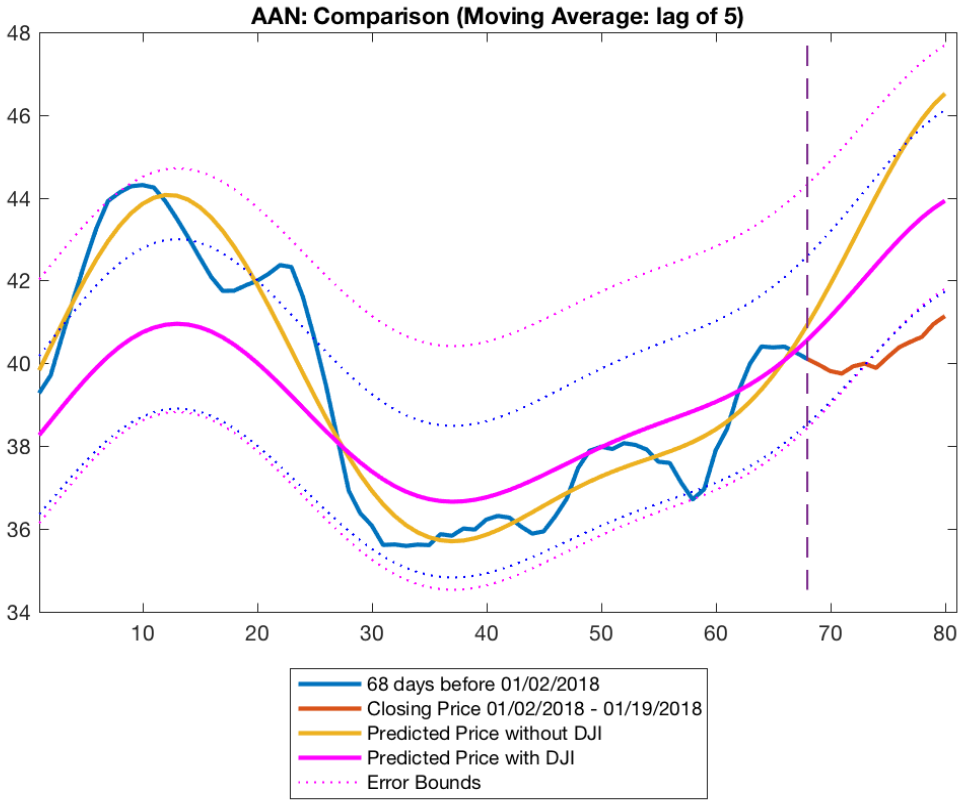

T: Comparison (Moving Average: lag of 5)
42
40
38
36
34
32
30
10
20
30
40
50
60
70
80
90
100
91 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds


T: Comparison (Moving Average: lag of 5)
42
40
38
36
34
32
30
10
20
30
40
50
60
70
80
90
100
91 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without DJI
Predicted Price with DJI
Error Bounds

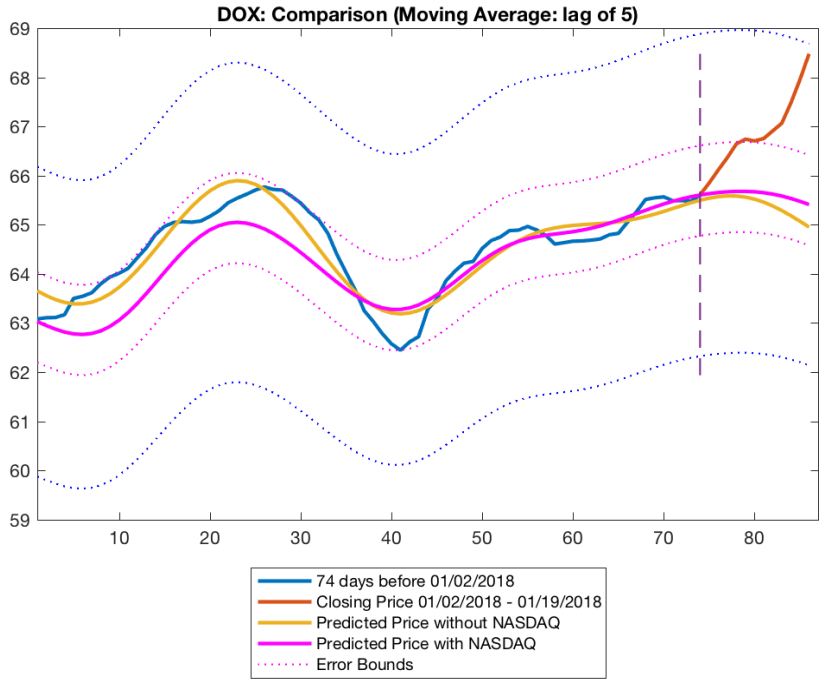
AT&T Inc. stock price prediction is fitted by using closing price of 91 days before Jan 2, 2018. The predicted price using Nasdaq correlation is almost exact with the actual price, but does not work better from the model without Nasdaq correlation. For the Dow jones model, the prediction is slightly better than the Nasdaq model. Both of their price boundaries are able to captures all of testing data.

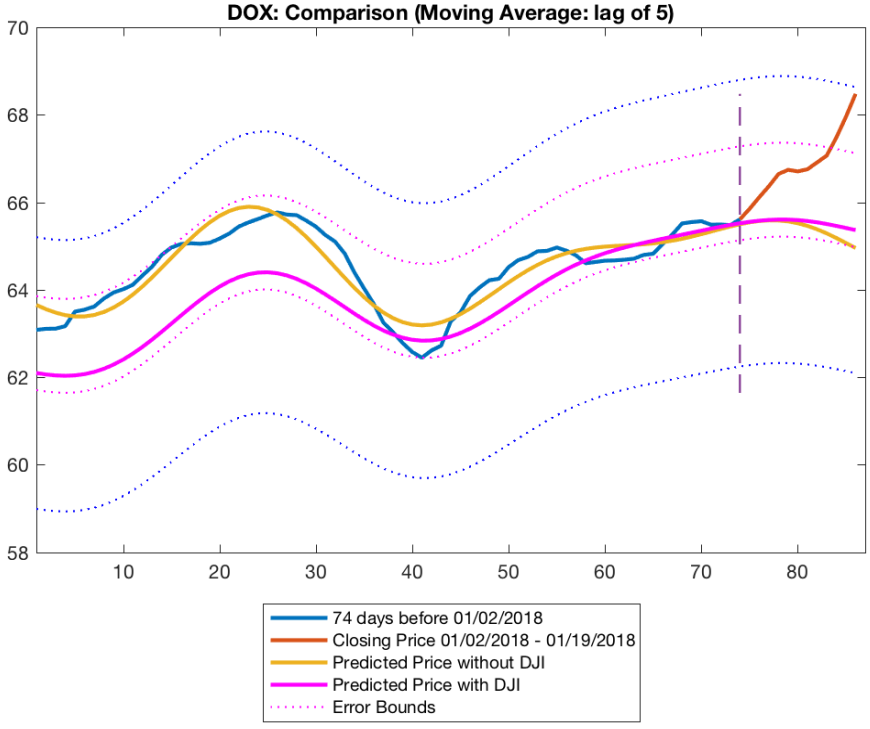
Portfolio 2:


AAN: Comparison (Moving Average: lag of 5)
48
46
44
42
40
38
36
34
10
20
30
40
50
60
70
80
68 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds

