HELLENIC OPEN UNIVERSITY

School of Social Sciences

Master in Business Administration (MBA)

Postgraduate Dissertation

"Exploring the Feasibility of Forecasting Stock Portfolio Returns: A Comparative Analysis of Traditional and Data-Driven Models in the Greek Stock Market."

"Alexandros KOTOUZAS"

Supervisor A: Dr. Timotheos ANGELIDIS

Supervisor B: Dr. Apostolos BALLAS

Patras, Greece, May 2026

“Exploring the Feasibility of Forecasting Stock Portfolio Returns: A Comparative Analysis of Traditional and Data-Driven Models in the Greek Stock Market.”

Alexandros KOTOUZAS

Supervising Committee

Supervisor:	Co-Supervisor:
Timotheos ANGELIDIS	Apostolos BALLAS
Professor, Hellenic Open University	Professor, Hellenic Open University

Patras, Greece, May 2026

“Acknowledgments

I would like to express my sincere gratitude to my Tutor, Dr. Timotheos Aggelidis, for his valuable contribution throughout my learning journey. Always available to share his knowledge and kind words unconditionally.

Dedication

To my little “monsters”. I hope they once forgive me for the time lost with them, but I am certain that they shall soon realise that an image sends the loudest message. And I was never really good in talking.

To the brightest star of my galaxy that is always there to shine for me in support, even in the darkest hours.

To my parents, who let me be who I want without judgement.”

Abstract

The construction of an optimized portfolio and the prediction of its future course has long been of interest among the finance community. The principal research question that this study aims to provide an answer to, is how feasible is it to forecast, and at which degree of accuracy and reliability, the range of returns generated by a portfolio consisting solely of Greece stocks.

To address this question, stock price and market data from the Greek market were collected for a specific sample period. A portfolio was initially built based on stock screening, personal preference, correlation analysis and weight optimization for the selected timeframe. Monthly returns were then calculated and used for model framing and forecasting. Capital Asset Pricing model and Fama French three factor model were implemented using regression analysis and were also used as prediction tools through a lagged factor technique. A Prophet model was also fitted on the same dataset that was further refined by applying Fama-French factors as external regressors to the base model. All models were then evaluated over the same out of sample period by employing specific evaluation metrics and compared against a naïve forecast.

What is realised through this study is that returns forecasting in the Greek market is feasible but with limited accuracy and reliability. Prophet models produced strong performance but their accuracy is highly dependent on the feature selection during the fitting process and the length of the training sample. Tradional models can be very effective in providing information for the characteristics of a portfolio based on past data. Still, the results suggest that, at least for the selected timeseries, the models deviate from the Efficient Market Hypothesis, since they generate abnormal returns as denoted by the intercept alpha. Unfortunately, they also exhibit minimal predictive accuracy for the testing period within the method context adopted in this research.

The researcher should therefore bear in mind that the models examined in this research may support portfolio analysis and provide an indicative range of returns for a portfolio in the Greek market. Nevertheless, these should be used with caution and in combination with a deeper analysis in the Greek market macroeconomic prospects and expectations.

Keywords

CAPM, Fama-French, Prophet, Portfolio, Stocks.

"Διερεύνηση της Δυνατοτητας Πρόβλεψης της Απόδοσης Χαρτοφυλακίου Μετοχών: Συγκριτική Ανάλυση Παραδοσιακών και Σύγχρονων Μοντέλων Δεδομένων με Εφαρμογή στην Ελληνική Χρηματιστηριακή Αγορά."

"ΑΛΕΞΑΝΔΡΟΣ ΚΩΤΟΥΖΑΣ"

Περίληψη

Η κατασκευή ενός βέλτιστου χαρτοφυλακίου και η πρόβλεψη της μελλοντικής του πορείας αποτελούν εδώ και καιρό αντικείμενο ενδιαφέροντος για τη χρηματοοικονομική κοινότητα. Το βασικό ερευνητικό ερώτημα στο οποίο επιχειρεί να απαντήσει η παρούσα μελέτη είναι κατά πόσο είναι εφικτή η πρόβλεψη, και με ποιο βαθμό ακρίβειας και αξιοπιστίας, του εύρους των αποδόσεων που παράγει ένα χαρτοφυλάκιο αποτελούμενο αποκλειστικά από ελληνικές μετοχές.

Για να απαντηθεί το παραπάνω ερώτημα, συλλέχθηκαν ημερήσια δεδομένα τιμών μετοχών καθώς και του Γενικού Δείκτη της ελληνική χρηματιστηριακή αγοράς για συγκεκριμένη χρονική περίοδο. Αρχικά κατασκευάστηκε ένα χαρτοφυλάκιο μετοχών βασισμένο σε επιλογή μετοχών μέσω, προσωπικής κρίσης, ανάλυση συσχέτισης και βελτιστοποίησης ποσότητας μετοχών κάθε εταιρίας για το επιλεγμένο χρονικό διάστημα. Στη συνέχεια υπολογίστηκαν οι μηνιαίες αποδόσεις όπου και χρησιμοποιήθηκαν για τη διαμόρφωση και πρόβλεψη των μοντέλων. Το Υπόδειγμα Αποτίμησης Περιουσιακών Στοιχείων (CAPM) και αυτό των Fama-French τριών παραγοντων εφαρμόστηκαν μέσω παλινδρόμησης και χρησιμοποιήθηκαν επίσης ως εργαλεία πρόβλεψης με τη χρήση μιας τεχνικής με παράγοντες υστέρησης. Παράλληλα, προσαρμόστηκε στο ίδιο σύνολο δεδομένων και ένα μοντέλο Prophet, το οποίο στη συνέχεια βελτιώθηκε περαιτέρω με την ενσωμάτωση των παραγόντων Fama-French ως εξωτερικών παλινδρομητών στο βασικό μοντέλο. Όλα τα μοντέλα αξιολογήθηκαν για την ίδια εκτός δείγματος περίοδο με τη χρήση συγκεκριμένων μετρικών αξιολόγησης και συγκρίθηκαν με ενα όριο πρόβλεψης αναφοράς.

Αυτό που προκύπτει από την παρούσα μελέτη είναι ότι η πρόβλεψη αποδόσεων στην ελληνική αγορά είναι εφικτή, αλλά με περιορισμένη ακρίβεια και αξιοπιστία. Τα μοντέλα Prophet παρουσίασαν ισχυρότερη απόδοση, ωστόσο η ακρίβειά τους εξαρτάται σε μεγάλο βαθμό από την επιλογή των χαρακτηριστικών κατά τη διαδικασία προσαρμογής στα δεδομένα, καθώς και από το εύρος της χρονοσειράς. Τα παραδοσιακά υποδείγματα μπορούν να είναι ιδιαίτερα χρήσιμα στην παροχή πληροφοριών για τα χαρακτηριστικά ενός χαρτοφυλακίου με βάση τα ιστορικά δεδομένα. Ωστόσο, τα αποτελέσματα δείχνουν ότι, τουλάχιστον για τη συγκεκριμένη χρονοσειρά, τα μοντέλα αποκλίνουν από την Υπόθεση Αποτελεσματικής Αγοράς, καθώς παράγουν μη κανονικές αποδόσεις, όπως υποδηλώνεται από τη σταθερά άλφα. Δυστυχώς, εμφανίζουν επίσης ελάχιστη προβλεπτική ακρίβεια για την εξεταζόμενη περίοδο πράγμα που οφείλεται στο μεθοδολογικό πλαίσιο που υιοθετήθηκε στην παρούσα έρευνα.

Λέξεις – Κλειδιά

Υπόδειγμα Αποτίμησης Περιουσιακών Στοιχείων, Χαρτοφυλάκιο, Μετοχές

Table of Contents

List of Figures

List of Tables

List of Abbreviations & Acronyms

CAPM	Capital Asset Pricing Model
EUR	Euro currency
FF3	Fama-French three-factor model
HML	High minus Low
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
OLS	Ordinary Least Squares
Prophet	Facebook Prophet
RMSE	Root Mean Square Error
SMB	Small minus Big
USD	United States Dollars currency

1 Introduction

The stock market, from its very beginning, has long served as a dynamic arena and fascinating playfield where individuals from diverse socio-economic backgrounds are attracted in pursuit of making profits. Although the driving forces for each person differ, ranging from fortune seeking and curiosity to professional ambition, all of them share the same goal of "beating the market" by gaining a slight edge through predicting market movements, even just one day in advance.

Over time, this pursuit evolved and matured within the academic society who developed increasingly sophisticated quantitative models (CAPM, FF3) to estimate expected returns and optimize asset allocation (Zhang, 2025). Nevertheless, the reliance of such models on linear trends and historic mean values often limits their applicability in unstable environments, particularly in emerging or semi-efficient markets like Greece as Maris (2009) and Michailidis (2009) highlight.

In recent years, breakthroughs in computation power and machine learning gave birth to advanced data driven algorithms, such as Prophet, which extended the boundaries of time series forecasting even further. This was made possible due to their capacity to model non-linear patterns and integrate seasonal and holiday factors as reported by Cheng *et al* (2023).

This research argues that auto-machine learning techniques, like Prophet, could offer increased predictive accuracy compared to traditional linear models (CAPM) and multifactor models (FF3) which mainly rely on financial time series forecasting. It focuses on a comparative evaluation of traditional linear time series (CAPM) and multi-factor models (FF3) and the Prophet tool, using historical data from an optimized stock portfolio constructed on the Athens Stock Exchange for a specific time frame. In addition, this thesis explores whether portfolio returns forecasting can be further improved by assigning the factors from FF3 as regressors in the Prophet model. The value of this research is concentrated in the belief that reliable forecasting techniques are important in taking informed decisions and shaping economic strategies for the future.

1.1 Background on portfolio return and forecasting challenges

Forecasting portfolio returns with accuracy is a major as well as challenging problem due to the inherently complex and uncertain nature of all financial markets around the globe. Financial time series are typically characterized as being non-linear, noisy and in most instances chaotic as per Thanh *et al.* (2023). This was realized and theoretically formed many years before with the Efficient Market hypothesis by Fama (1991) where it is stated that asset prices fully and instantaneously reflect all available information. In practice, this implies that markets follow a "random walk" and are unpredictable.

Nevertheless, the previous discussion has not deterred the academic community from trying to forecast stock prices. Early theoretical developments in modern portfolio management and asset pricing like Markowitz's (1952) introduction to diversification, Sharpe's (1964) revolutionary model of capital asset pricing and the Fama-French (1992) multi-factor model examined a structured way to associate expected returns to risk factors and introduce the essence of diversification in finance.

Progressively, it was realised that these models failed to consistently capture the financial market trend and fluctuations. As computational power advanced with time, the focus increasingly shifted towards exploring new approaches to financial market prediction. Specifically, the most commonly used models for prediction nowadays involve support vector machines and neural networks, as per the recent study of Henrique *et al* (2019). However, such modelling requires from the analyst to have advanced mathematical skills and deep knowledge in machine learning techniques and programming.

In response to this challenge, user friendly open-source tools such as Prophet emerged, offering an alternative that combines time series forecasting with elements of machine learning algorithms. By simplifying model development and results interpretation, sophisticated algorithms can now be utilised by users with limited programming experience but with strong interest in quantitative analysis, thereby broadening the cycle of people involved in financial research.

1.2 Research aim and objectives

The primary aim of this thesis is to investigate the feasibility of forecasting stock portfolio returns using Prophet as well as tradition models (CAPM, FF3) and offer a practical solution for capturing market movements in the context of the volatile Greek stock market.

The research questions implicate the following:

- Can traditional models be applied in forecasting stock portfolio returns in the Greek market and if yes how do these perform?
- Can data driven models, like Prophet offer equivalent or even better predictive accuracy and reliability than traditional ones and therefore be rendered useful to the financial community?

To address the latter, the current study employs a comparative analysis between the CAPM and FF3 with Prophet which model portfolio returns from the Athens stock exchange. The key objectives of the research are summarised hereunder:

- Create a diversified portfolio of stocks from theATHEX Composite Share Price Index pool by collecting and processing historic stock data prices and aggregate these in monthly returns.
- Create the CAPM model assuming the Market portfolio as the ATHEX Composite Index and the Greek three-months Treasury Bill rate as a proxy for the risk-free rate. Perform regression analysis to estimate beta.
- Create the FF3 model using already constructed factors for Greek stocks from reputable internet sources. Find the relevant slopes (betas) and intercept (alpha) using regression analysis.
- Create the Prophet model built upon optimal feature engineering and improve its predictive accuracy using additional regressors from FF3.
- Forecast monthly returns for each model for a specific time frame and evaluate these using standard statistical metrics (MASE, RMSE etc.).

1.3 Scope

This section delineates the boundaries and scope of the present study. The central focus of this research is the empirical analysis and forecasting feasibility of a single diversified stock portfolio that exclusively contains equities listed at the Athens Stock Exchange. The main concept is summarised in whether portfolio returns can be forecasted and up to which degree of reliability with the use of traditional asset pricing models, such as the CAPM and FF3, as well as contemporary data-driven algorithm, namely Prophet.

The dataset is comprised from daily historic stock prices spanning from January 2018 to December 2023 for the training period and January 2024 to December 2024 for the testing period. This includes both "bear" and "bull" markets as well as the COVID-19 pandemic and aftershocks.

The thesis further includes portfolio weights estimation, regression analysis for CAPM and FF3 for a testing period as well as an attempt to forecast monthly returns using a lagged explanatory variables concept. In parallel, a Prophet model is instantiated, trained and tested in forecasting monthly portfolio returns on the same period as the traditional models. In extent, a refined version of the Prophet model, using FF3 variables as external regressors, is fitted and tested to the same dataset.

Typical performance metrics, like RMSE, MAE and MASE, were employed to assess the predictive accuracy of all the models while these were also compared against a naïve forecasting benchmark.

1.4 Thesis outline

Apart from the current introductory section the thesis outline is organised as follows:

Chapter two begins by summarising the literature, starting from the basic concepts of modern portfolio theory up to contemporary models for stock price forecasting.

The third chapter details the presumptions and methodology employed for constructing the financial models under review. All challenges faced during the modelling phase are also included in this section.

Section four presents the results of the analysis, discussed the forecasting accuracy and lays out limitations and challenges faced.

The thesis is concluded in the fifth section with proposals and a roadmap for further research.

2 Literature Review

2.1 Overview of modern portfolio theory and stock returns

The founding of Modern Portfolio Theory, by a young post-graduate student called Harry Markovitz back in 1952, remains one the most significant breakthroughs in the finance industry. It transformed investment selection from a subjective practice into a coherent quantitative science. Before Markovitch, investors focused on picking stocks based on their highest expected returns. Modern Portfolio Theory introduced a structured framework for constructing an efficient investment portfolio.

Markowitz's (1952) fundamental idea was that risk and return are inherently and inseparably linked based on the traditional notion "Nothing ventured, nothing gained". He formalised the concept of a portfolio that produces the maximum expected return for a given level of risk measured in terms of standard deviation of returns. The difficulty faced is that risk and return are mutually exclusive. In essence, no investor can achieve above average gains without assuming above average risk. In his paper he argued that the two critical inputs associated with this analysis are the expected returns of each security and the covariances between security returns i.e. the variance-covariance matrix.

Reily and Brown (2000) argue that the Markowitz portfolio selection model is structured on the following assumptions regarding investor behaviour:

- Investors view each asset as having a probability distribution of possible returns over a given holding period. Since the future is uncertain, every asset return is a collection of possible outcomes with probabilities.
- Investors maximise expected utility over a single period and their utility functions exhibit diminishing marginal utility of wealth. In essence, the satisfaction gained for every unit of profit diminishes as the profits increase.
- Portfolio risk is measured by the variability of returns and is captured by the variance or standard deviation.
- Investment decisions are based only on expected return and risk, so investor preferences can be represented as a function of these two variables. In this context, they process stock information correctly (i.e. they are rational) and are risk averse.

Other general presumptions of the model are:

- No friction in market i.e. there are no taxes and transaction costs on trading. Assets can be bought or sold in any quantity.
- All investors have the same costless and instant access to all relevant information.
- Prices can not be influenced by a single investor i.e. everyone is a price taker.

Although modern portfolio theory provides the framework for informed investing decisions, its application in real world and specially in markets like Greece can be challenging. As mentioned above the theory relies on assumptions, among other, such as perfectly efficient market conditions and investors who act based on fundamental data and rationale rather than emotions. Yet, in most cases, these conditions do not apply in the Greek reality. Following the arguments put forward by Giouvritsa (2016), it appears that the Greek stock market behaves inconsistently while it generally displays low liquidity and strong reactions to economic and political events. In fact, it shows signs of both efficiency and inefficiency, reflecting its mixed status between a mature and an emerging market. Despite these limitations, the concept of balancing between risk and return (suggested by Modern Portfolio theory) remains highly relevant and shall be used in the empirical study section of this document for the selection of the optimal portfolio.

The following section shall now focus on one of the most important ideas derived from the Modern Portfolio theory and that is the relationship between risk and return and how pivotal diversification is in forecasting portfolio performance.

2.2 Risk-return trade off and diversification

The finance community is used to think of risk in terms of statistical metrics, like variance or volatility or standard deviation. In our view, the most intuitive definition of risk is given by Aswath Damodaran (2014) as “the mix of danger and opportunity” inspired by the ancient Chinese symbol for crisis.

Risk is therefore neither good nor bad. It is simply an integral feature of investing activity. The puzzle that investors have to answer is, therefore, not how to avoid risk but how to integrate it effectively into their decision-making process. Markowitz’s portfolio theory addresses this issue directly by modelling risk as a marginal cost and return as a marginal benefit. The output of this optimisation process is called the Efficient Frontier and in reality is that set of stock portfolios that offers the maximum anticipated return for every given level of risk or the minimum risk for every level of return (Ganti, 2025). The Efficient Frontier (otherwise called the minimum variance portfolio) is graphically represented as a curved line with risk (standard deviation) on the x-axis and expected returns on the y-axis. A non-numeric example of an Efficient Frontier is illustrated in Figure 1 where portfolios lying below the rising upper part the curve are suboptimal since they provide inferior returns for the same level of risk.

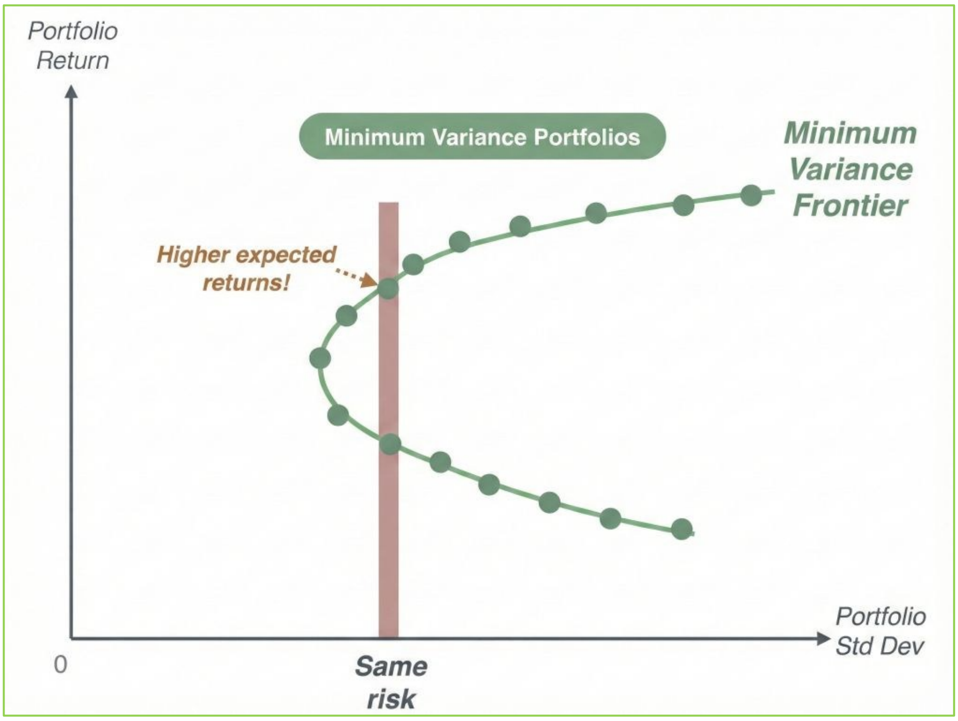


Figure 1: Efficient Frontier

Source: (PrepNuggets, 2022)

Markowitz's contribution was to demonstrate that the risk of a portfolio is not a simple weighted average of the risks of individual assets but rather a function of the covariances among assets. This insight laid the foundation for the concept of diversification, which remains one of the cornerstones for modern asset management.

Diversification can be defined as the combination of securities that comprise an efficient portfolio which has a correlation coefficient less than positive one. The effectiveness of diversification increases as the correlation between assets decrease since lower correlations generate a greater reduction in total portfolio risk (Papadimitriou, 2019). This concept is clearly illustrated in Figure 2 where a set of possible portfolios formed from risky assets (1 and 2) is plotted for various correlation coefficients (denoted as r_{ij}). Expected return is once again in the y-axis and risk, measured as the standard deviation of returns, in the x-axis. When correlation is positive and close to $+1$ there is no effect on diversification i.e. no risk reduction for greater returns. On the contrary, as correlation declines towards -1, the curve bends to the left suggesting lower risk for the same expected return. Hence, the benefits of diversification are critically dependent on the correlation between assets.

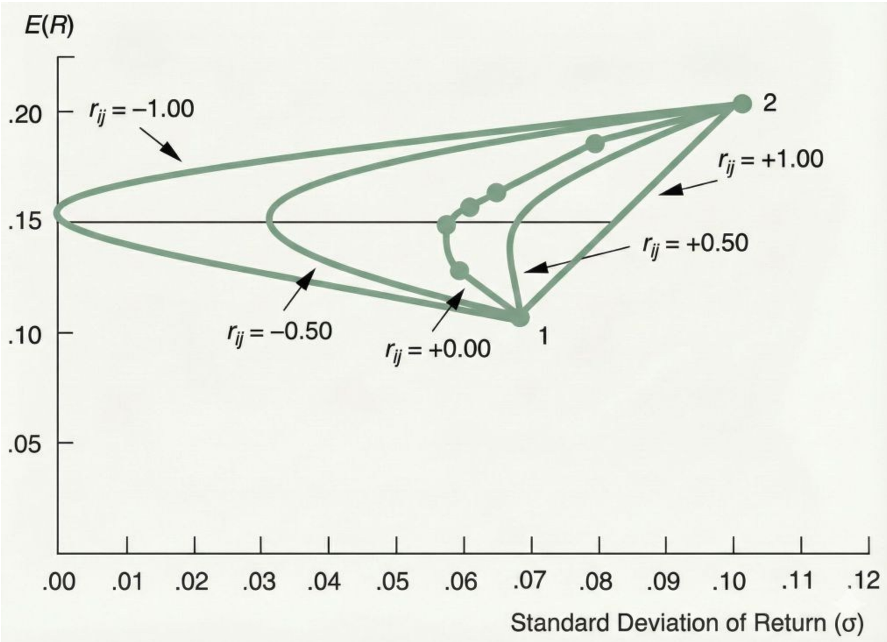


Figure 2: Risk-return plots for different portfolio weights

Source: (Reilly & Brown, 2000) page 226

Consequently, the next important question to consider at this point is how many stocks are required to achieve a completely diversified portfolio. Brealey *et al.* (2015) attempted to provide an answer to this question through their study. By computing the standard deviation for portfolios of increasing numbers up to twenty U.S. stocks, while excluding transaction costs, they found that approximately 90% of the diversification benefit was derived from portfolios containing twelve to eighteen stocks.

In the context of the Greek stock market, Papadimitriou (2019) argues that diversification plays a vital role in enhancing portfolio efficiency. By selecting stocks, during the period 2014 to 2018, with varying behaviour and by applying modern portfolio techniques she showed that properly diversified portfolios (like the tangent and efficient portfolios) outperformed others in terms of risk adjusted returns. Despite the general decline and low liquidity during the Greek financial crisis, her findings confirmed that combining assets with low correlations leads to better risk-return trade offs. Unfortunately, this study also omits to quantify the number of securities at which diversification benefits are diminished. From our knowledge, this specific research question has not been broadly investigated for Greek securities. Anyhow, most study results are particularly encouraging and support the argument that a diversified portfolio sets the foundations for accurate performance assessment and reliable forecasting.

Moving on now to consider the two primary types of economic risk and further elaborate on their importance in financial modelling.

2.3 Systematic vs unsystematic risk

Building on the discussion of diversification and risk from the previous section, it is now useful to place these in a specific financial dimension. Financial textbooks decompose the total risk on an investment in unsystematic and systematic segments.

Unsystematic risk, otherwise known as firm specific risk, affects specifically a single or a small group of assets with similar characteristics and typically results from the way each firm is managed (Koutsianikouli, 2022). It arises from factors such as management decisions, competition, selected financial model etc. Since unsystematic risk affects a few firms, it tends to average out. For every company where something better than expected happens then in another company something worst than expected will happen (Damodaran A. , 2014).

In contrast, systematic risk is the variability in stock returns due to changes in economic factors that affect the entire business environment. From data collected from various sources these could typically be associated with changes in macroeconomic variables like:

- Interest rates
- Government spending
- Oil prices
- Foreign exchange rates

Systematic (market) risk can change over time depending on the market conditions where the business operate in. This type of risk affects all business operating in a market but not at the same level and can not be diversified away. The impact of systematic risk could, therefore, be condensed in tracking the rate of return in a market portfolio which can be simulated as a portfolio including all assets of the economy (bonds, stocks etc.). In practice this is represented by a market index (Brealey et al., 2015).

The relevant measure of an asset's exposure to systematic risk is the beta coefficient (β), which captures the sensitivity of a stock's return to the directions (highs or lows) of the market portfolio. Betas usually vary from values between zero and two. The greater the value of beta from one, the riskier and more volatile this security is considered. Similarly, portfolio betas are the average of the betas on that portfolio weighted by the investment in each security. In practice, it portrays the volatility of the entire portfolio. Practically, portfolios with betas greater than one are considered aggressive and can produce greater profits in a rising market but also sustain greater losses at times when markets decline. Betas close to zero are thought to be defensive and provide more stable returns whereas negative betas are highly unusual and signify an asset whose performance is inverse in relation to the market movement (Forjan, 2025).

The empirical studies performed within the Greek Stock Market by Giouvritsa (2016) and Maris (2009) suggest that unsystematic risk tends to remain at relatively high levels. This is mainly attributed to the following structural characteristics:

a) the small size of the Greek market with few listed companies as opposed to major international markets which enlist category "killer" enterprises

b) the greatest share of the market value is highly concentrated in few industrial sectors such as banking, construction and retail.

Collectively, from the discussion in §2.2 hereabove and the present section, we can argue that although diversification benefits are present in the Athens stock exchange, they are considerably limited. This essentially arises from the facts that investors have fewer investment and sectors options to choose from and that the market is primarily guided by a small number of core industries.

In view of all that has been mentioned so far, it becomes important to assess how risky assets interact with a potentially riskless asset. From the perspective of modern portfolio theory, the introduction of a risk-free asset enables investors to establish a benchmark against which they can decide how much risk they are willing to undertake and the return that is associated with that risk within their selected portfolio. The next section analyses this "riskless" rate and its fundamental importance to the set-up of financial models like the CAPM and FF3.

2.4 Risk-free rate

What is commonly accepted is that there is no risk in the past, all risk is in the future. Sadly, all risk measures depend on past performance but uncertainty always lays ahead in time. On this basis, investors try to mitigate this unsettling situation by finding a safehouse in risk-free assets which provide risk free returns.

Based on a recent article from Adam Hayes (2025), the concept of the risk-free rate is based on the existence of a hypothetical risk-free asset. Such risk-free asset actually represents an investment that almost guarantees a specific future return with virtually no risk of financial loss. Given that the risk-free rate is a theoretical concept and accepting that an investment that bares no risk does not exist, in practice the risk-free asset is approximated with highly secure government bonds or treasury bills that have infinitesimal chances of default. Countries around the world set their own level for the risk-free rate as the economic conditions and geopolitical challenges vary between them.

While a truly risk-free investment is not realistic, as no one in any market can be compensated for all the risk they undertake, it still has a profound importance in the construction of financial modelling. The risk-free rate forms the baseline against which other investments are compared with. It is the key ingredient to form the Capital Market Line, a line with a starting point from the risk-free rate and tangent to the efficient frontier. So, the Capital Market Line denotes the expected return of a portfolio as a linear function of the trade off between risk and return for a portfolio that consists of the risk-free asset and the market portfolio discussed previously. This tangency portfolio represents the optimal portfolio where anyone can choose their preferred point on the Capital Market Line by selecting their own asset weights depending on their risk tolerance.

Investors require practical tools that systematically classify the relationship between risk and return, especially when a risk-free alternative is available. One of the most widely used metrics that captures exactly this concept is the Sharpe Ratio. It provides a reliable indicator of a portfolio's excess return relative to its overall volatility while it represents a key metric for evaluating portfolio performance in risk adjusted terms. As such, it is worthwhile to investigate its function and usefulness at the next section.

2.5 Portfolio performance metrics

Evaluating portfolio performance is a central concern both in the financial industry and throughout the academic community. Simple portfolio returns fail to quantify the risk undertaken to achieve those returns and are therefore insufficient indicators for assessment of portfolio performance. To tackle this issue, several risk-adjusted performance measures have been developed, indicatively naming the Treynor Ratio, Jensen Alpha and the Sharpe ratio. Each of these metrics adopts different characteristic and approach portfolio evaluation uniquely while they gradually became indispensable tools for analysts.

The Sharpe ratio is probably the most widely used scalar metric of risk adjusted performance. Originally formalised by William Sharpe (1966), this ratio gauges a portfolio's excess return over total risk (measured in standard deviation units) in a systematic way. In principle, the Sharpe ratio is the percentage of excess return for every unit of exposure to total risk. This makes it particularly useful for ranking portfolios, like mutual funds for which it was originally designed for, with various risk profiles as it encompasses both return and risk into a simply interpretable metric.