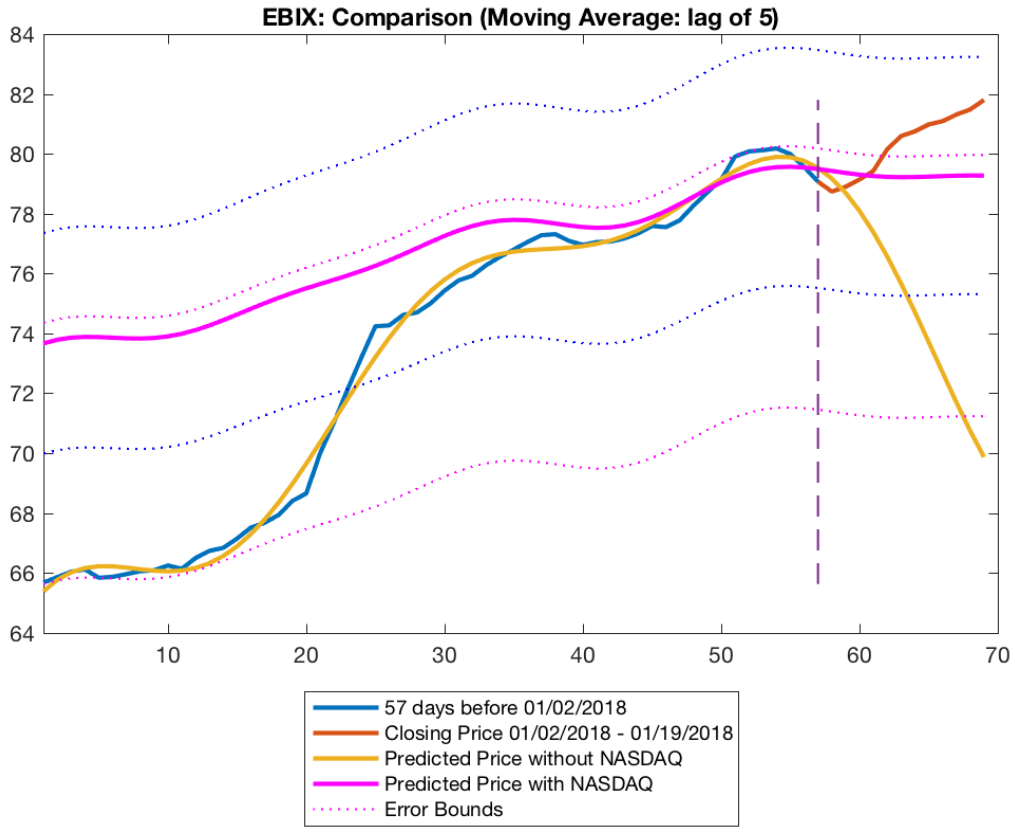


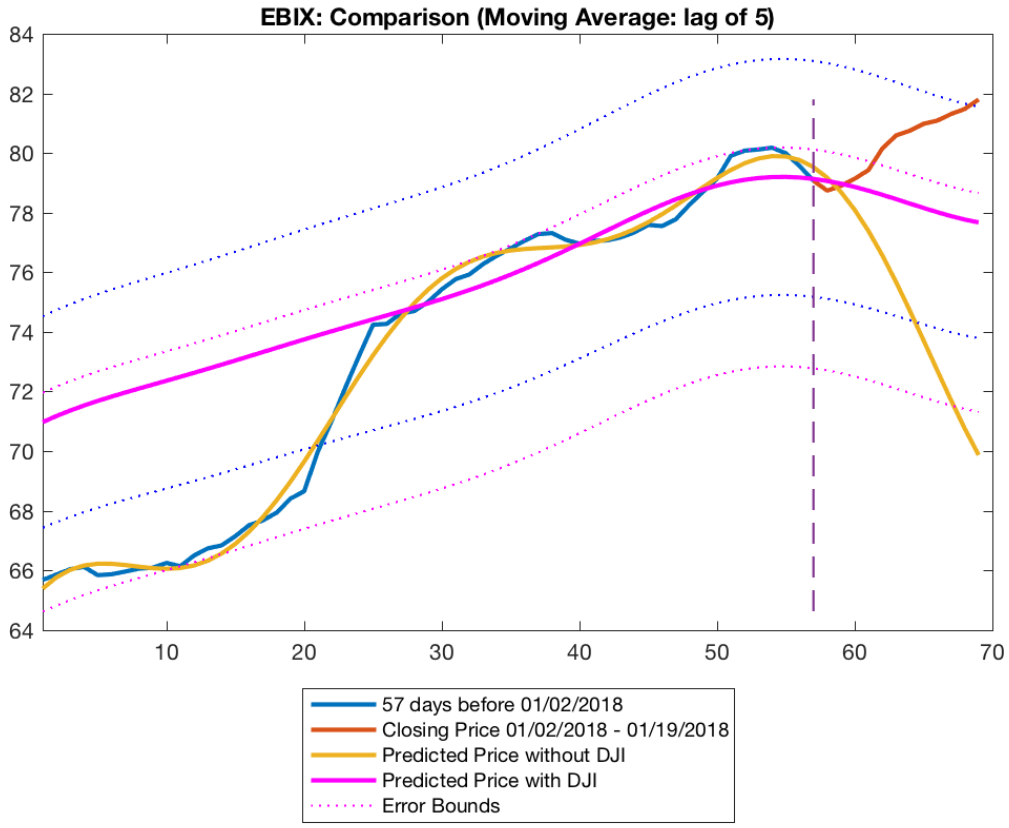
Aaron's Inc is fitted by using closing price of 68 days. For model involving Nasdaq, the prediction is clearly better with Nasdaq than without Nasdaq. Similar to model including Dow Jones, the model seems to predict better with Dow Jones. However, the Nasdaq's error boundaries seem to capture the testing data better than Dow Jones' error boundaries. AAN stock proves that including Nasdaq and Dow Jones index seem to be effective.


DOX: Comparison (Moving Average: lag of 5)
69
68
67
66
65
64
63
62
61
60
59
10
20
30
40
50
60
70
80
74 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds

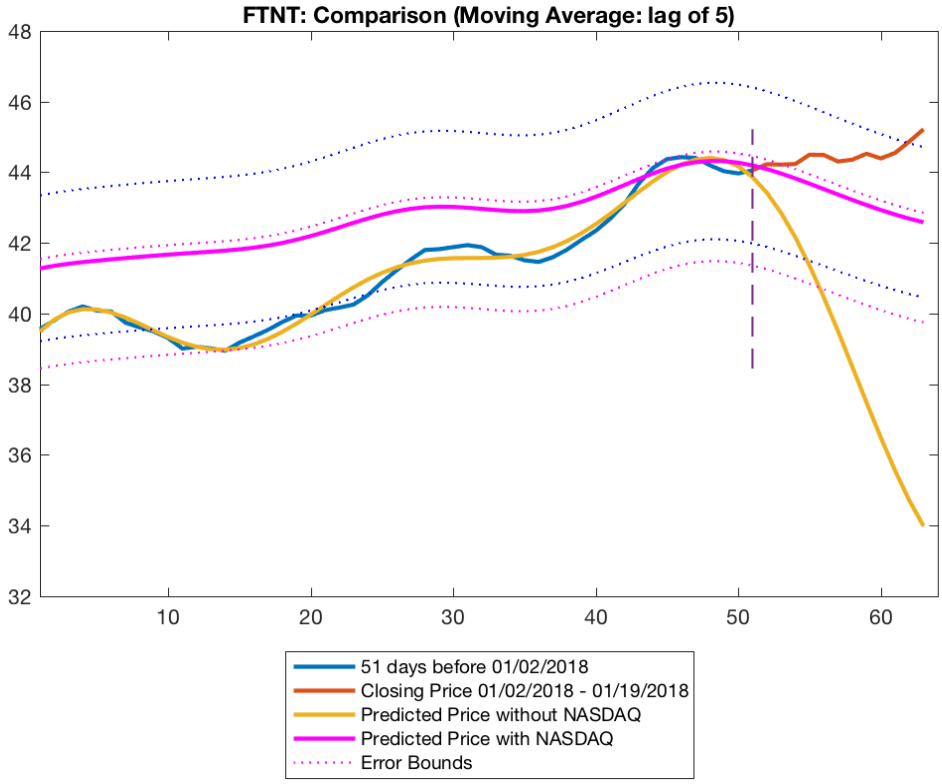

DOX: Comparison (Moving Average: lag of 5)
70
68
66
64
62
60
58
10
20
30
40
50
60
70
80
74 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without DJI
Predicted Price with DJI
Error Bounds

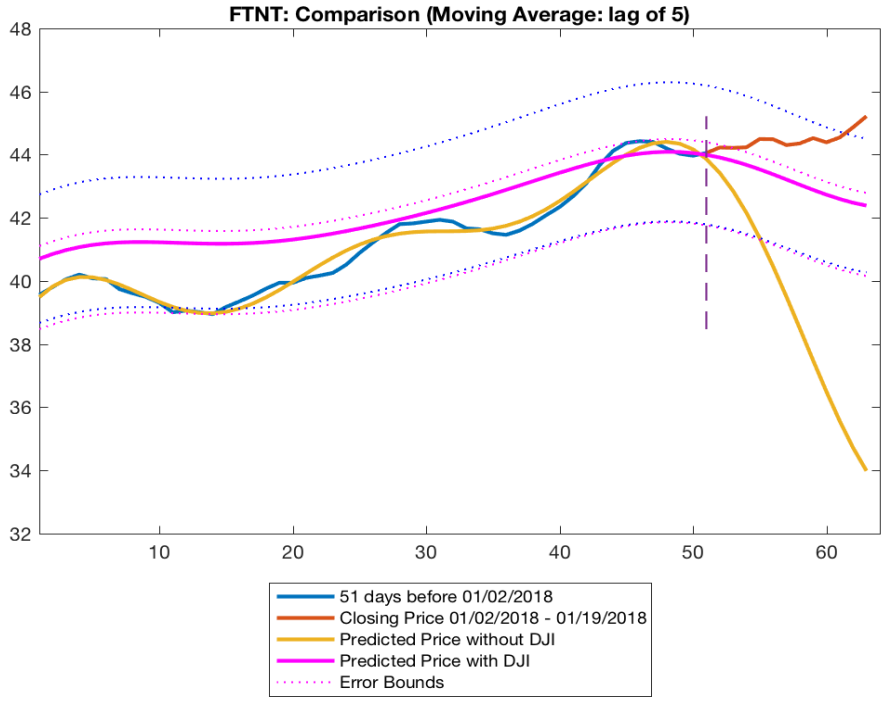
Amdocs Corporation is fitted by using closing price of 74 days. Both Nasdaq and Dow Jones index seem to be ineffective in helping the price prediction. Although it does slightly better job in matching the real closing price, the results from the graph does not good enough. The error boundaries from both models still be able to keep the real closing price data inside the boundaries in the beginning but the closing price still manages to spike out of the boundaries at the end.