To gain a better understanding of the Sharpe ratio, a visual interpretation of it is illustrated in Figure 3. The y-axis represents the average portfolio returns per annum against their annual risk, which is in standard deviation terms. So, the Capital Market Line is drawn from a hypothetical risk-free asset as the tangent to the Efficient Frontier. Geometrically, the property of the tangent is that it is the steepest possible line thus denotes the highest maximum slope. This is mathematically modelled as:

$$Slope = \frac{\Delta y}{\Delta x} \tag{2.1}$$

and by substituting the annual portfolio risk (Δx) and the annual mean returns (Δy), the maximum attainable Sharpe ratio is obtained as:

$$\frac{\Delta y}{\Delta x} = \frac{Portfolio\ Return - Risk\ Free\ Asset}{Portfolio\ Risk\ (standard\ deviation)} \tag{2.2}$$

The numeric interpretation of the Sharpe ratio is straightforward. The portfolio with the highest Sharpe ratio is the best performing one. This simplicity has made the ratio the most popular for academic studies and portfolio managers specially when portfolio performance is evaluated across different markets or investment strategies.

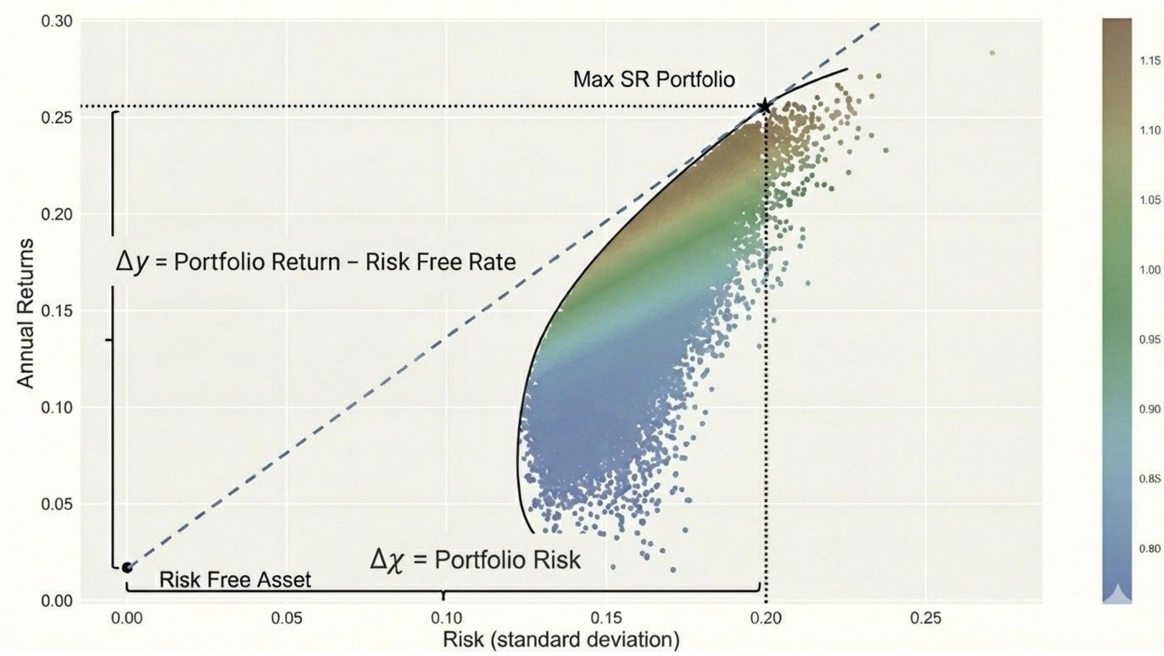


Figure 3: Visualization of the Sharpe ratio

Source: (Hagmann, 2025)

Nonetheless, the academic community's criticism on the Sharpe ratio, highlighted in the study of Kousiou (2005), focuses on the fact that it utilises an investment's total risk in the denominator. This implies that the examined portfolios are not diversified or alternatively that portfolios constitute an individual's total personal wealth. This presumption is unrealistic as a single investment is just one component of the investor's total equity. For this reason, critics argue that the focus should be shifted mainly to the systematic (market) risk that an investment adds to a diversified portfolio.

Despite this concern, the Sharpe ratio remains a key metric in portfolio analysis. In essence, it laid the groundwork for asset pricing models, as the CAPM, to refine the risk-return analysis and focus exclusively on systematic risk measured by volatility (beta). The following section expands on the CAPM principles and explores how systematic risk effects expected returns.

2.6 Overview of traditional models

2.6.1 Capital Asset Pricing Model

The development of the oldest risk and return model in finance, called the Capital Asset Pricing Model or CAPM, signalled a new era in capital market theory. Building on the foundations of modern portfolio theory, William Sharpe (1964) envisaged a model which, despite its weaknesses, is still widely used particularly in cost of capital estimation for corporate finance and portfolio performance evaluation.

According to Zhang (2025), the central idea behind CAPM is that diversification can phase out unsystematic risks and that investors are rewarded only for taking systematic risks. The simplicity of this concept has contributed for the CAPM to be one of the most appealing among finance community. In essence, CAPM states that the expected return of any individual stock is a linear function that begins with the risk-free rate and adds the difference between market return and risk-free rate (called market risk premium) weighted by the level of systematic risk denoted by beta. Hence, the CAPM equation is mathematically expressed as:

$$E(R_i) - R_f = \alpha_i + \beta_i \cdot \left[E(R_m) - R_f\right] \tag{2.3}$$

where

$E(R_i) - R_f$ is the expected return of an investment

R_f is the risk-free rate of the market that the security is associated with

$E(R_m)$ is the expected return of the market

α_i is the intercept term, which denotes the abnormal performance of asset i

β_i is the sensitivity of the asset i to market volatility and can be further decomposed as the ratio of the covariance of asset i with the market portfolio over the market's portfolio variance as Manolakis (2012) elegantly describes. The methodology employed to estimate betas, in both international and Greek empirical studies, is typically through time series regression analysis of excess stock returns on excess market returns.

CAPM can be visualised by the plot of a Security Market Line that represents the relationship between risk and expected return of an individual asset (Forjan, 2025). A typical example of the Security Market Line is displayed in Figure 4. Essentially an investment plotted above the Security Market Line is considered undervalued because it offers greater return against the risk it carries whereas if it is located below, it is overvalued as the compensation is insufficient for the risk it bears.

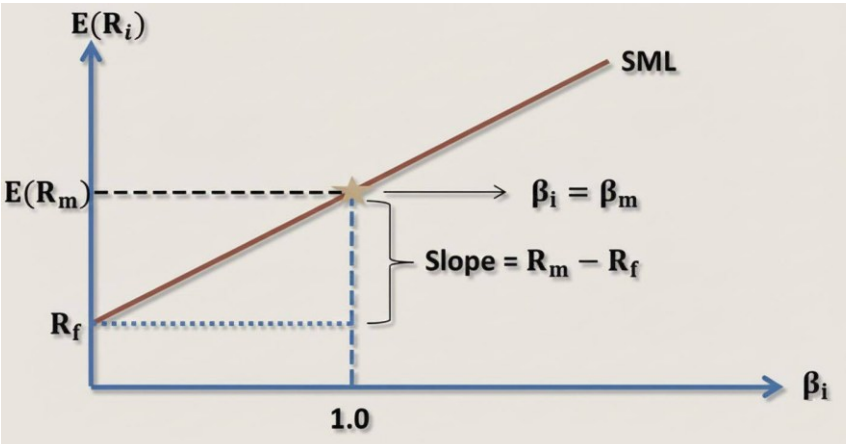


Figure 4: Security Market Line (SML)

Source: (Forjan, MPT and CAPM (FRM Part 1 2025 – Book 1 – Chapter 5), 2025)

Among several others, for CAPM to hold the following main underlying presumptions should apply, as reported by founder:

- market is frictionless (i.e. no transaction costs or taxes exist)
- investors have equal and free access to all data and information related to securities (i.e. no one knows which stocks are overvalued or undervalued)
- no individual investor can influence market price
- investors are rational and risk averse i.e. try to mitigate risk.

It is logical to consider that the previous assumptions are in most instances unrealistic and that, generally, the model would not perform well in practice. However, as Reily and Brown (2000) argue in their book on investment analysis and portfolio management, two points are important. First, by relaxing most of these assumptions the impact on the model would be limited as its main implications and output would be essentially unchanged. Second, a theory is tested on its predictive accuracy in real world examples rather than on the premises it is based on.

With regards to the application of the CAPM in the Athens stock exchange it can be argued that most studies fail to converge in similar conclusions. For example, Angelopoulou (2018) applied the CAPM to equities belonging to the Greek General index over the period 2007-2016. She pointed out that testing the validity of CAPM within the Greek Stock Market might not generate reliable results as certain aspects of the market produced noise in her analysis. She further argued that the Greek market is volatile, driven by psychological factors, inefficient and there seemed to exist a lot of information asymmetry. She eventually suggested that the CAPM would be more suitable for mature markets.

A detailed summary of various studies on empirical application of the CAPM in the Athens Stock Exchange is also done in the postgraduate dissertation of Michailidis (2009). The author identifies that:

- there is no clear relationship between expected returns and systematic risk
- during periods of macroeconomic instability, such as the Greek debt crisis, systematic risk dominates the market, neutralising diversification benefits and subsequently reducing the explanatory power of the CAPM.

From all the studies reviewed here, it is observed that the CAPM displays a lack of consistency when applied to less mature markets as the Greek one. This aligns with other international evidence showing that it occasionally underestimates the influence of firm-specific factors and macroeconomic fluctuations. For instance, Khoa and Huynh (2023) confirmed similar limitations in the Hanoi stock market. They emphasized that the CAPM could not rigorously explain why some portfolios achieved higher or lower projected returns than others and this was attributed to its lack in explanatory variables. Their conclusion was that if the CAPM was the best forecasting model it should explain most of the difference in returns across stocks over long periods of time.

The issues and reservations, with respect to CAPM forecasting ability, raised so far by the researchers mentioned are well noted and shall be considered within the present study. Apart from that, the CAPM remains a fundamental tool in portfolio performance evaluation and is a reference point for returns forecasting in the Greek stock market as well. Its simplicity allows us to define a baseline expected return against which a comparative analysis can take place with multifactor models and data-driven approaches that attempt to capture stock return variations more efficiently.

This naturally leads to the following section where another important model of financial analysis is presented. Using the CAPM as the starting point, FF3 further extends it by adding more explanatory variables in the linear equation. This transition from a single factor to a multifactor model signifies a turning point in asset pricing theory, leading to a broader and more pragmatic approach in the comprehension of the effect that different determinants have on stock returns. Whether such an approach could be more adaptable in the Athens stock exchange remains to be determined in the next chapters of this study.

2.6.2 Fama-French Three Factor Model

In the early 1990s several financial researchers started to question whether the risk of an asset is only associated with the risk it contributes to the market portfolio as proposed by the CAPM. The concern emerged from noticing that assets with similar returns could have very different correlation with the market and since then a long-lasting debate started on whether CAPM betas could sufficiently explain returns.

As Damodaran (2014) perfectly describes on his online corporate finance course, the Fama-French proxy model critique is based on the notion that if CAPM betas do not have enough predictive accuracy then other factors should be explored that may generate explanatory power to the model. On this basis, Fama and French (1992) developed an extended version of the CAPM by adding two more explanatory variables, namely the firm size and the book value to market equity ratio.

By analysing several years of data, Fama and French discovered that there is a higher risk associated with smaller companies as opposed to larger ones hence small companies yield higher expected returns than larger ones. Looking at the value factor, companies with a high ratio of book equity to market equity have higher returns than those with lower.

Regarding the latter, all these factors contribute to risk and appears to better explain why some stocks have high (or low) average returns than the CAPM predicts. In practice, FF3 refines the risk measure of the CAPM by considering that, apart from the market, risk is also linearly correlated with the size and value factors. These observations were formalised through the following expression:

$$E(R_i) - R_f = a_i + \beta_{i1}\left[E(R_m) - R_f\right] + \beta_{i2}SMB + \beta_{i3}HML + \varepsilon_i \qquad (2.4)$$

where

$E(R_i)$ is the expected return of an investment on an asset (or portfolio)

R_f is the risk-free rate of the market that the security is associated with

$E(R_m)$ is the expected return of the market

α_i is the intercept term, which denotes the abnormal performance of asset i after controlling its exposure to the market, firm size and book to market equity factor. It should be equal to zero provided that the three factors capture the totality of systematic risks (Forjan, 2019).

β_{i1} is the sensitivity of the asset i to market volatility as per the CAPM

β_{i2} is the sensitivity of the asset i to the firm size (small or large in terms of capitalisation)

β_{i3} is the sensitivity of the asset i to the value factor ("growth" or "value" stock)

SMB is the size factor $(Price\ of\ share \times Number\ of\ shares\ outstanding)$ constructed as the difference in returns between small capitalisation and large capitalisation stocks (Small minus Big)

HML is the value factor, constructed as the difference in returns between portfolios of high and low book to market value stocks (High minus Low) where book to market value is book value per share divided by the stock price

ε_i is the error term otherwise known as the random noise around asset i.

Maris (2009), in his effort to investigate how the FF3 models applies in the Athens stock exchange, tested it over a period of ten years (Jully 1999 to June 2009). According to him, the timeframe included both periods of stability and volatility. Unfortunately, the results of his research were rather discouraging and puzzling as he reports that:

- all three factors were found statistically insignificant

- returns of large cap firms were greater than those of small ones i.e. opposite to what the model expects
- the error terms exhibited heteroscedasticity

Angelopoulou (2018), computed stock returns from 2007 to 2016 on a yearly basis using the FF3 model. The study concluded that certain factors of the model agree with international empirical findings. In particular:

- the market risk premium was statistically significant and had a positive impact on returns
- the size factor was also positively statistically significant

However, the book to market value was found to be statistically insignificant while the overall explanatory power of the model was as low as 24%. An interesting remark on the study is that an investor who selects small cap stocks of high beta in the Greek market is expected to receive greater returns.

With time, the FF3 evolved and into a four and eventually a five factor version integrating factors related to momentum effects, profitability and investment mode. Further empirical evidence presented by Mougiou (2020) do not support the existence of momentum profits in the Greek stock market. In addition, Zhang (2025) documents that as the number of factors in the Fama French model increases then it:

- becomes more complex
- requires a vast amount of historic data to evaluate all parameter
- is more difficult to interpret it.

On this basis, there seems to be no added value in using more than three factors in the present study.

Overall, the FF3 is still highly appreciated and frequently used by the finance community in our days. Provided we acknowledge that it carries the same theoretical limitations as the CAPM, it has been a common belief among most scholars that it could also offer better predictive accuracy. Evidence from research on the Greek market show that the predictive power of the model is limited mainly due to the small size of the market and its high volatility. For these reasons, once the model is applied to our portfolio it is imperative to critically assess the output and recall that the model is used as a useful benchmark rather than a flawless forecasting tool.

If findings of Greek studies are accurate, traditional regression based linear models interpret returns without satisfactory accuracy on the Greek market. The complexity and randomness of the data lead to nonlinear patterns. This drives us to look beyond conventional approaches and investigate alternatives that could be more efficient in capturing such complex patterns. A very promising option would be that of Prophet, which shall be detailed in the following section.

2.7 Overview of data-driven time series models – Prophet

In recent years, significant advancement in the field of data science has produced numerous business-oriented algorithms. This trend came along with the boosting in computation power as well as the development of modern computer languages giving rise to highly sophisticated predictive tools. Such tools have become widely known because they are open source and therefore costless, while they are matched with a vast availability of data that is now, more than ever, freely accessible in many cases.

One such algorithm is Facebook Prophet (Prophet). Prophet is a time series forecasting model developed by Taylor and Leatham (2017), during their employment for Facebook social media platform, designed to handle common features of business time series data. According to their published instructional paper, where much of the information presented in this section is extracted from, Prophet was built with the objective of "forecasting at scale", which essentially means that:

- the model facilitates active engagement from the analyst requiring only basic familiarity with time series models while retaining the benefits from their domain expertise to critically evaluate the output of the forecasting
- a wide range of forecasting problems can be modelled by multiple users
- a large number of forecasts is feasible whereas through model cross validation and by using human feedback the predictive accuracy could be further improved.

Prophet unites structural times series components such as trend, seasonality and the effects of holidays into an additive model. These are combined in the following equation:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \qquad (2.5)$$

where:

Trend component $g(t)$: Captures non-periodic changes and identifies points (called “changepoints”) at which the time series displays sharp declines or rises

Seasonality component $s(t)$: Models (using Fourier series) periodic seasonal patterns examining weekly, monthly, yearly effects

Holiday component $h(t)$: Represents the effect of non-periodic seasonal effects (holidays) occurring over one or more days

Error term ε_t: refers to any idiosyncratic changes which can not be explained by the model (noise)

The underlying mathematical concepts behind the components mentioned hereabove are beyond the scope of this dissertation. Nonetheless, it could be useful to point out that Taylor and Leatham report that the entire model fitting is formulated in matrices built in “Stan” code, a Bayesian library designed for statistical modelling with high-performance computations.

The innovative notion of “analyst in the loop”, as Jahangiry (2025) aptly states on his online course, upgrades the role of the analyst within the forecasting at scale approach as demonstrated in the following schematic view.

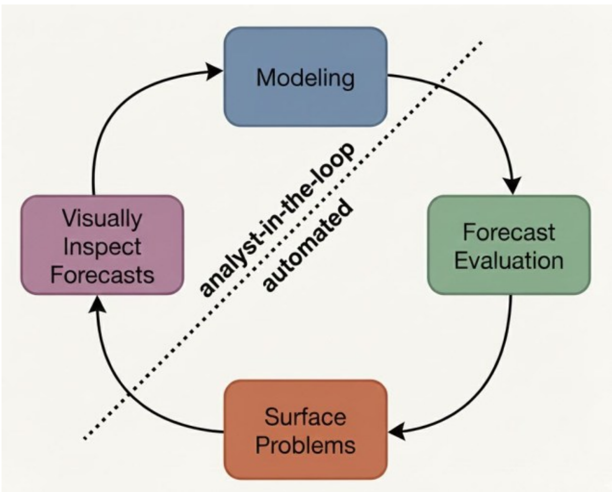


Figure 5: Analyst in the loop automation

Source: (Taylor & Letham, 2017)

We could practically summarise the key advantages of Prophet as flexibility, fast model fitting, ability to cope with irregularly spaced observations without the need for interpolation and domain knowledge utilisation as analysts can apply their insights to the model by adjusting specific parameters. Besides these, its uniqueness lies in its ability to capture certain nonlinear trends and multiple seasonality where traditional models fail to cope with.

Undoubtably, like any other forecasting model, Prophet has limitations that need to be referred and taken into account during the model fitting and output review. The biggest caveat is that the model is not autoregressive i.e. its predictions are not directly based on the input of recent historic values e.g. yesterday's, last week's etc. Instead, the model is primarily designed as a curve fitting approach. Also, the default settings involve a model where trend, seasonality, holiday event effects and noise components are combined additively or through simple transformations. In result, the model can not properly evaluate how strong each effect actually is. Another issue occurs when time series data present very unstable or highly irregular seasonality, like stock prices in certain occasions, which could eventually negatively affect the model's predictive strength.

In spite of these limitations, the inherent ability of Prophet to process non-stationary, high variance data even when these exhibit multiple seasonality, has made it particularly attractive in the finance world, specially in the domain of cryptocurrency price forecasting.

Angelo *et al.* (2023) compared Prophet with classical time series models (AutoRegressive Integrated Moving Average – ARIMA) for Bitcoin (type of cryptocurrency) price forecasting and found that Prophet outperformed ARIMA over a two year period for daily and weekly horizons in terms of predictive precision. Though, they also concluded that other methodologies could be further reviewed as Prophet is essentially a univariate approach with the closing price being the single variable in their study.

Similar research was performed by Indulkar (2021) on a panel of cryptocurrencies, namely Bitcoin, Bitcoin Cash, Ethereum, Chainlink and XRP. A model was constructed for each cryptocurrency for which the data were split into training and testing sets. The models were then used to make a prognosis of prices for the following fifteen days while the mean absolute error was employed as the evaluation metric. Patterns of predicted and actual prices were quite similar indicating a good overall convergence of the model with reality, as shown by the low error values, suggesting that Prophet could effectively track short horizon trends when seasonality is correctly specified.

More recent evidence, however, suggest otherwise. Prophet was outperformed by deep learning algorithms when applied to daily trading data and volatility metric (Garman-Klass) of Bitcoin from January 1st 2017 to October 30th 2022, in a recent study by Cheng *et al.* (2023). The authors also report that Prophet was also found underperforming in periods of extreme turbulence such as the Russian invasion in Ukraine and the COVID-19 pandemic.

Apart from cryptocurrencies, Prophet precision has also been tested in stock price forecasting from Annapoorna *et al.* (2024). The authors proposal focused in predicting Tesla's stock price over a thirty day horizon. They showed that the model managed to effectively predict the overall trend, on a short-term basis, as the average absolute difference between the forecasted and actual stock price was around 15.7%. They also emphasize Prophet's strengths in data visualization easiness and data interpretability while claiming that the model fitting weakens when faced with non-stationary long-term trends or extremely noisy and complex patterns.

The applicability of Prophet in the Greek stock market remains to be examined as no relevant study appears to have been published by the time this thesis was being drafted. The last few years, research on predictions of stock returns in the Athens stock exchange seems to be primarily performed through complex machine learning models as reported in relevant publications by Tzanetis (2024) and Ioannidis (2024). What is likely expected is that Prophet could have a slight advantage over regression models in data fitting over short term periods that exhibit moderate fluctuations in trend. Whereas it shall probably deteriorate in long term forecasting as macroeconomic factors start to play a more significant role in returns. Nevertheless, as Zhang (2025) notes, combining traditional asset pricing models with data driven forecasting techniques can improve portfolio stock price predictions, specially in emerging and semi efficient markets that could simulate to the Greek one.

The key thing about Prophet is that it represents a package able to handle changes in trend, multiple seasonality, missing data and outliers which makes it particularly attractive for forecasting stock returns in markets with attributes like the Greek one. By understanding its limitations, including it alongside with CAPM and FF3 models in the assessment of feasibility of forecasting stock portfolio returns in the Greek stock market is strongly supported.

For the objective assessment of all three models, we need to rely on statistics and more specifically on quantitative evaluation metrics that measure the difference between actual and forecasted returns. The following chapter reviews the literature for such error based metrics.

2.8 Evaluation metrics for predictive models

As previously stated, most authors have recommended specific metrics to evaluate forecasting accuracy. All such metrics show how close are predictions to actual realised data and each offers specific advantages and limitations. Generally, a lower metric value suggests that predictions are closer to reality. For all metric that follow, it stands that:

y_i is the actual value of the response variable

$\hat{y}_i$ is the predicted value of the response variable

N is the total number of observations in the dataset

m is the number of forecasts evaluated in the dataset

2.8.1 Mean Absolute Error

Measures the average difference between predicted and actual values ignoring if these are positive or negative. It measures in the same units as the data and is generally less sensitive to outliers than other metrics, but it does not penalise large errors as much (NeuralNine, 2024).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{2.6}$$

2.8.2 Mean Absolute Percentage Error

Measures the average size of errors in absolute terms. Very useful when comparing forecasts between difference dataset as it is scale independent and less sensitive to outliers. The disadvantages are that if the actual values are close to zero the metric becomes indefinitely high while it tends to underestimate rare but large errors (The Friendly Statistician, 2025).

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{2.7}$$

2.8.3 Root Mean Square Error

Emphasizes large errors due to squaring, it is easily interpretable as it has the same scale as the response variable, but it is highly sensitive to outliers (NeuralNine, 2024).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{2.8}$$

2.8.4 Mean Absolute Scaled Error

A really interesting approach, conceptually different than the aforementioned forecast accuracy metrics, is proposed by Hyndman *et al.* (2006). The question that this metric answers is "how does the involved model compared to the naïve forecast model". Here, the naïve model assumes that the next observation is equal to the most recently observed value ($\hat{y}_i = y_{i-1}$). In principle, MASE uses the naïve forecast as a baseline and scales other forecasting models against it as illustrated in the expressions below:

$$MASE = \frac{MAE_{Model}}{MAE_{naive\ forecast}} = \frac{\frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i|}{\frac{1}{N-1}\sum_{i=2}^{N}|y_i - y_{i-1}|} \tag{2.9}$$

In practice, the result can be interpreted as:

- $MASE < 1$: the forecast model performs better than the naïve baseline
- $MASE > 1$: the forecast model performs worse than the naïve baseline
- $MASE = 1$: the forecast model performs equally as the naïve baseline

The benefits of this metric are summarised in that it is unit independent (similar to MAPE), easy to interpret, does not breakdown when actual values of the response variable are equal to zero and sets an equal penalty for both positive or negative as well as large or small errors.

To conclude this section, conventional measures used in inferential statistics such as the Pearson correlation coefficient (r), coefficient of determination (R^2) etc. are not discussed here. They are considered as fundamental concepts that can be easily recalled in multiple statistic bibliography such as the foundational textbook from Berenson *et al.* (2020).

2.9 Summary and research gap

Having scrutinised classical theorems, revolutionary empirical studies, innovative models and recent computation advances, it is evident that practising stock return forecasting has been a subject of extensive research for the last seventy years.

Starting with modern portfolio theory and by applying the risk-return principle through the differentiation of systematic and unsystematic risk, a breakthrough occurred on how investors perceived capital placement in stocks and how this should be effectively diversified. Using those building blocks, CAPM was brought to the picture providing a simple, yet effective option, for portfolio evaluation that is still pertinent nowadays. Its simplicity allows researchers to establish a baseline expected return against which more advanced models can be evaluated. The validity of this model has been applied by several researchers in the Greek market as well.

Previous studies from contemporary financial researchers have shown that a single factor model can not explain all variation in portfolio returns. The FF3 model was designed to resolve this inconsistency by adding more factors, namely the size and value of firms, in the model and managed to somewhat improve its explanatory power in many markets, including the Greek one.

More recently, many academics turned their focus on data driven time series forecasting models in their pursue of an alternative to traditional forecasting methods. Prophet is considered as a prominent approach that offers new dynamic in financial time series forecasting. Although the model is largely applied on cryptocurrency price forecasting in international and domestic literature, it appears to be able to capture short term price fluctuations consistently and outperform classic time series models under certain conditions. This makes Prophet particularly suitable for our empirical analysis, subject that we accept that every time we try to estimate the future, we might be wrong.

To the best of our knowledge, in spite extensive research performed on portfolio forecasting internationally, Prophet has not been applied in the Greek stock exchange yet. The past two decades, traditional pricing models (the CAPM and FF3 in particular) have been extensively applied in Greek finance academic research. Recently, financial publishes have shown an increased interest in returns forecasting using advance machine learning algorithms in the Athens stock exchange. As such, it becomes relevant to explore whether Prophet workflow can be adapted in the process of portfolio returns forecasting, if it can outperform traditional models (CAPM, FF3) and whether factor data calculated from these models could be efficient regressors in Prophet thus improving its predictive accuracy.

Another novelty of our study is the use of MASE as a metric used in the evaluation phase of the models. Even though the metric is generally known, it has received little empirical follow-up. This has significant gravity as the naïve forecast serves as a threshold that all prediction models should be able to outperform for these to be considered successful, yet analogous empirical studies make no attempt in using it.

In the chapter that follows we present the research context and settings as well as the data modelling process employed for our empirical study, detailing the data selection approach, model specification, evaluation techniques as well as the setbacks encountered throughout this workflow.

3 Data modelling and methodology

3.1 Research design overview

Our research study begins by defining a reference period testing that is used to construct a single portfolio from equities of the Athens stock exchange. Stocks of large, medium and small capitalisation were carefully screened based on specific attributes (volume traded, sector, capitalisation) as detailed in the relevant sections hereunder. From the remaining pool of stocks, six of these were selected to form the optimal portfolio on the grounds of the lowest correlation existence between them in an attempt to maximise the diversification benefits in line with the Modern Portfolio Theory. Next, the optimal portfolio weights were determined, at the point where the Sharpe ratio reaches its maximum value on the efficient frontier, aiming to capture the highest risk adjusted returns.

Using the adjusted close price of each of the six stocks multiplied by the optimal weights, simple returns were calculated and aggregated per month. This resulting dataset was then processed according to the explicit requirements of each modelling approach, that is the CAPM, FF3 and Prophet.

Once the models were built, they were used to generate short-term future return predictions. Each prediction was evaluated against the realised performance of the portfolio. Parameter tuning and optimisation was then applied and models were re-assessed.

3.1.1 Software and tools

Data collection, processing, modelling, future returns forecasting and assessment, were conducted in programming language *Python* (version 3.11). The following key libraries were used:

- *yfinance* for historic daily stock price data download in a usable format
- *Pandas* and *NumPy* for efficient data wrangling and tabular numeric manipulations
- *Matplotlib* and *Seaborn* were employed for data visualization and charting
- Prophet for modelling and time series forecasting of portfolio returns and
- regression analysis package (*statsmodels*) for setting-up the CAPM and FF3 models.

The benefits for having selected *Python* are multiple. The syntax is user-friendly and simple, while it became increasingly popular among data scientists and researchers due to its versatile tools for data management tools. The open-source nature of Python provides access to various modules and packages capable of performing advanced computations quickly and effectively. These features make *Python* particularly suitable for the quantitative analysis demands contained in our study.

All *Python* code of the current study is written using *Jupyter* notebooks hosted in Visual Studio Code. This set-up was selected in an attempt to combine executable code with dataframes, statistic analysis and charting within the same file. In addition, Visual Studio Code served as the primary Integrated Development Environment that provided all those necessary tools for the project organisation, debugging, testing and code assistance for *Python*.

3.1.2 Data variables

The current study adopts daily adjusted closing price for stocks as the target variable. It is one of the most common variables used to establish portfolio returns in literature. The benefit of this approach is that any dividends paid or stock splits have been integrated in this variable thereby it captures the total return realised to an investor.

These daily prices are then aggregated into simple arithmetic monthly returns. The use of simple arithmetic monthly returns was preferred over logarithmic returns (very common specially for machine learning algorithms) as:

- our objective to use the optimised weighted average monthly returns for the stock portfolio is straightforward when using simple returns and has no application with log returns
- they are more intuitive to analyse and interpret.

3.2 Portfolio formation and optimisation

3.2.1 Main data sources and data management

The primary data source for our study, as for many other researchers like Ioannidis (2024), is *Yahoo Finance* (https://finance.yahoo.com). This platform provides open-access historic data for the Athens Stock Exchange. Financial data acquisition, specifically daily adjusted close prices and volume traded, was automated using *yfinance Python* library (https://pypi.org/project/yfinance) that ensured a reproducible download of the required data in a specified tabulated format. As the Greek stock price data retrieved from *Yahoo Finance* are denominated in EUR, we had to convert these into USD using the daily exchange rate EUR to USD. This ascertains that the dataset is homogenous as all variables in the regression analysis have to be at the same currency to produce meaningful results.