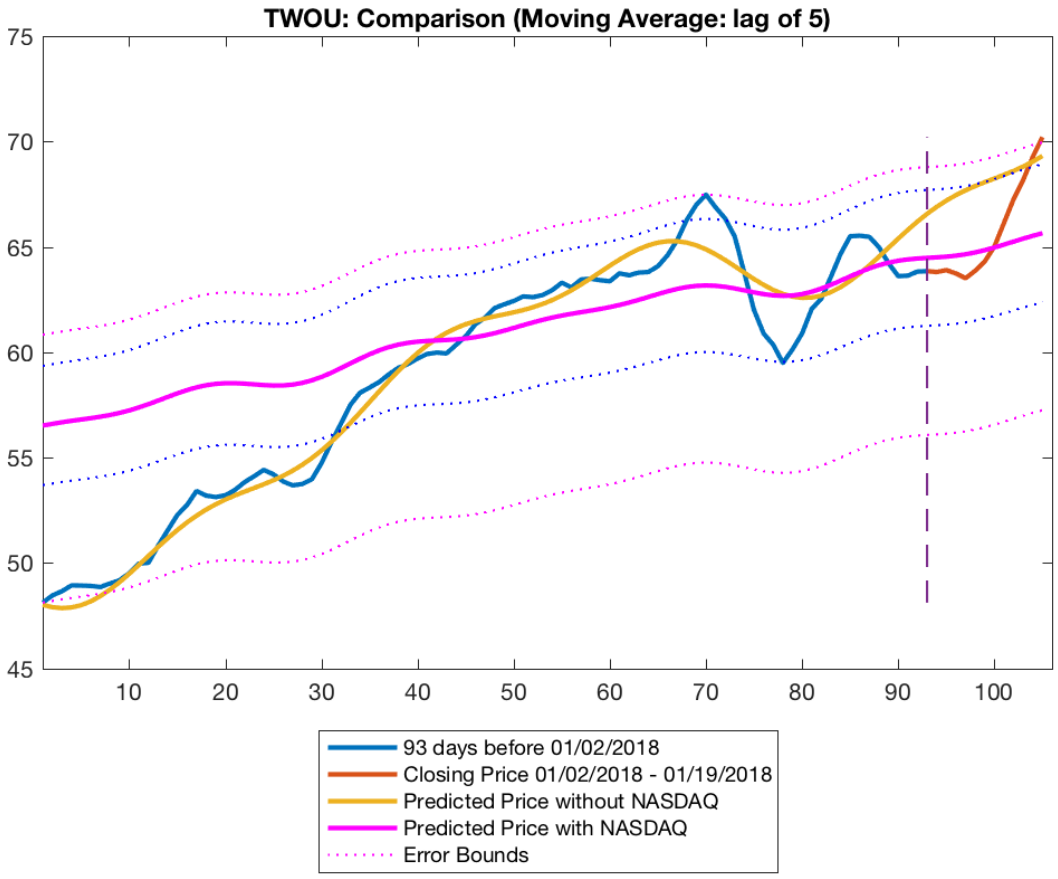


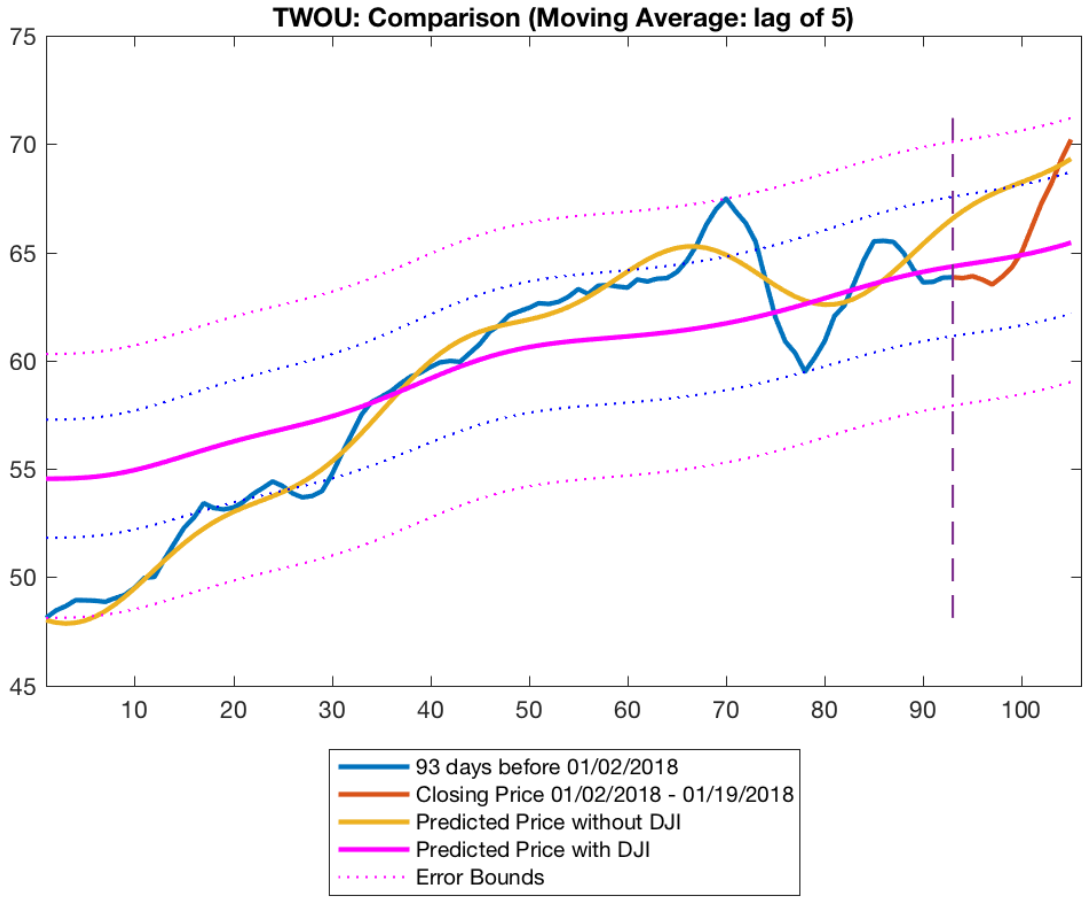
Ebix Inc is fitted by the closing price of 57 days. It is clear on both graphs that Nasdaq and Dow Jones index do help improving the model in predicting the prices. Although both indices did improve the model, but the trend of the actual closing price and prediction from the indices are still not the same. The actual price also seems to be out of the error boundaries of both models.


FTNT: Comparison (Moving Average: lag of 5)
48
46
44
42
40
38
36
34
32
10
20
30
40
50
60
51 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds


FTNT: Comparison (Moving Average: lag of 5)
48
46
44
42
40
38
36
34
32
10
20
30
40
50
60
51 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without DJI
Predicted Price with DJI
Error Bounds

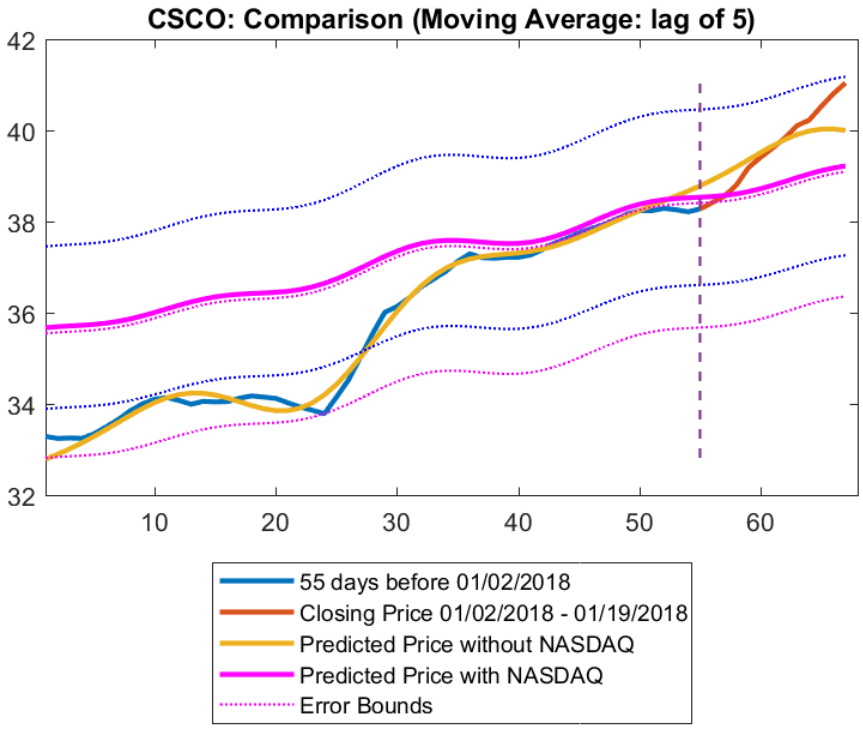
Fortinet Inc is fitted by closing price of 51 days. Similar to Ebix's models, Nasdaq and Dow Jones index prove to be really effective in improving the prediction model. The trends of actual closing price and prediction price are still not the same and the error boundaries are not able to capture the testing data. This means there is still inconsistency inside the model that uses to predict the data.

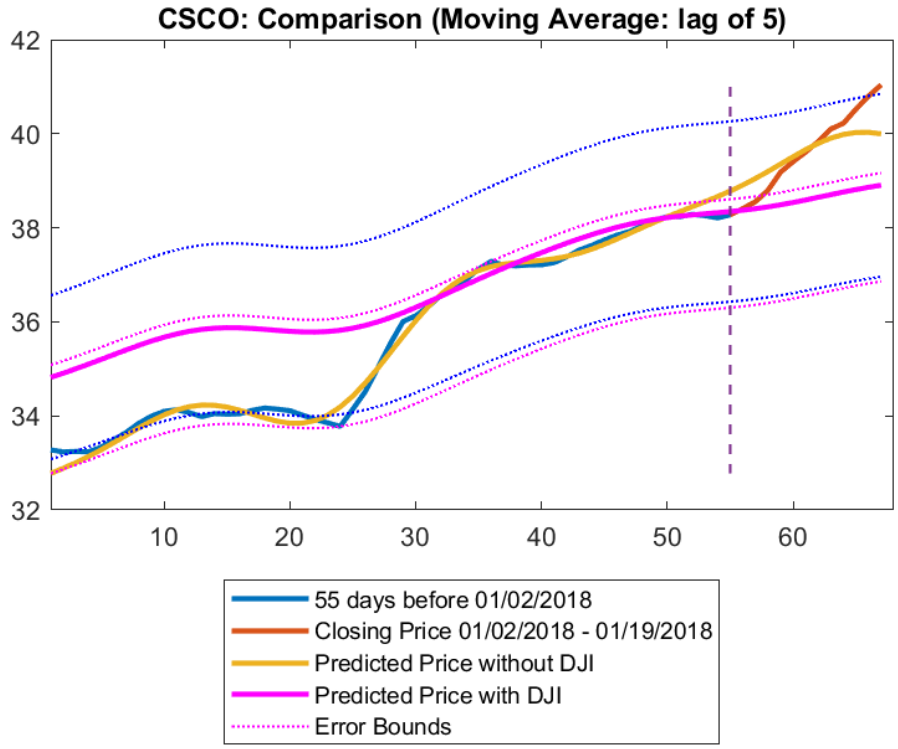




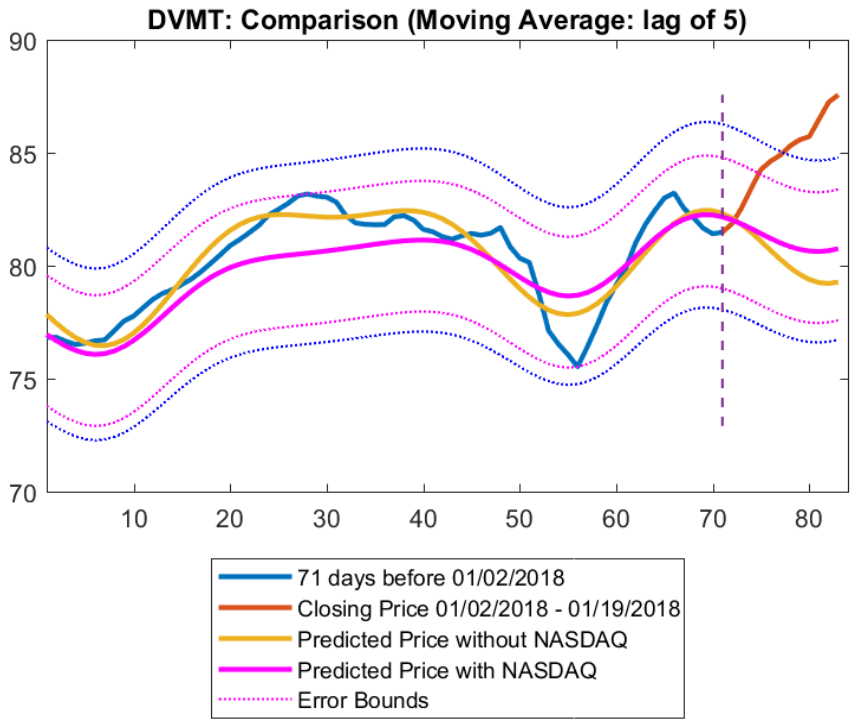
2U Inc is fitted by closing price of 93 days. If focus at the beginning period of the prediction, models with Nasdaq and Dow Jones index seem to make predictions that are closer to the actual price than models without Nasdaq and Dow Jones index. Although, the actual price spikes up in the later period, which causes the models without Nasdaq and Dow Jones index to have closer predictions that the models that do include there indices. However, the error boundaries in these two models still be able to capture the testing data.

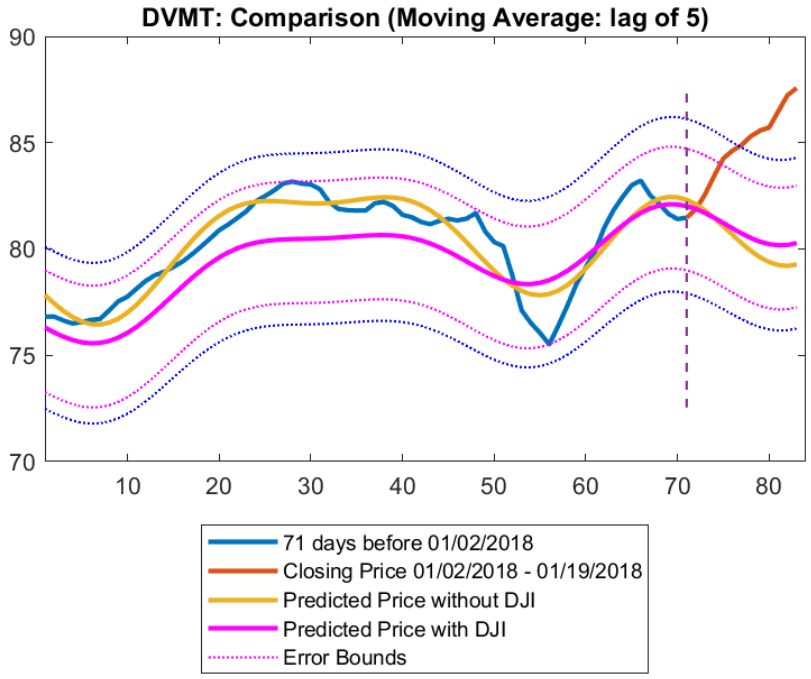
Portfolio 3:


CSCO: Comparison (Moving Average: lag of 5)
42
40
38
36
34
32
10
20
30
40
50
60
55 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds

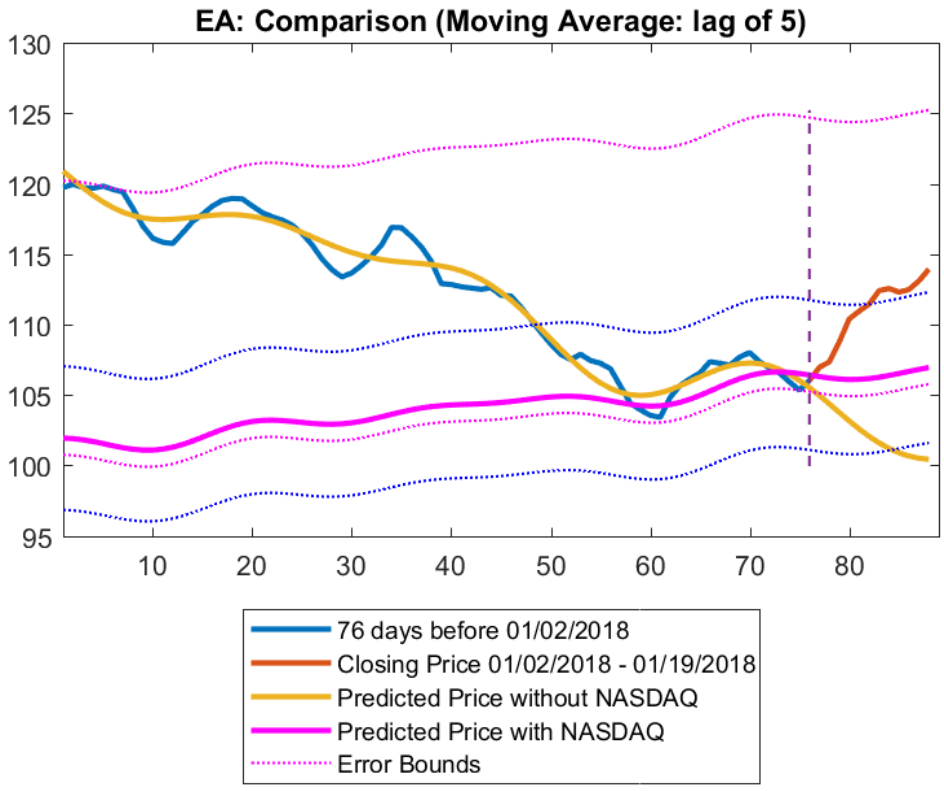

CSCO: Comparison (Moving Average: lag of 5)
42
40
38
36
34
32
10
20
30
40
50
60
55 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without DJI
Predicted Price with DJI
Error Bounds