The risk-free rate is taken from the three-month Greek T-Bill rate retrieved from the Greek Public Debt Management Agency (https://www.pdma.gr). Accordingly, the proxy used for the market returns was the Athens General (composite) Index as in most studies involving the Greek market.

FF3 model factor inputs, namely SMB and HML, specific to the Greek market have been sourced from the Global Factor Data repository (https://jkpfactors.com). These factors are already constructed in USD currency so there is no requirement for any further processing.

3.2.2 Data timeline and frequency selection

A drawback with older observations is that they may lack relevance to today's macroeconomic circumstances and realities. So, the reasoning behind the selection of the sample period in this study was guided by the fact that data from recent years reflect active market trends which can be better understood and is the environment in which corporations operate.

Considering that the course of the Athens General Stock Index acts as a proxy of the actual condition of the economy and public sentiment, it was decided to plot the closing price over the last 11 years (Figure 6). From this chart it is noticed that an interesting sample period, for training the Prophet model and performing the regression analysis for CAPM and FF3, would be between years from January 2018 to December 2023. The out of sample prediction period is set between January 2024 to December 2024 were both major uptrends and steep downtrends are recorded.

The selected seven year timeframe window encompasses distinct phases of the modern economic environment including:

- Increased market variability affected by the exit from a long-lasting period of great financial distress in Greece
- an unprecedented pandemic period, due to COVID-19, which dismantled the markets globally
- a subsequent recovery period displaying an impressive rebound.

Selecting this seven year interval supports the objectives of this study for a number of reasons. It provides a sufficient number of monthly observations for the execution of the CAPM and FF3 linear regression. It also offers more than four years of data for Prophet to assess whether any type of seasonality occurs.

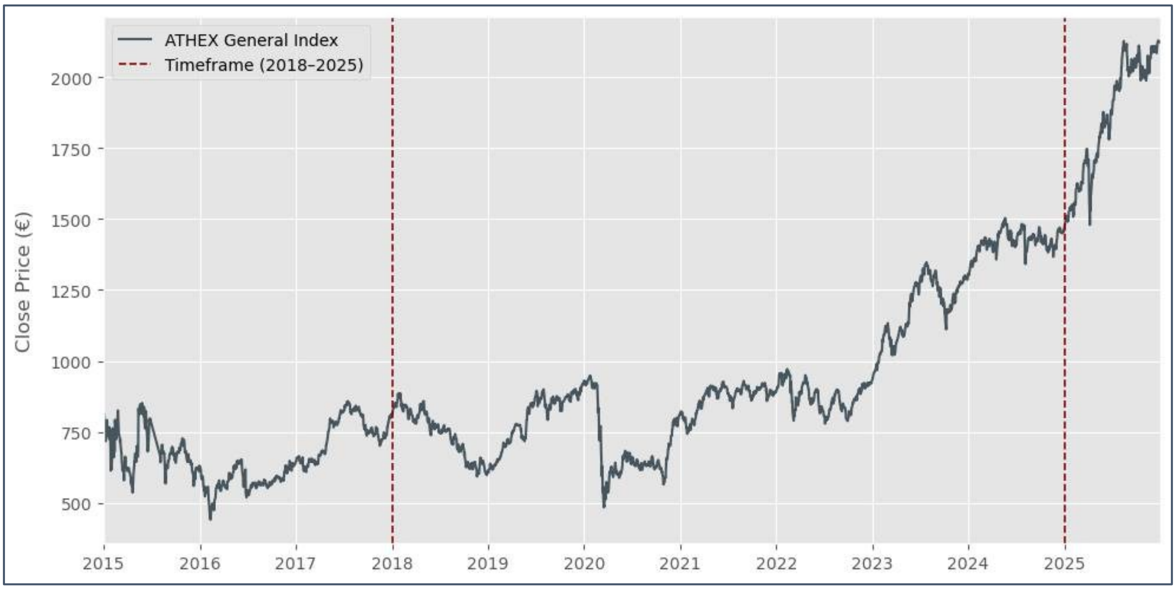


Figure 6: ATHEX Composite Index timeframe from 2015

3.2.3 Stock selection process

At the time this dissertation is drafted, there are 145 stock trading in the Athens stock market according to *Stock Analysis* website (https://stockanalysis.com). Our first step was to align the stocks to our selected timeline based on the volume traded. By eliminating stocks for which no data where available during this period we reached down to 119 stocks. These were ranked based on volume traded from the highest to the lowest. In fact, our algorithm removes any day where more than 90% of the stocks had zero volume so that the dataset is perfectly aligned across 1732 trading days. From this dataset, 26 stocks were shortlisted. Criteria for selecting this subset were as follows:

- In search of an active trendline in stock returns, trading volume was treated as a key variable for our approach. The last 19 firms were filtered out from the original cohort of 119.
- Variability in capitalization of firms (Small, Mid and Large) since this is a prerequisite for the modelling of FF3.
- Mixing industrial sectors would hopefully produce a stratification of the sample. The expectation was that firms from various sectors would possibly exhibit negative correlation.
- The last choosing level came down to personal preference and point of view on the quality of the firms. The criteria for this review stage were based on an assessment of their beta coefficient and future prospects which are out context of the present study.

These 26 stocks along with a brief presentation, sector and capitalisation are presented in Appendix A: “Athens Stock Exchange: Eligible Companies”.

As detailed previously from §2.2, the diversification benefit is diminished at a range from 12 to 18 stocks which sets the maximum threshold for mature markets. Similarly, Michailidis (2009, p. 26) identified four stocks as the minimum threshold required for a diversified stock portfolio.

Based on this argumentation, it was decided to further condense our portfolio down to six stocks. In a market like the Greek one, which is mainly driven by few large companies, there seems to be no added value for diversification in holding a large number of stocks.

To determine the six equites that form the optimal portfolio we relied on correlation analysis. From the pool of 26 equities, we evaluated all six stock combinations and selected the one with the lowest average correlation for the period between 1st of January 2018 to 1st of January 2025.

The first step was to compute the daily simple returns for each stock s using the formula below:

$$return_{s,t} = \frac{P_{s,t} - P_{s,t-1}}{P_{s,t-1}} \tag{3.1}$$

where $P_{s,t}$ represents the adjusted close price of stock s on day t which for our examination period yields 1737 daily returns (five more than the volume dataset) per stock. The Pearson correlation coefficient r (Berenson et al., 2020, p. 181) between each pair of stocks X and Y was then calculated as per the formula hereunder:

$$r = \frac{cov(X,Y)}{S_X S_Y} \tag{3.2}$$

where the numerator is the covariance of stocks X and Y with respect to their daily returns and S_X and S_Y are the standard deviations of daily returns for each stock respectively.

The next step was to build a 26 by 26 grid matrix containing all unique pairwise corelations. To find number of pairs in a group of 26 stocks we used the combination without replacement counting formula (Berenson et al., 2020, p. 223) which results in:

$$\binom{26}{2} = \frac{26!}{2!\,(26-2)!} = \frac{26 \times 25}{2} = 325 \text{ corelations} \tag{3.3}$$

The corelation matrix for all 26 stock is documented in Appendix B: “Pearson Correlation Matrix for 26 stocks”

The total number of possible portfolios that can be generated by choosing six stocks from a total of 26 is:

$$\binom{26}{6} = \frac{26!}{2!\,(26-6)!} = 230{,}230 \text{ portfolios} \tag{3.4}$$

Each of these portfolios contains $\binom{6}{2} = 15$ pairwise correlations from which we derive the mean. The portfolio with the minimum average corelation is the following (Figure 7):

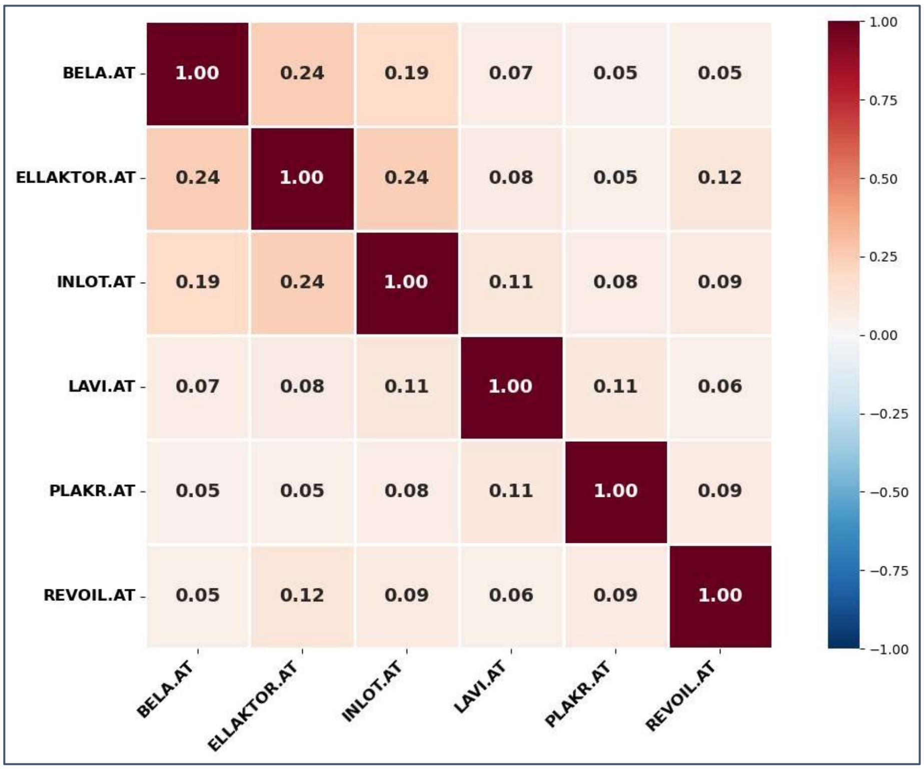


Figure 7: Correlation heatmap for the six stock asset optimized portfolio

The details of the constituent portfolio stocks are described below:

Yahoo Fin. Ticker	Greek Company Name	Sector/Industry	Capitalisation	Brief Description
BELA.AT	Jumbo S.A.	Retail	Large cap	A major retail group specializing in toys, baby products and seasonal items with a strong presence in Southeastern Europe.
ELLAKTOR.AT	Ellaktor S.A.	Construction	Large cap	A major infrastructure group with operations in concessions, real estate and environmental services
INLOT.AT	Intralot S.A. Integrated Lottery Systems and Services	Gaming Technology	Large cap	A global technology supplier and operator in the lottery and gaming sector.
LAVI.AT	Lavipharm S.A	Healthcare	Small cap	Integrated pharmaceutical group specializing in medicine and distribution.
PLAKR.AT	Plastika Kritis S.A.	Plastics Industry	Mid cap	A global manufacturer of plastics and agricultural films with multiple international production sites
REVOIL.AT	Revoil S.A	Oil & Gas	Small cap	Energy company focused on storage and distribution of petroleum products.

Figure 8: Profile of the selected portfolio equities

Using the *describe* method from *pandas* library in *Python*, Table 1 presents the summary descriptive statistics for the monthly returns over the entire dataset i.e. including both the training and testing period. One interesting finding is the highest mean return of LAVI (2.44%) which, however, comes at the highest risk measured by the standard deviation (19.61%). Another important finding is that PLAKR is the least volatile stock offering a mean monthly return of 1% approximately which is the lowest between the rest of the stocks. Another finding that stands out, is BELA's performance which appears to provide consistent returns (1.49%) at very low risk (8.06).

	BELA	ELLAKTOR	INLOT	LAVI	PLAKR	REVOIL
Count	84	84	84	84	84	84
Mean	1.49%	1.65%	1.33%	2.44%	0.97%	1.93%
Standard Deviation	8.06%	13.33%	17.14%	19.61%	6.81%	10.61%
Minimum	-22.59%	-51.40%	-43.05%	-43.30%	-21.20%	-21.01%
Maximum	24.92%	51.36%	62.67%	62.33%	20.00%	37.94%

Table 1: Summary statistics of stock portfolio monthly returns for January 2018 to December 2024

What is worth exploring at this stage are the cumulative returns of each stock over the study period. Rather than plotting monthly returns separately, it was decided to assess total growth by compound each stock returns from start the of 2018 to end of 2024 as illustrated in Figure 9. As a general comment, the chart shows significant variation in stock performance. A closer inspection though, reveals a steep decline for all stocks between years 2020 to 2021 which can be most certainly attributed to the COVID-19 pandemic. Following this shock, most stock exhibit a strong rebound with similar paths, excepting PLAKR which continued to underperform although to a lesser extend. What is also essential to report from this chart is that by the year 2024 nearly all companies have positive returns again though excluding PLAKR. Further analysis leads to the conclusion that despite the fluctuations faced from 2018, this portfolio of stock ends with a positive sign in terms of returns.

Figure 9: Cumulative returns of stock portfolio for January 2018 to December 2024

3.2.4 Portfolio assets weights evaluation

To determine the optimal weight for each constituent stock in our portfolio, the past performance of each stock was analysed within the time frame from January 2018 to December 2024. Annualised expected returns and volatilities were computed from daily adjusted close prices in order to express the performance for a single year. This practice is commonly exercised in portfolio optimisation as demonstrated by Hagmann (2025).

The idea is to assign random weights for each stock and create a series of clone portfolios each having a corresponding Sharpe ratio by using equation (2.2) and assuming an annual risk free rate of 2% as a reasonable (although arbitrary) assumption since it does not affect this part of the analysis. The weight of the portfolio with the max Sharpe ratio shall be selected as it is the one that provides the highest returns at the lowest possible risk.

The constraints that we have introduced in the algorithm are that:

- The weights of the stocks are randomly generated (100 thousand simulations executed)
- Each single weight is between zero and one.
- The portfolio weights remain constant over time i.e. every day we would need to sell winning stocks and buy losing ones without cost and taxes. Although it is understood that this is not a realistic approach it is imperative since our focus is essentially on forecasting returns and not in portfolio management.

Figure 10 plots 100,000 dots representing the Sharpe ratio of a random weight portfolio. The red diamond denotes the portfolio with the maximum Sharpe ratio meaning the best annualised percentage return on y-axis per unit of annualised risk on x-axis.