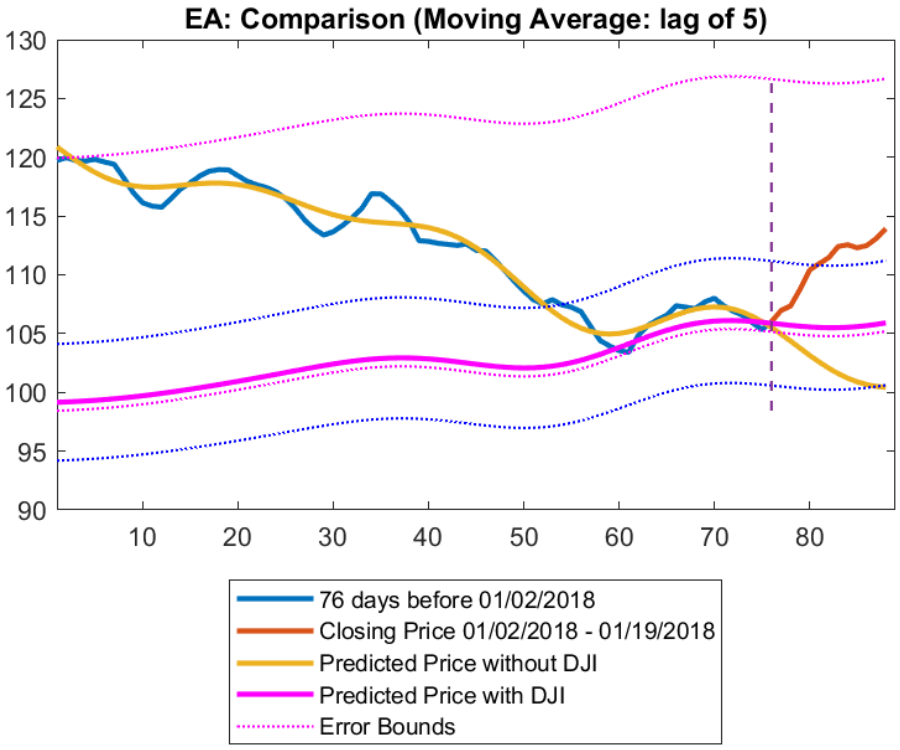
CSCO stock price prediction is fitted by using the closing price of 55 days before Jan 2, 2018. The predicted price without using NASDAQ is more accurate than using NASDAQ. For Dow Jones model, the result is repeated. Price prediction without using DJI is more accurate than using DJI.



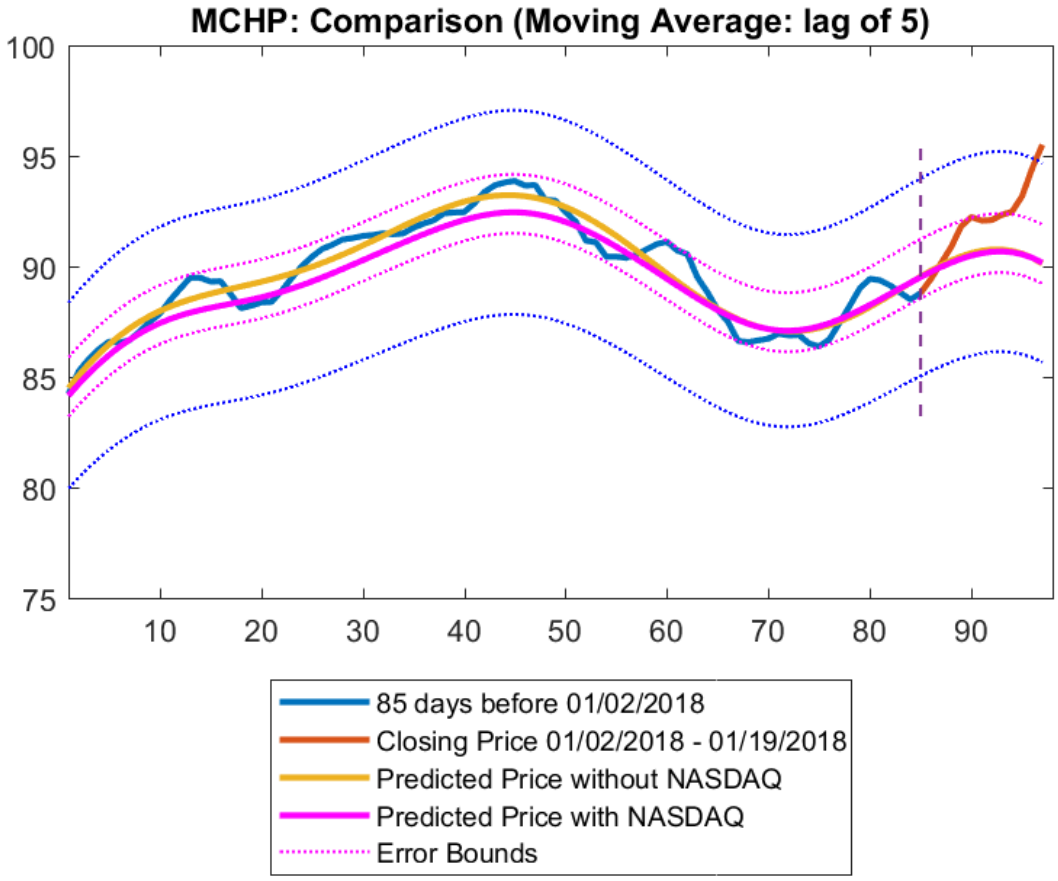


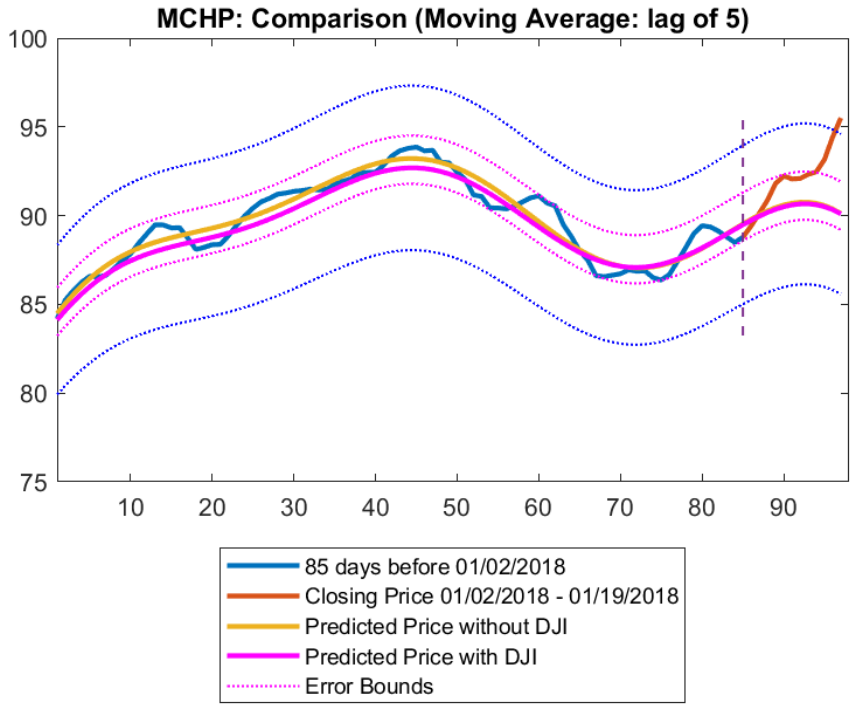
DVMT stock price prediction is fitted by using the closing price of 71 days before Jan 2, 2018. All the predicted models fall short from this because of the sudden rise in the stock price of DVMT. However, using NASDAQ and DJI to help predict the price is more accurate than not using them.


EA: Comparison (Moving Average: lag of 5)
130
125
120
115
110
105
100
95
10
20
30
40
50
60
70
80
76 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds

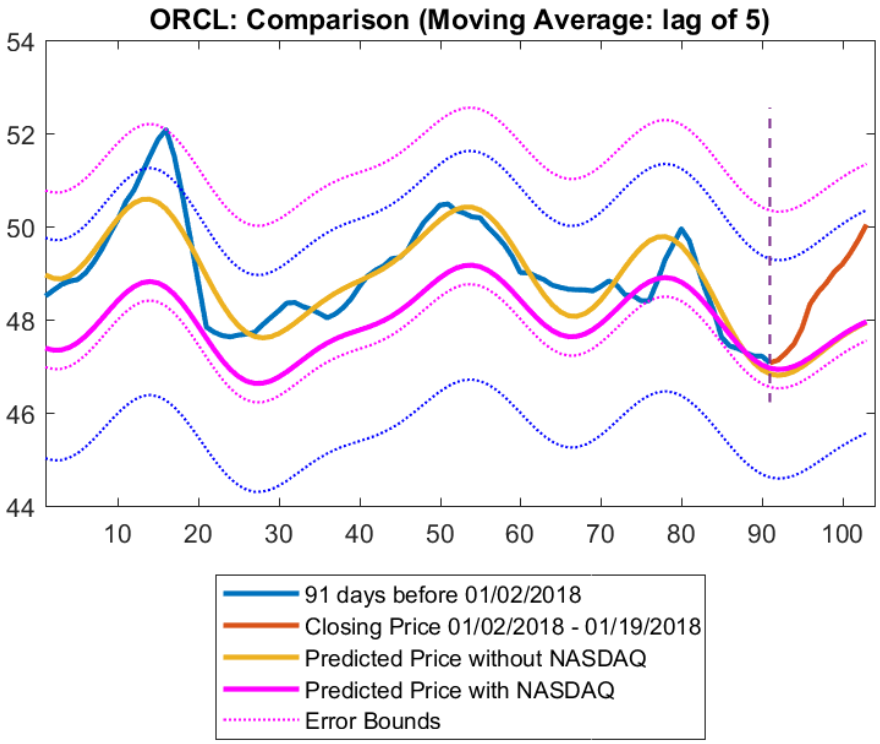

EA: Comparison (Moving Average: lag of 5)
130
125
120
115
110
105
100
95
90
10
20
30
40
50
60
70
80
76 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without DJI
Predicted Price with DJI
Error Bounds