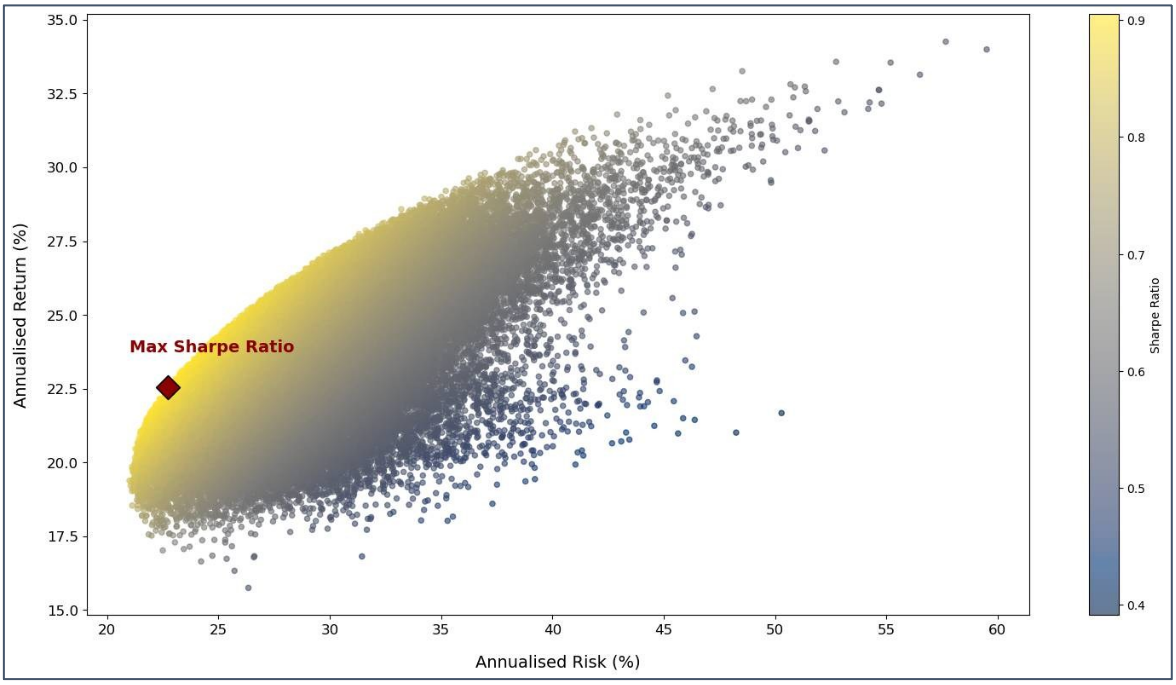


Figure 10: Random weight portfolio return vs risk (100K simulations)

The weights per stock that correspond to the maximum Sharpe ratio above is illustrated below (Figure 11).

Yahoo Fin. Ticker	Greek Company Name	Sector/Industry	Capitalisation	Weights
BELA.AT	Jumbo S.A.	Retail	Large cap	29.62%
ELLAKTOR.AT	Ellaktor S.A.	Construction	Large cap	10.64%
INLOT.AT	Intralot S.A. Integrated Lottery Systems and Services	Gaming Technology	Large cap	2.38%
LAVI.AT	Lavipharm S.A	Healthcare	Small cap	10.72%
PLAKR.AT	Plastika Kritis S.A.	Plastics Industry	Mid cap	21.36%
REVOIL.AT	Revoil S.A	Oil & Gas	Small cap	25.28%
TOTAL				100%

Figure 11: Optimal weights per stock in the optimal portfolio

3.3 Data preprocessing and return calculation

Using the *Python* library *yfinance*, daily adjusted close prices were retrieved for the six constituent stocks (refer to Figure 8) of our optimal portfolio for the period January 2018 to December 2024 inclusive. These daily prices were then converted from EUR to USD by multiplying with the daily exchange rate that was downloaded from the same *Python* library. This choice was dictated by the fact that the available Fama-French factors are determined in USD terms and our decision to express the dataset in a specific base currency.

Daily simple returns, for each stock, were then computed for the entire portfolio using formula (3.1) and any missing values were removed from the dataset. Each of these returns was then multiplied to its corresponding weight (indicated on Figure 11) under the admission that at each trading day the weights are rebalanced. The daily returns were then aggregated (compounded) to monthly returns by multiplying all daily return factors and subtracting from one as per the following rule:

$$Monthly\ returns = (1 + r_{day\ 1}) \cdot (1 + r_{day\ 2}) \dots (1 + r_{day\ n}) - 1 \tag{3.5}$$

The following chart (Figure 12) visualises the trend of the monthly returns of our portfolio in the time frame set initially. It includes both in-sample (January 2018 to December 2023) and out of sample (January 2024 to December 2024 inclusive) data.

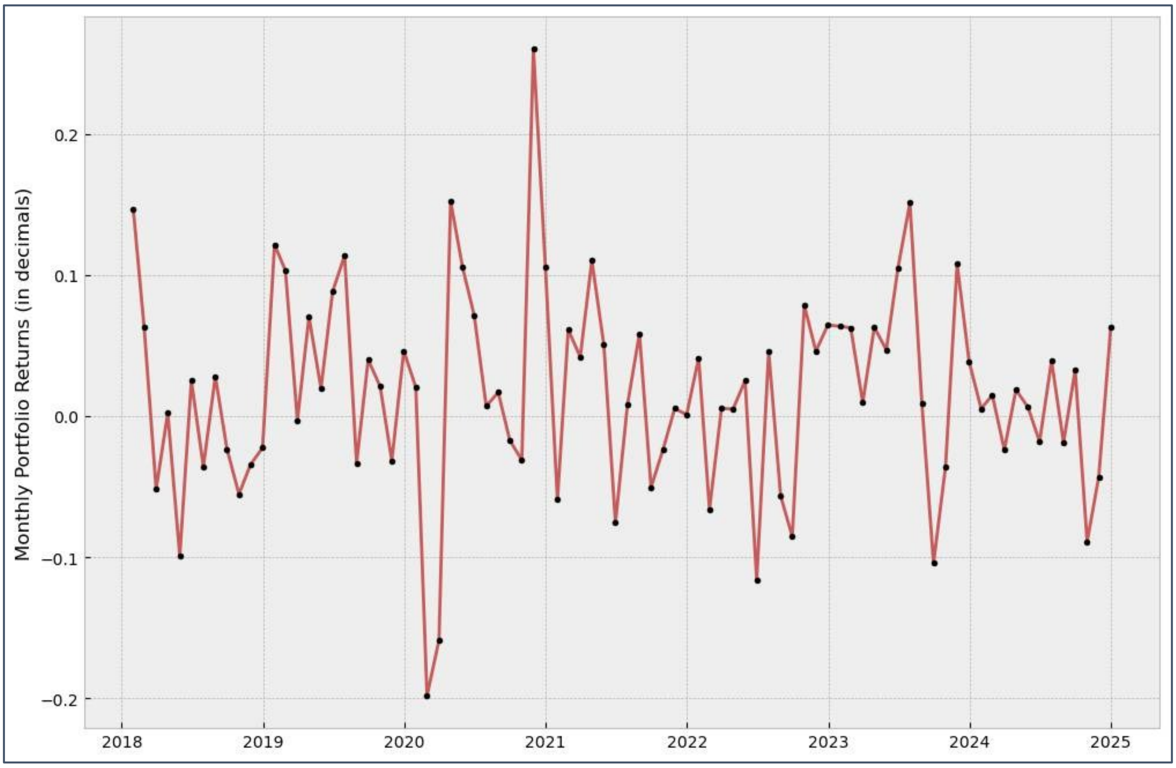


Figure 12: Actual Monthly Portfolio Returns for the entire sample period

3.4 Model development, assessment and refinement

Once our portfolio returns were established our focus then shifted to the development and optimisation (where feasible) of the forecasting models included in this thesis.

3.4.1 Prophet model

The objective at this stage is to use Prophet to build a model that would forecast the monthly return of our optimised portfolio of six stocks. The implementation algorithm was designed as per Figure 13:

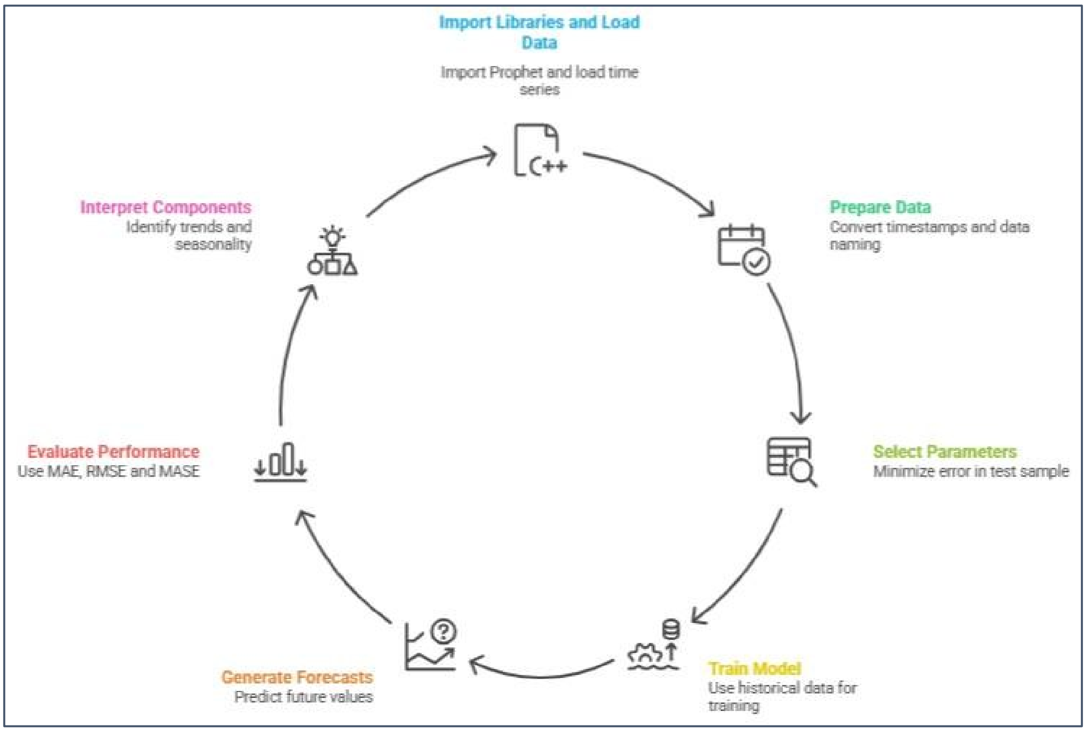


Figure 13: Workflow for Prophet modelling

From §3.3 we have saved our monthly returns data of our optimal six stock portfolio. These data were then imported in *Python*. The input to Prophet is always a dataframe with two columns, namely "ds" and "y". In our case the dates refer to the end of each month and the "y" column refers to the monthly returns which is our target variable.

A critical step in the workflow was the calibration of Prophet's parameters. As Prophet has several controlling parameters, it was decided to concentrate on the ones affecting the trend flexibility and the seasonality strength. The holiday layer was not included in the model as the monthly returns could not be aligned with the daily frequency of holidays. Following the suggestions of Taylor and Letham (2017), a grid search was executed using a set of values from seasonal and trend parameters. Specifically, a series of models were created from a combination of the following parameters:

- *seasonality_prior_scale*: It controls the strength of the seasonality. The higher the value, the stronger the seasonality component hence the faster the model can capture larger seasonal fluctuations. The grid used was [0.1, 0,5, 1.0, 2.0, 5.0, 10.0, 20].
- *changepoint_prior_scale*: Denotes the points (cut-off) at which the trend changes. As the value increases it allows the model to detect and adapt to changes more easily but also increases the risk of overfitting. List of values was [0.001, 0.005, 0.01, 0.05, 0.1, 0.5].

For each combination of parameters, a single Prophet model was initialised in which:

- The model was trained over a four-year period starting from January 2018 and ending early January 2022 (1460 days from start).
- The model then forecasted monthly portfolio returns for the next 12 months. These forecasts were compared with the true monthly returns of the portfolio and an RMSE and MAE was calculated per month. These errors were averaged so that a single RMSE and MAE was stored for those 12 months horizon, corresponding to that specific train/test split. Note that the MAPE was excluded from this evaluation process as monthly returns could be close to zero where this metric becomes undefined providing misleading output.
- The model then "rolled forward" by adding the next month to the training set (i.e. January 2022) and forecasts the next 12-month horizon as described above. This process is repeated until the end of the dataset in December 2023. Refer to Appendix C: "Rolling Window Cross Validation Chart" for a visual interpretation of the process.
- For each parameter subset, a total of 13 split periods were created resulting in 13 RMSE and 13 MAE scores. For each metric, their scores were then averaged to a single final performance measure covering all the split periods. The parameter settings that produced the lowest average RMSE and MAE were selected for the optimal model.

The average errors and best settings for the optimal Prophet model are presented in Table 2. Closer inspection of the table shows that the trend component is relatively rigid to prevent overfitting and the seasonal effect does not react strongly to occasional seasonal spikes.

RMSE	MAE	seasonality_mode	seasonality_prior_scale	changepoint_prior_scale
6,76%	5.50%	Additive	0.11	0.1

Table 2: Metrics results and parameter values of the optimal model

The next step forward was to fit the model with the parameter settings of Table 2. Model training was performed using only the in-sample data from January 2018 to end of December 2023 (sample size $N = 72$ monthly returns). The fitted model was then used to generate monthly returns over a horizon of one year. Specifically, the prediction covered a period from January 2024 to December 2024 without using any rolling window updates.

As can be seen from the following forecast chart (Figure 14), the actual monthly portfolio returns are plotted as black points joined by a red line, Prophet predictions as a blue continuous line while the shaded blue area denotes the 95% confidence interval.

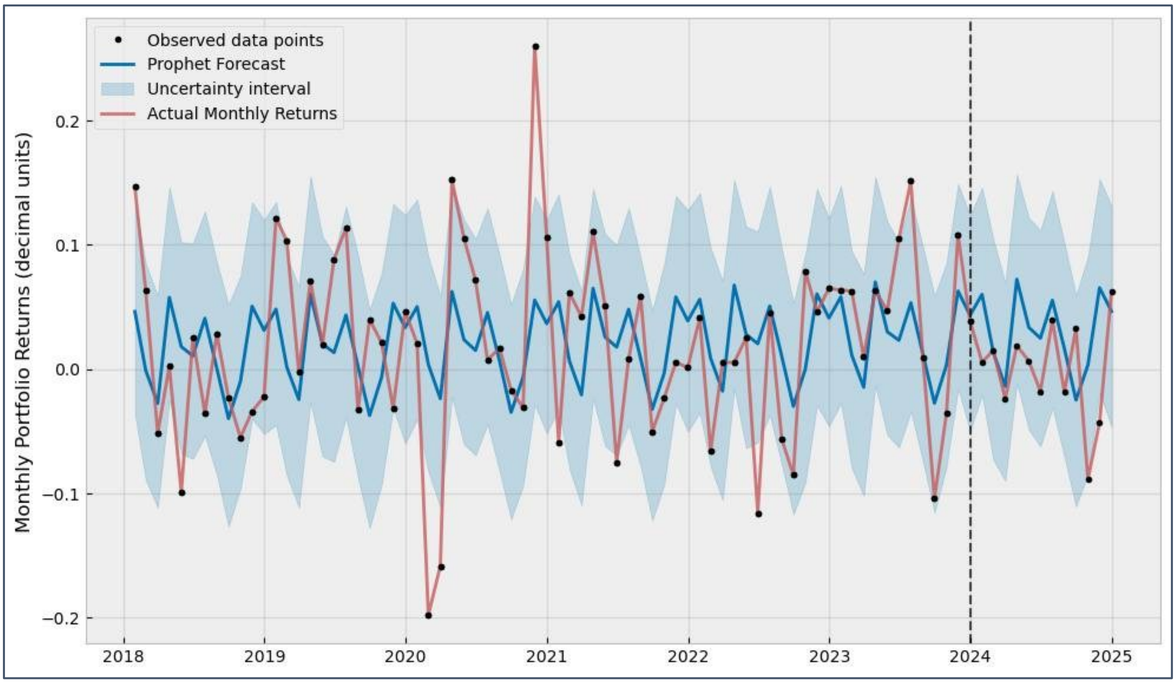


Figure 14: Prophet and actual monthly returns forecast for the entire sample period

Before expanding further on the analysis of this outcome, we shall first discuss the CAPM methodology set up for this study.