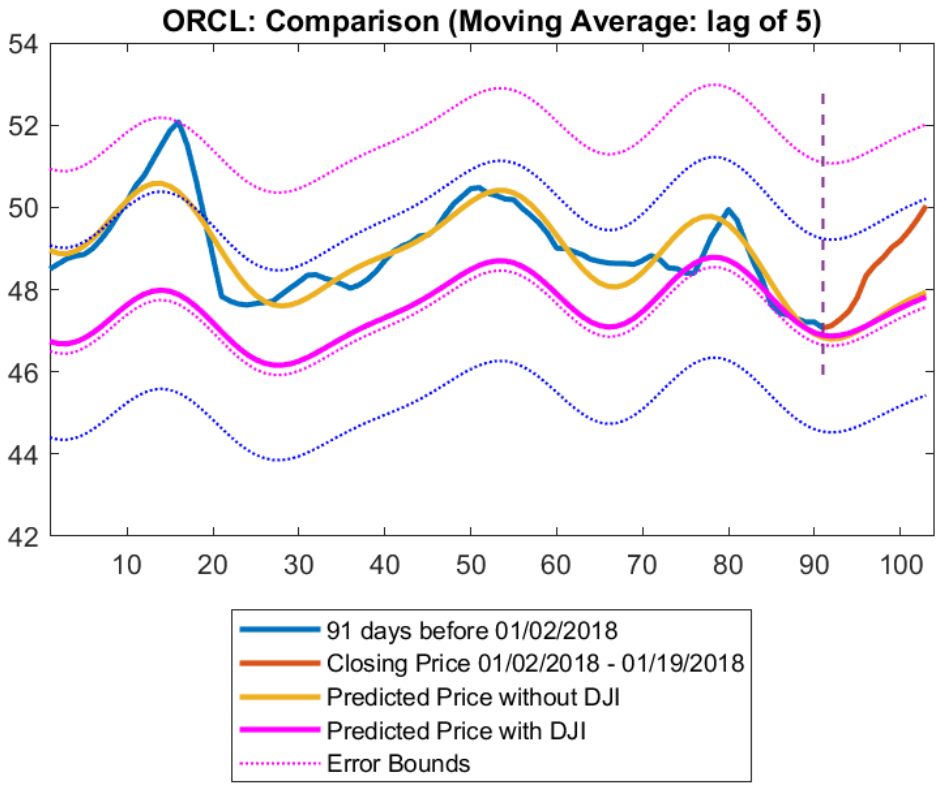
EA stock price prediction is fitted by using the closing price of 76 days before Jan 2, 2018. Same as DVMT prediction, the predicted models fall short from this because of the sudden rise in the stock price of EA. However, using NASDAQ and DJI to help predict the price is more accurate than not using them. Without NASDAQ and DJI, the predicted price will keep dropping.





MCHP stock price prediction is fitted by using the closing price of 85 days before Jan 2, 2018. The predicted models fail to predict the price of MCHP. As the stock price of MCHP increases, the predicted price for all the graphs drops down. . Both of their price boundaries are able to captures all of testing data.

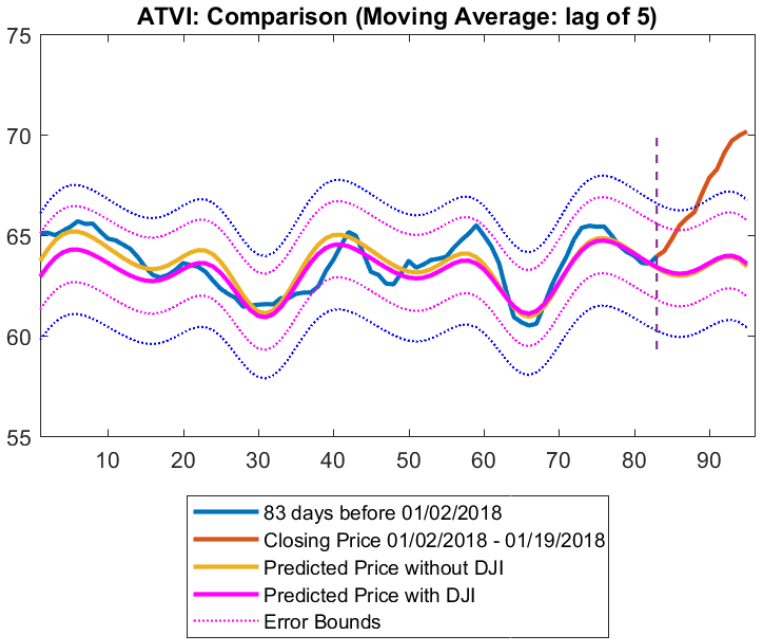

ORCL: Comparison (Moving Average: lag of 5)
54
52
50
48
46
44
10
20
30
40
50
60
70
80
90
100
91 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without NASDAQ
Predicted Price with NASDAQ
Error Bounds

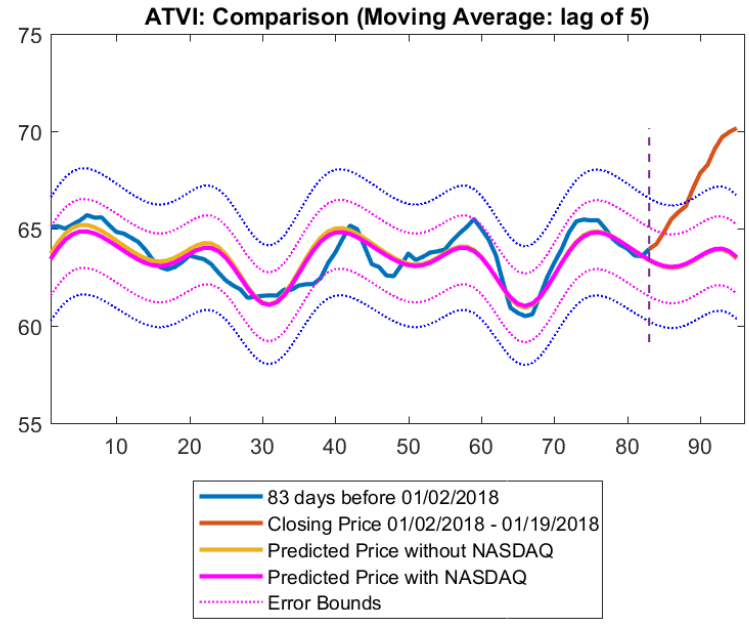

ORCL: Comparison (Moving Average: lag of 5)
54
52
50
48
46
44
42
10
20
30
40
50
60
70
80
90
100
91 days before 01/02/2018
Closing Price 01/02/2018 - 01/19/2018
Predicted Price without DJI
Predicted Price with DJI
Error Bounds

ORCL stock price prediction is fitted by using the closing price of 91 days before Jan 2, 2018. The predicted models fail to predict the price of ORCL. As the price of ORCL increases, all models, including with and without both Nasdaq and DJI, increase too. This is acceptable to a certain extent since we can use this to judge whether the price will go up or down but can't pinpoint the exact price.

Portfolio 4:

Activision Blizzard, Inc.

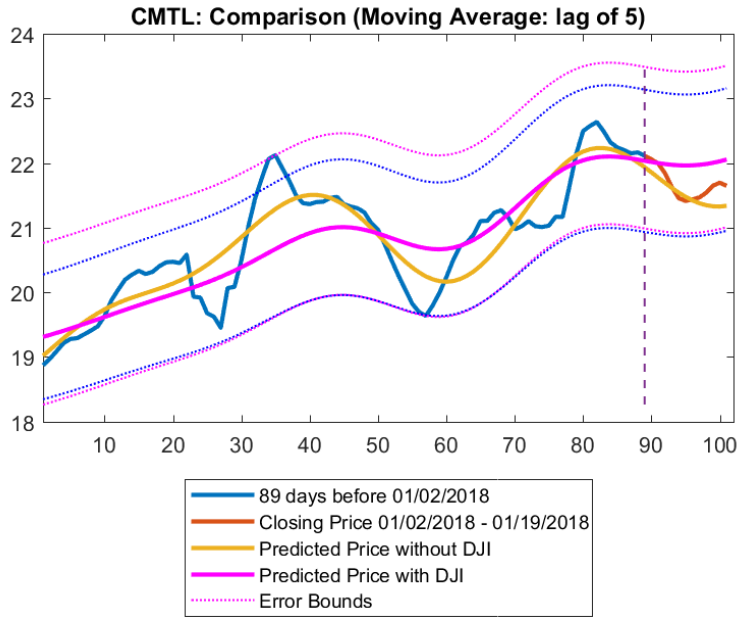


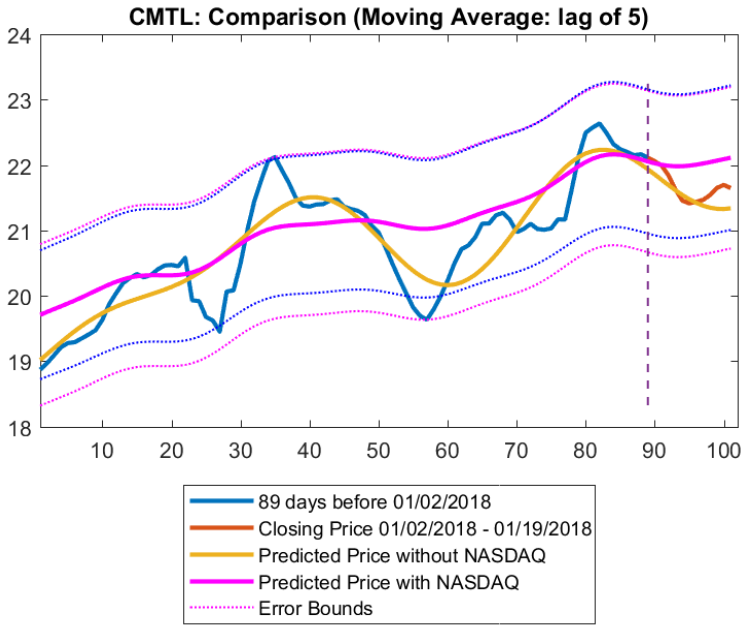


Activision Blizzard, Inc. stock price prediction is fitted by using closing price of 83 days before Jan 2, 2018. The predicted price using Nasdaq correlation is slightly higher than the actual price. For the Dow jones model, the prediction does not differ from the original

model, which is without index stocks. Both of their price boundaries cannot really capture the testing stock price as the real price went while all of the predictions roughly guessed that the price would have remain stable.

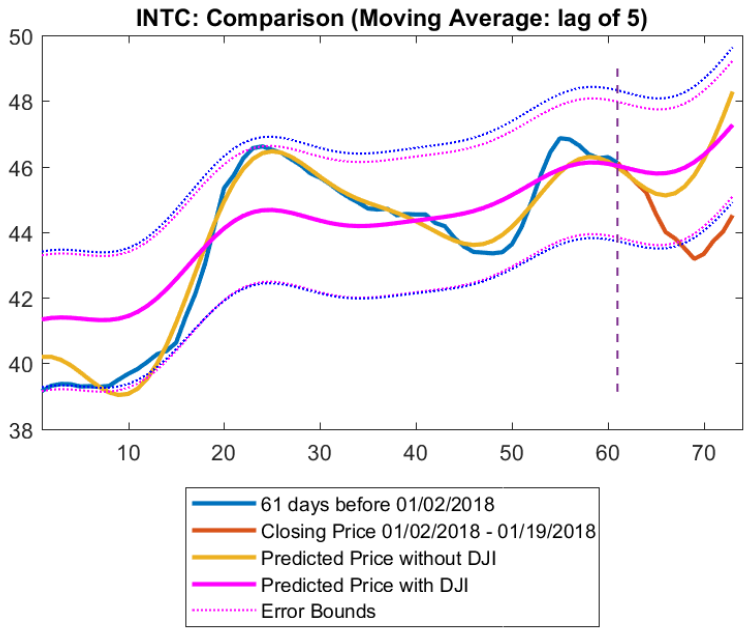
Comtech Telecomm. Corp.

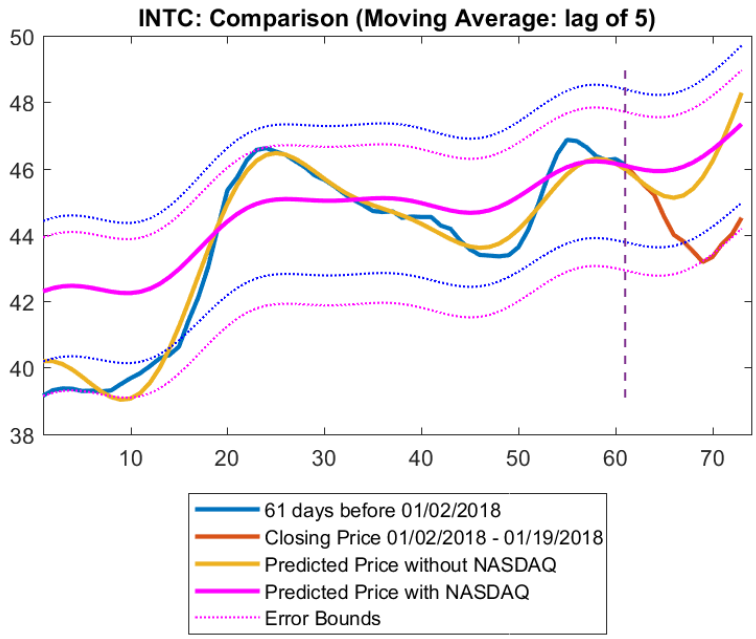




Comtech Telecomm. Corp. stock price prediction is fitted by using closing price of 89 days before Jan 2, 2018. The predicted price using Nasdaq correlation is slightly higher than the actual price. Similarly, the Dow jones model had slightly higher prediction. Both of their price boundaries can completely capture the testing stock price. However, the prediction without index stocks is more precise.

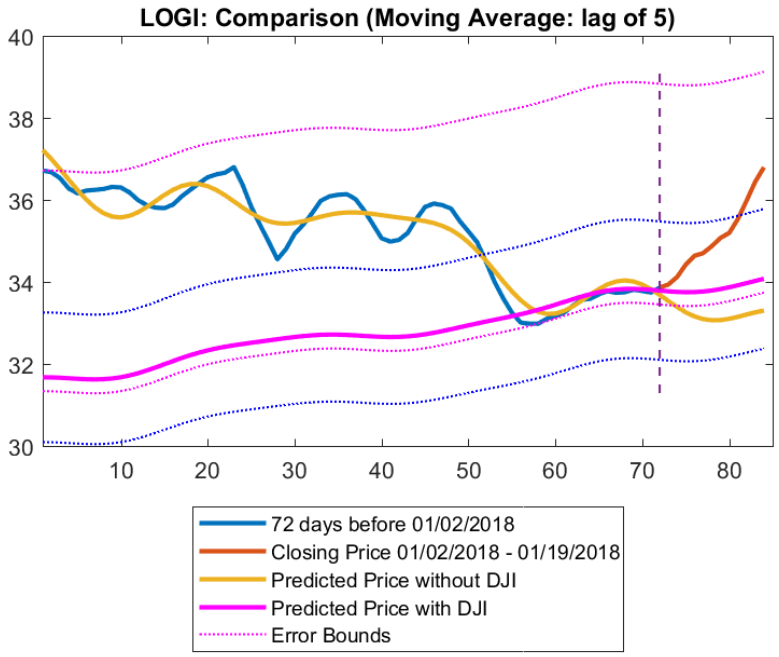
Intel Corporation

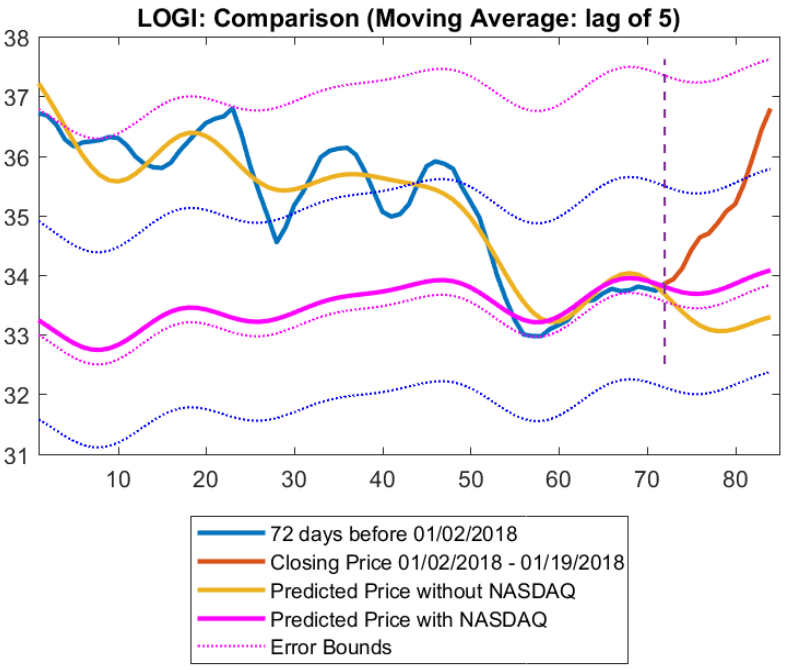




Intel Corporation stock price prediction is fitted by using closing price of 61 days before Jan 2, 2018. The predicted price with index stocks went wider than the original predicted price, but the trends were similar. Only did the predicted price with Nasdaq index remain in the error boundaries. However, all of them cannot acceptably predicted the price stock as it dropped far below the predicted prices.