3.4.2 CAPM

Having concluded the modelling of Prophet, let us now turn to defining the CAPM. Once again, we started out with the optimal portfolio monthly returns dataset from §3.3 and split the time series to the training period of the model (spanning from January 2018 up to December 2023 inclusive) and the test period (spanning from January 2024 to December 2024).

As discussed from §3.2.1, the R_f was selected as per the three-month Greek T-Bill yield in EUR terms admitting that a U.S investor can have the exact same rate in USD terms. As this rate is quoted in an annualised basis, for instance a quote of 2% refers to the return over a single year if the rate remains constant, it had to be downsized to a monthly level (de-annualised). This was done by simple approximation $R_f = \frac{yield}{12}$ as the annual yield of our sample period is generally very low (not greater than 5%). The challenge was that the T-bill rates occur only on specific auction dates which are irregular and did not match with the dates of our portfolio excess returns series. To address this issue, the last observed auction rate within each a month was assigned to that month end. For months without an auction, the latest available rate was copied forward until updated with a future one. This setting produced a complete risk-free rate dataset that could be properly merged (aligned) in the portfolio monthly returns dataset.

For the market return proxy, the Athens Stock Exchange General (Composite) Index was selected. Following the same principle as with the data preprocessing in §3.3, daily closing prices of the index (ticker: GD.AT) were downloaded using *yfinance*. These daily prices were then converted from EUR to USD by multiplying with the daily exchange rate and aggregated to monthly returns through compounding. This dataset was then aligned by date and merged with the monthly portfolio returns and risk-free rate dataset series for the execution of the regression analysis to follow.

On the basis of equation (2.3), the dataset was rearranged so that the dependent variable is excess portfolio returns $\left(E(R_i) - R_f\right)$ and the independent variable is the market risk premium $\left(E(R_m) - R_f\right)$. A preliminary analysis using a scatterplot (Figure 15) illustrates a positive linear relationship between the variables.

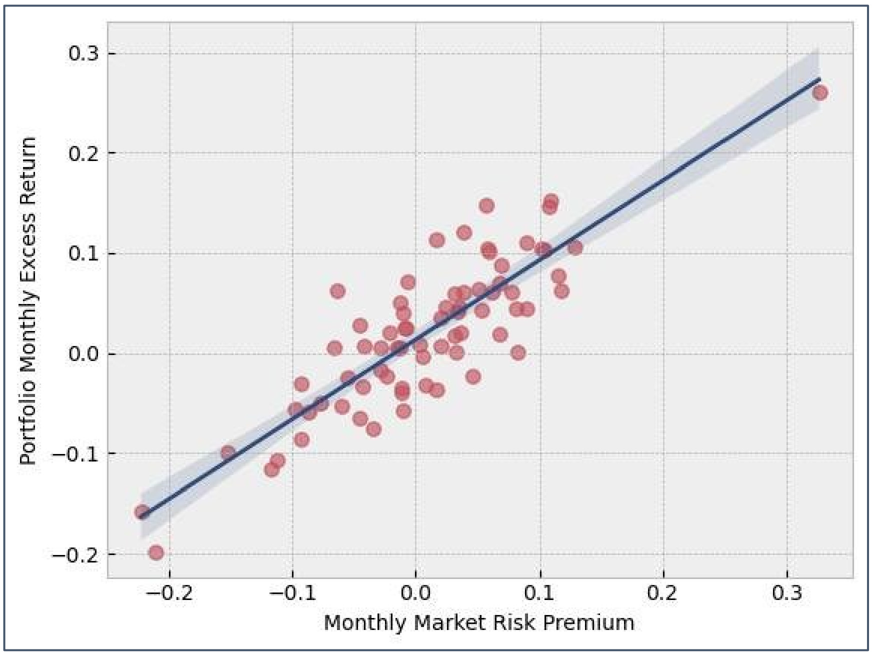


Figure 15: Portfolio Excess Return vs Market Risk Premium (in decimals)

CAPM was then estimated on the training dataset (72 monthly observations) using Ordinary Least Squares (OLS) simple linear regression through the *statsmodels* library in *Python*. By running the regression we estimated the beta coefficient that measures the portfolio's sensitivity to market risk premium movements and the intercept (alpha) that expresses abnormal returns not explained by the market. Regression residuals were evaluated through Durbin-Watson statistics while the data distribution was checked through skewness and kurtosis. The regression results are presented and discussed in §4.1

After the CAPM parameters were evaluated, the model had to be applied to the testing period. The method adopted was to use a lagged value of the market risk premium so that the excess portfolio return in month m is predicted using the information available at the end of month $m - 1$. This singe shift forward of the market risk premium, combined with the slope and intercept estimated from the regression analysis on the training period, were inserted into the CAPM to generate 12 monthly forecasts of portfolio excess returns for January 2024 till December 2024.

3.4.3 Fama-French three factor model

Having discussed the methodology for modelling the CAPM of our optimal portfolio, it is now time to expand it by building the FF3 model.

Like the CAPM, the Athens General Index, converted to USD currency, served as the market proxy while the three-month Greek Treasury Bills decomposed in monthly frequency do for the risk-free rate.

As was mentioned in §3.2.1, the SMB and HML factors for Greece were downloaded from the *jkpfactors* library. The factors frequency is monthly and the currency in USD. Capped value weighting was used, a process which prevents a firm with an extremely high capitalisation to dictate the factor return. In consequence, imposing a weight limit (cap) on each stock allows the overall market to form the factors instead.

Similarly to the CAPM, the ordinary least squares regression was utilised for the period spanning from January 2018 to December 2023 inclusive. The dependent variable was once again expected portfolio monthly returns and the independent variables where the Market risk premium, the SMB and the HML.

The FF3 model's predictive accuracy was then tested over January 2024 to December 2024. A similar approach with the CAPM was also adapted. A monthly lagged value of the market risk premium, the SMB and HML was followed so that the excess portfolio return in month m is predicted using the information available at the end of month $m-1$. A summary of the main findings shall be provided in the chapter that follows.

3.4.4 Comparative model evaluation using relevant metrics (MASE)

Prior to analysing our findings, a description of the process by which the MASE metric was evaluated is necessary. Starting from the Prophet model and by using equation (2.9) the process was carried out in the sequence provided below:

1. Forecast each period i.e. 12 monthly portfolio returns ($\hat{y}_{test}$) from January 2024 to December 2024
2. Compute the absolute forecast error for each of the 12 months in the testing period completing the nominator part, $|e_i| = |y_{test} - \hat{y}_{test}|$.
3. Obtain the naïve forecast baseline for the training period time series as suggested by Hyndman *et al* (2006, p. 684), $MAE_{naive} = \frac{1}{72-1}\sum_{2}^{72}|y_{train} - y_{train-1}| = 0.07514$
4. Monthly MASE for the testing period is then $MASE_i = \frac{|e_i|}{0.07514}$

5. The mean MASE across the 12 test months is $MASE = \frac{1}{12}\sum_{1}^{12} MASE_i$

Figure 16 summarises the data from the previous calculations on the MASE:

Month	y_{test}	y_hat_{test}	$\lvert e_i \rvert$	MAE_{naive}	$MASE_{monthly}$	MASE
Jan	0.0054	0.5860	0.0532	0.0751	0.7084	0.5617
Feb	0.0149	0.0123	0.0026	0.0751	0.0348	
Mar	-0.0238	-0.0153	0.0085	0.0751	0.1137	
Apr	0.0187	0.0708	0.0521	0.0751	0.6935	
May	0.0065	0.0321	0.0257	0.0751	0.3414	
Jun	-0.0180	0.0235	0.0415	0.0751	0.5523	
Jul	0.0392	0.0539	0.0147	0.0751	0.1957	
Aug	-0.0186	0.0152	0.0338	0.0751	0.4500	
Sep	0.0326	-0.0263	0.0589	0.0751	0.7832	
Oct	-0.0889	0.0020	0.0908	0.0751	1.2087	
Nov	-0.0430	0.0639	0.1070	0.0751	1.4233	
Dec	0.0627	0.0450	0.0177	0.0751	0.2354	

Figure 16: MASE calculation for Prophet model

The MASE calculation for the CAPM and FF3 follows the same principles as the one for Prophet with the main difference being that the target variable is excess monthly returns instead of total monthly returns. So, with reference to equation (2.9):

- The numerator is the sum of the absolute errors between the actual and forecasted excess monthly returns of the portfolio for the testing period of 12 months.
- The denominator is the naïve baseline evaluated as the average absolute difference between consecutive monthly returns for the entire training period (72 excess monthly returns)

The output data are summarised in Figure 17. Note that y_{test} is the same for both models.

	CAPM						FF3				
Month	y_{test}	y_hat_test	$\lvert e_i \rvert$	MAE_{naive}	$MASE_{monthly}$	MASE	y_hat_test	$\lvert e_i \rvert$	MAE_{naive}	$MASE_{monthly}$	MASE
Jan 24	0.0022	0.0290	0.0268	0.0751	0.3571	0.5816	0.0517	0.0495	0.0751	0.6593	0.55619
Feb 24	0.0117	0.0395	0.0278	0.0751	0.3696		0.0075	0.0041	0.0751	0.0552	
Mar 24	-0.0270	0.0439	0.0709	0.0751	0.9433		0.0422	0.0692	0.0751	0.9208	
Apr 24	0.0156	0.0077	0.0079	0.0751	0.1053		0.0250	0.0094	0.0751	0.1253	
May 24	0.0034	0.0183	0.0149	0.0751	0.1979		0.0081	0.0046	0.0751	0.0616	
Jun 24	-0.0211	0.0107	0.0318	0.0751	0.4227		0.0087	0.0298	0.0751	0.3962	
Jul 24	0.0365	-0.0134	0.0499	0.0751	0.6640		-0.0212	0.0577	0.0751	0.7678	
Aug 24	-0.0213	0.0615	0.0828	0.0751	1.1017		0.0426	0.0639	0.0751	0.8502	
Sept 24	0.0299	0.0049	0.0250	0.0751	0.3334		0.0099	0.0200	0.0751	0.2658	
Oct 24	-0.0911	0.0293	0.1204	0.0751	1.6021		0.0314	0.1224	0.0751	1.6299	
Nov 24	-0.0452	-0.0472	0.0020	0.0751	0.0261		-0.0475	0.0023	0.0751	0.0300	
Dec 24	0.0604	-0.0039	0.0643	0.0751	0.8557		-0.0082	0.0685	0.0751	0.9122	

Figure 17: MASE calculation for CAPM and FF3 models

3.4.5 Model refinement (Prophet model tuning)

Returning to the Prophet model presented in §3.4.1, we have observed how in its basic form it can generate forecasts using historical data from the target variable (monthly returns) through trend and seasonality components. As Taylor and Leatham (2017) explain, Prophet allows the integration of additional variables, called external regressors, in the fitting process of the model to improve its explanatory and predictive ability.

In this study, the regressors feature was simulated by adding the SMB and HML factors to the modelling process. In this way the model accounts for additional characteristics that may affect the portfolio monthly returns. Using this concept useful conclusions on how traditional theory connects with contemporary data-driven algorithms could be derived.

Practically, by using the same code structure and optimal parameter values obtained from the initial Prophet model, we introduced two additional columns from the FF3 model, the SMB and HML. This refined Prophet model was then trained from January 2018 to December 2023. Keeping the same testing period of one year (2024), SMB and HML were shifted by one month ahead to generate the monthly portfolio returns prediction. The forecasting results are presented below in Figure 18.

Month	y_{test}	y_hat_{test}	$\lvert e_i \rvert$	MAE_{naive}	$MASE_{monthly}$	MASE
Jan 24	0.0054	0.0554	0.0500	0.0751	0.6652	0.5434
Feb 24	0.0149	0.0144	0.0005	0.0751	0.0060	
Mar 24	-0.0238	-0.0179	0.0060	0.0751	0.0792	
Apr 24	0.0187	0.0715	0.0528	0.0751	0.7027	
May 24	0.0065	0.0296	0.0232	0.0751	0.3082	
Jun 24	-0.0180	0.0245	0.0425	0.0751	0.5656	
Jul 24	0.0392	0.0511	0.0119	0.0751	0.1581	
Aug 24	-0.0186	0.0143	0.0329	0.0751	0.4381	
Sept 24	0.0326	-0.0271	0.0597	0.0751	0.7946	
Oct 24	-0.0889	-0.0014	0.0875	0.0751	1.1639	
Nov 24	-0.0430	0.0586	0.1016	0.0751	1.3517	
Dec 24	0.0627	0.0411	0.0216	0.0751	0.2878	

Figure 18: MASE calculation for Prophet refined model

4 Data Analysis and Discussion

This chapter presents and discusses the empirical results obtained from the modelling processes described in the previous section of this thesis. Starting with the Prophet models, a cross validation of the prediction accuracy is performed to strengthen our confidence on the forecasts. Following, the regression output for the CAPM and FF3 are examined so as to determine the relationship between portfolio excess returns and the associated risk factors. The implications of these results on our portfolio shall also be reviewed. All models shall be compared to the naïve benchmark to establish whether predictions are meaningful or simply follow a random walk. The chapter concludes with the consideration of the main limitations involved and sensitivity analysis of the research.

4.1 Interpretation of model outputs (patterns, deviations, over/underperformance)

It is interesting to compare Figure 16 and Figure 18 as they show how similar the MASE values for the base and refined Prophet models are. Contrary to initial expectations, their results are almost identical (0.5617 and 0.5434 respectively). Because of this, only a single cross-validation was carried out to check the forecasting errors distribution of the base Prophet model, which should also be pertinent for the refined model.

For this evaluation process, the module library includes functionality for time series cross validation that measures prediction errors from the historical data. Technically, Prophet executes cross validation in a daily frequency. Since the frequency of our dataset is monthly, a conversion into days was necessary. The cross-validation is by default limited on the training dataset (January 2018 to December 2023) to prevent any data leakage of the test data. For this purpose, we need to introduce the following terminology used by Prophet:

$Initial$: This parameter defines the training period for the model. 1460 days were selected as a reasonable amount of historical data used. This setting forces the algorithm to train on the first four years of data (January 2018 through early January 2022) before making its first validation forecast.

$Period$: This is the forward step size i.e. the number of days for which a forecast shall be done. The period is selected to be approximately one month (30 days), so the algorithm shifts its timeline forward by exactly 30 days to run the next simulation.

$Horizon$: This attribute determines how many days in the future we want to forecast for. For every period, the model is programmed to predict the portfolio return for the next 12 months (365 days).

By applying this forward step process, it becomes possible to plot the performance metrics for RMSE and MAE over the 365 days horizon. Note that on the two following plots, the grey points represent the individual forecast errors obtained from each cross validation period.

As discussed in §2.8.3 , the RMSE metric is sensitive to outliers. Figure 19 illustrates how stable the RMSE remains over the entire 12 months horizon, ranging from slightly lower than 6% to nearly 8%. One of the issues that emerge from this observation is that there does not seem to be very strong outliers within this period that can shadow other significant effects on portfolio returns. What is also reassuring is that the model's reliability does not deteriorate in the long term since the RMSE does not increase with time.