Logitech International SA (USA)

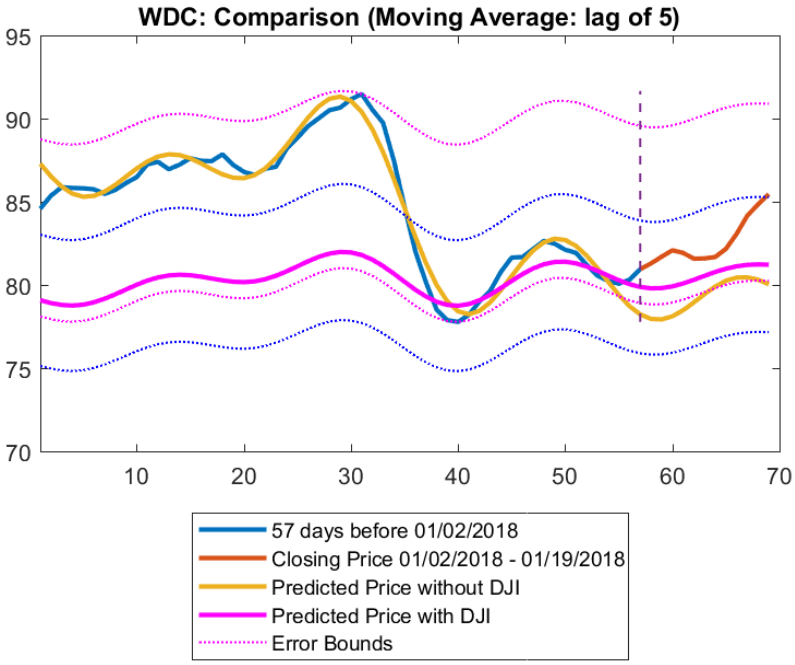


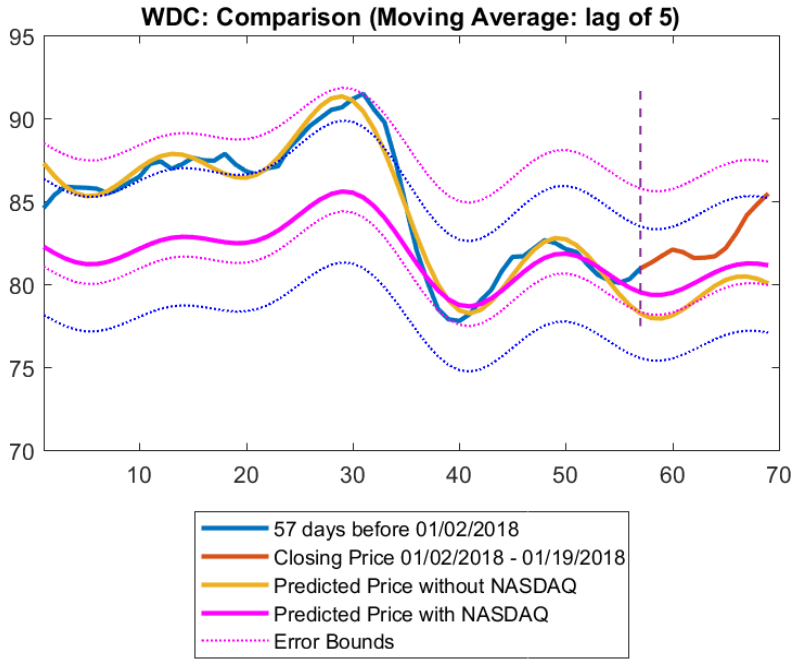


Logitech International SA (USA) stock price prediction is fitted by using closing price of 72 days before Jan 2, 2018. The index stocks strongly affected the predicted prices; they

reduced values of the real price stock. However, they gave more accurate predictions than the original one, but they still could not match the testing data, real price after predictions.

Western Digital Corp.





Western Digital Corp. stock price prediction is fitted by using closing price of 57 days before Jan 2, 2018. Both Nasdaq and Dow john neutralized the sudden drop and made predicted price more stable. Regardless, all of the models correctly predicted the stock price.

Bibliography

1. Tsay, Ruey S. An introduction to analysis of financial data with R. John Wiley & Sons, 2014.
2. Perchanok, K., and I. Hrytsyuk. "The Encyclopedia of the Indicator RSI (Relative Strength Index)." Charleston: CreateSpace (2011).
3. Derryberry, DeWayne R. "Fitting Fourier Series." Basic Data Analysis for Time Series with R: 115-124.
4. Staff, Investopedia. "What is the difference between the Dow and the Nasdaq?" Investopedia, 10 Apr. 2017, https://www.investopedia.com/ask/answers/difference-between-dow-and-nasdaq/
5. Picardo, CFA Elvis. "Moving Average ." *Investopedia*, 7 Dec. 2016, www.investopedia.com/terms/m/movingaverage.asp.
6. Ibbotson, Roger G. "Why does market volatility matter?" *Yale Insights*, 23 Nov. 2011, www.insights.som.yale.edu/insights/why-does-market-volatility-matter.
7. Staff, Investopedia. "Volatility." *Investopedia*, 4 Aug. 2016, www.investopedia.com/terms/v/volatility.asp.
8: Graham, Benjamin. The intelligent investor: A book of practical counsel. Prabhat Prakashan, 1965.
9. Otani, Akane, and Jon Sindreu. "U.S. Stocks Slide as Technology Drops --- Nasdaq Index Falls 1.4% as Alphabet, Nvidia Slide; S&P 500, DJIA Log Worst Fall in Month." Wall Street Journal, Eastern edition ed.Jun 30 2017. ProQuest. Web. 19 Feb. 2018 .
10. Kilgore, Tomi. "U.S. Stocks Higher, Led by Nasdaq; Technology Sector Leads Advances as Investors Await Earnings Season." Wall Street Journal (Online)Apr 08 2014. ProQuest. Web. 19 Feb. 2018 .
11. Otani, Akane, and Mike Bird. "Nasdaq Composite Hits another Record; Gains in Technology Shares Lift Nasdaq Index to 33rd Closing High of the Year." Wall Street Journal (Online)May 16 2017. ProQuest. Web. 19 Feb. 2018 .

12. Pant, Paula. “The S&P 500, NASDAQ, Dow Jones - What is This Stuff?” *The Balance*, 3 June 2017, www.thebalance.com/the-sandp-500-nasdaq-dow-jones-what-is-this-stuff-453745.
13. Kennon, Joshua. “Why Do Stock Prices Fluctuate?” *The Balance*, 30 June 2017, www.thebalance.com/why-do-stock-prices-fluctuate-356347.
14. “Autocorrelation.” *Statistics Solutions*, www.statisticssolutions.com/autocorrelation/.
15. Krambeck, Donald. “Learn About Fourier Coefficients.” *All About Circuits*, 5 Jan. 2016, www.allaboutcircuits.com/technical-articles/the-fourier-coefficients/.
16. Rick. “Trading Psychology: Setting Stops The Stress-Free Way.” *How To Find The Best Entry Points For Short Selling Stocks*, Jan. 11AD, www.morpheustrading.com/blog/best-entry-short-selling-stocks/.
17. Lough, William H. B. 1881. *Business finance, a practical study of financial management in private*. Nabu Press, 2010.
18. Zweig, Jason. “Making Sense of Market Forecasts.” *The Wall Street Journal*, Dow Jones & Company, 8 Jan. 2011, www.wsj.com/articles/SB10001424052748703675904576064320900295678.
19. SCHWERT, G. WILLIAM. “Why Does Stock Market Volatility Change Over Time?” *The Journal of Finance*, Blackwell Publishing Ltd, 30 Apr. 2012, onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1989.tb02647.x/full.
20. Fama, Eugene F. *The behavior of stock-Market prices*. 1966.
21. Staff, Investopedia. “Autocorrelation.” *Investopedia*, 9 Oct. 2015, www.investopedia.com/terms/a/autocorrelation.asp.
22. Weisstein, Eric W. "Least Squares Fitting." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/LeastSquaresFitting.html
23. Thenub314, . *Fourier series for square wave*. Retrieved from https://en.wikipedia.org/wiki/Square_wave
24. Weisstein, Eric W. "Fourier Series." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/FourierSeries.html

25. Picardo, CFA Elvis. “Moving Average - MA.” *Investopedia*, 23 Jan. 2018, www.investopedia.com/terms/m/movingaverage.asp.

26. Staff, Investopedia. “Volatility.” *Investopedia*, 4 Aug. 2016, www.investopedia.com/terms/v/volatility.asp.

27. “Q: What is a Fourier transform? What is it used for?” *Ask a Mathematician / Ask a Physicist*, 18 July 2017, www.askamathematician.com/2012/09/q-what-is-a-fourier-transform-what-is-it-used-for/.

28. Bevelacqua, Pete. “Fourier Transforms.” *Fourier Transform*, www.thefouriertransform.com/.

29. Weisstein, Eric W. "Fourier Transform." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/FourierTransform.html