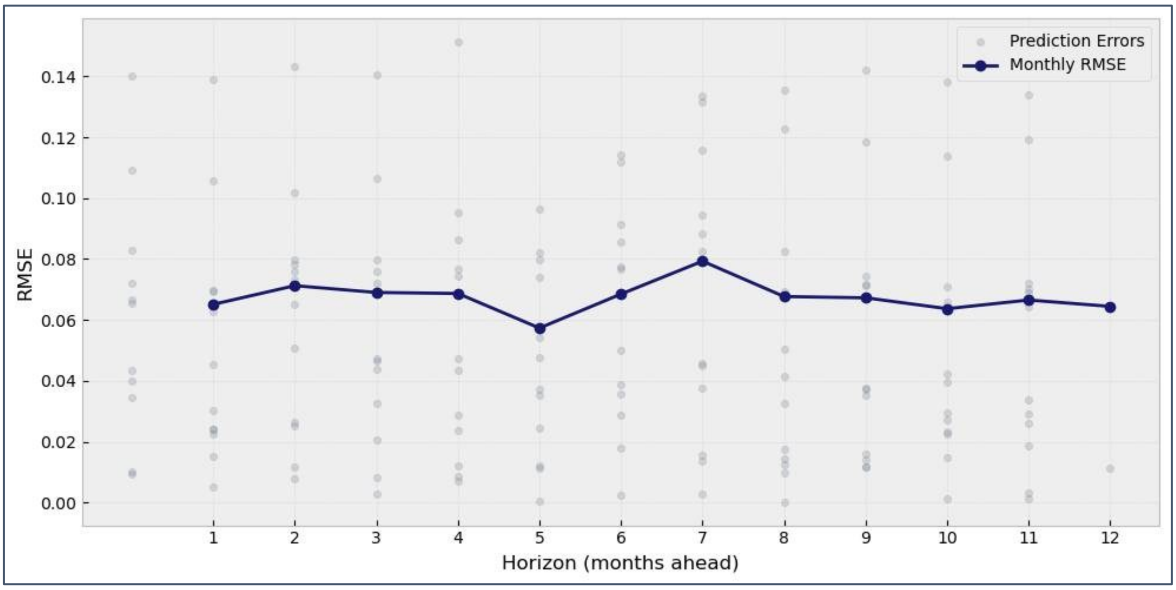


Figure 19: Prophet cross-validation - RMSE over a 12 month forecast horizon

Turning into the next evaluation metric, Figure 20 displays the progression of MAE across the forecast horizon. Similarly to the RMSE, MAE follows the same trajectory, though at a lower range of roughly 4.5% to 6.5%. The lowest error is observed at around five months, corresponding to the horizon midspan, while the highest error appears near seven months. Given that MAE is less sensitive to outliers than RMSE, the observed pattern is suggestive of a relatively uniform error across the 12 month horizon with limited variation in predictive performance.

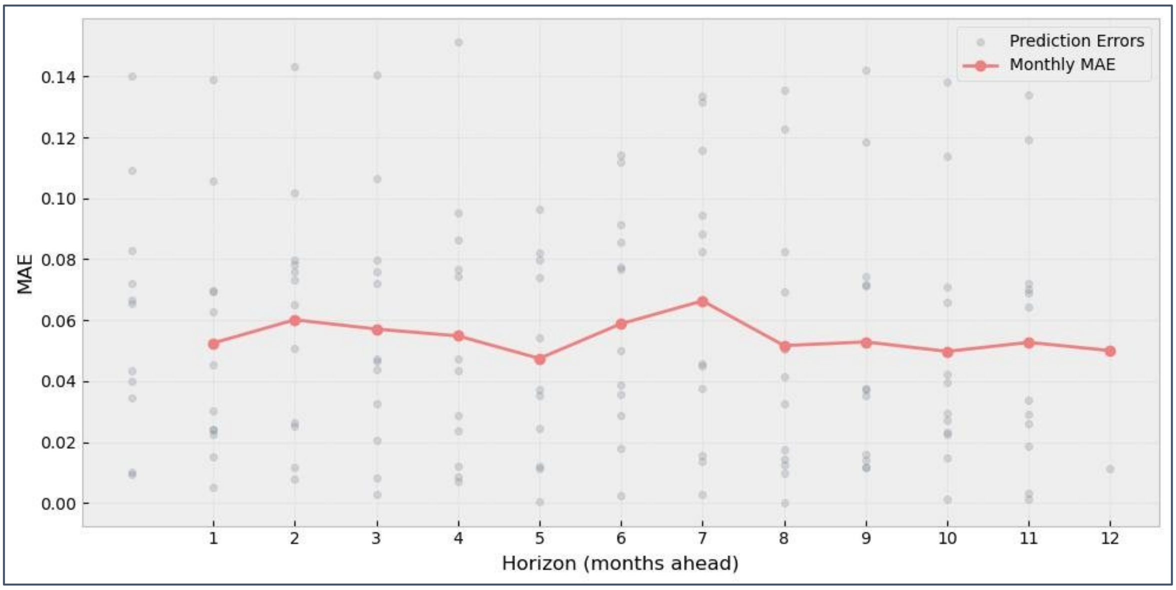


Figure 20: Prophet cross-validation - MAE over a 12 month forecast horizon

In this paragraph, the empirical results of the CAPM regression for the training period are presented. The output is illustrated in Figure 21.

```
==============================================================================
Dep. Variable:          Excess_Return   R-squared:                       0.743
Model:                            OLS   Adj. R-squared:                  0.739
Method:                 Least Squares   F-statistic:                     294.7
Date:                Sun, 08 Mar 2026   Prob (F-statistic):           8.54e-27
Time:                        16:45:58   Log-Likelihood:                 132.64
No. Observations:                  72   AIC:                            -261.3
Df Residuals:                      70   BIC:                            -256.7
Df Model:                           1
Covariance Type:                  HC3
==============================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const             0.0134      0.005      2.908      0.004       0.004       0.023
Market_Premium    0.7938      0.046     17.168      0.000       0.703       0.884
==============================================================================
Omnibus:                        1.706   Durbin-Watson:                   1.996
Prob(Omnibus):                  0.426   Jarque-Bera (JB):                1.465
Skew:                           0.348   Prob(JB):                        0.481
Kurtosis:                       2.943   Cond. No.                         12.2
==============================================================================
```

Figure 21: CAPM regression analysis output

Applying the intercept and the slope to equation (2.3), the CAPM equation is expressed as:

$$E(R_i) - R_f = 0.0134 + 0.7938 \cdot \left[E(R_m) - R_f\right] \tag{4.1}$$

The diagnostics of the main regression metrics resulting from Figure 21 are analysed below:

$R - squared = 0.743$ the market premium explains 74.3% of the variation in the portfolio returns, indicating that the model provides a good fit for our data.

$Intercept\ (a) = 0.0134$ the intercept is statistically significant at the 95% confidence level since the p_{value} (0.004) is lower than 5%. This denotes that the portfolio generates abnormal returns (1.34% on average) that can not be explained by the market risk (beta). This finding contradicts the expectations from the efficient market hypothesis which suggests that no one can generate returns beyond the systematic risk (beta) undertaken. Even if it happens it should be either for a very short time or attributed to chance. Hence this finding should be treated with caution.

$Market\_Premium = 0.7938$ the slope of the regression (beta) is highly significant given the low $p - value$ $(0 < 0.05)$. As the beta is lower than one, we can safely estimate that the portfolio is less volatile than the market.

$Skewness = 0.348$ the mild positive skew is suggestive of a marginally higher probability of extreme positive returns rather than negative ones. This is also indicative of a less risky asset.

$Kurtosis = 2.943$ matches a behaviour similar to the normal distribution at the tails as this value is very close to three.

$Durbin - Watson = 1.996$ as this figure is nearly two, we conclude that there is very little chance of autocorrelation in the residuals. This is fundamental for confirming the reliability of our regression output.

The next section of our empirical analysis focuses in the output of the FF3 regression as outlined in Figure 22:

```
==============================================================================
Dep. Variable:          Excess_Return   R-squared:                       0.800
Model:                            OLS   Adj. R-squared:                  0.791
Method:                 Least Squares   F-statistic:                     128.4
Date:                Sun, 08 Mar 2026   Prob (F-statistic):           5.99e-28
Time:                        17:25:44   Log-Likelihood:                 141.62
No. Observations:                  72   AIC:                            -275.2
Df Residuals:                      68   BIC:                            -266.1
Df Model:                           3
Covariance Type:                  HC3
==============================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const              0.0114      0.004      2.707      0.007       0.003       0.020
Market_Premium     1.0087      0.070     14.489      0.000       0.872       1.145
SMB                0.3921      0.091      4.292      0.000       0.213       0.571
HML               -0.2136      0.074     -2.870      0.004      -0.359      -0.068
==============================================================================
Omnibus:                        3.713   Durbin-Watson:                   2.284
Prob(Omnibus):                  0.156   Jarque-Bera (JB):                2.962
Skew:                           0.473   Prob(JB):                        0.227
Kurtosis:                       3.301   Cond. No.                         28.6
==============================================================================
```

Figure 22: FF3 regression analysis output

Replacing the intercept and independent variables, the FF3 equation is expressed as:

$$E(R_i) - R_f = 0.0114 + 1.0087\left[E(R_m) - R_f\right] + 0.3921 \cdot SMB - 0.2136 \cdot HML \qquad (4.2)$$

Turning to the interpretation of the regression results for FF3 model, it is observed that 80% of the variability in the portfolio excess monthly returns is explained by the totality of the independent variables. This is strong evidence to argue that the model is highly explanatory. Adjust R-squared (79.1%) is closely tracking the R-squared indicating that the additional factors (SMB and HML) add predictive value in the model. This observation is further supported by the large number of the F-statistic (128.4), an indicator that reveals how much the independent variables contribute to the data explanation collectively.

The intercept is significant at the 5% level, in accordance with the CAPM regression results, so it requires cautious treatment. The market premium (1.0087) is significant and slightly higher than the CAPM prediction. Since this coefficient is practically one, the portfolio returns appear to follow the market exactly. The SMB coefficient has a slightly positive value and is statistically significant. This could also imply that our stock portfolio is more invested in small cap stocks. The next factor, HML, is largely negative and statistically significant. It signifies that the portfolio tends to be growth rather than value oriented.

4.2 Forecasting performance comparison between models

In this section we examine how the empirical results of our thesis illuminate the research questions established in §1.2. The first question sought to determine whether traditional models can be used to forecast stock portfolio returns in the Greek market, and if so, how well do they perform.

The top panel of Figure 23 depicts the comparison between actual portfolio excess returns (grey dotted line) and the lagged CAPM forecast for year 2024. Our decision to forecast excess returns using the one month lagged market premium is clearly imprinted on this chart. After the first two months of the forecast, it becomes clear that model’s predictions are strictly guided by the previous month’s market performance. The model fails to capture changepoints in trend on time but rather acts more as a delayed predictor of the returns.

The lower panel of Figure 23, illustrates the monthly MASE over the 2024 testing period. At first sight the model appears to outperform the naïve benchmark since at ten months out of twelve it scores a MASE less than one. A note of caution is due here though as a closer inspection of the data indicates that the model is not actually anticipating future return movement, but it is rather following the market with one month delay. This is especially evident in November where the model portrays a sharp decline in excess returns in contrast with the portfolio’s actual returns which are rebounding from a previous low. The fact that the MASE is so small in this case can be attributed merely to chance. Contrary to our expectations, even though the MASE results appear to be very good, the forecasting performance can’t be considered reliable.

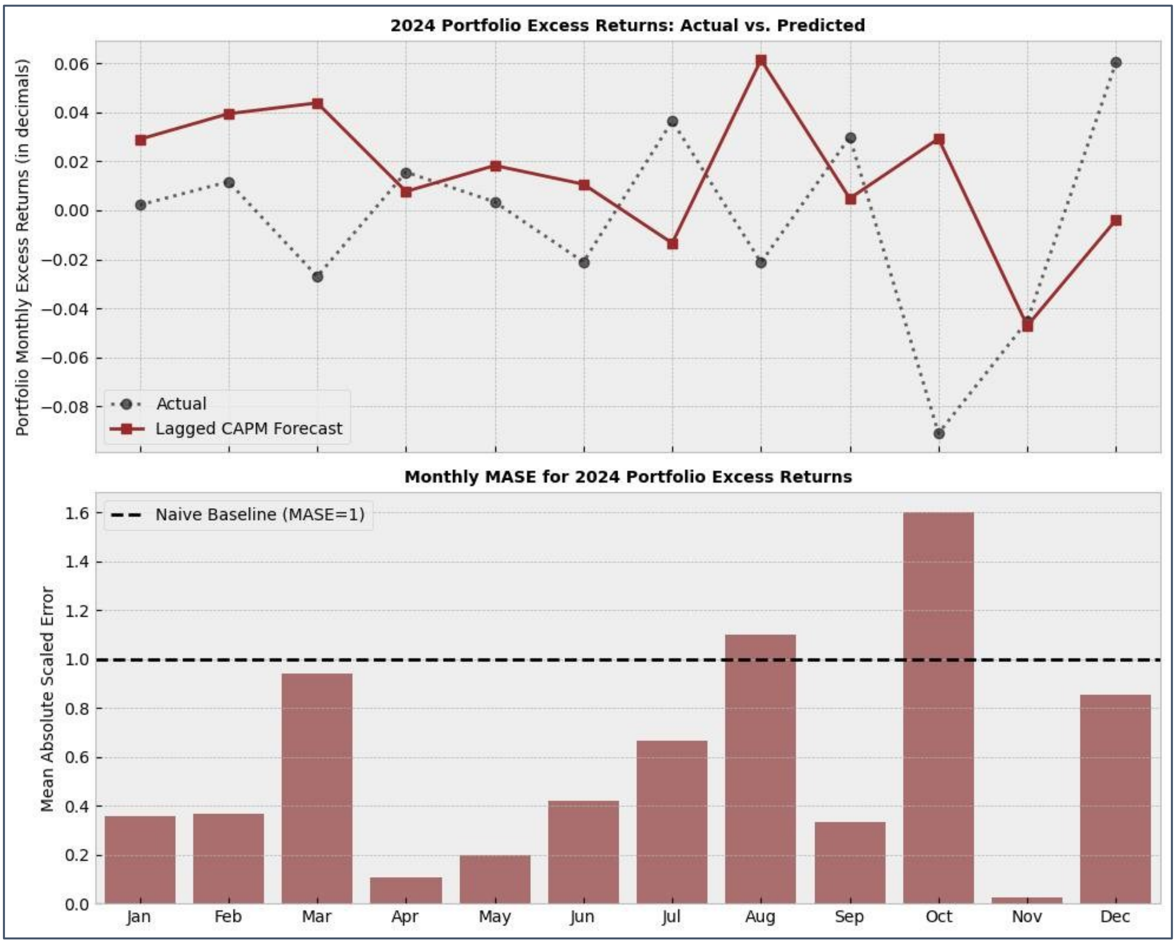


Figure 23: CAPM - Comparative analysis of portfolio excess returns and MASE for 2024

Continuing our review on the traditional asset pricing models, attention now turns on the predictive performance of the FF3 model shown in Figure 24. Comparing the top panel with that of the CAPM it is noticed that the FF3 forecasts start quite differently, for January and February, as they move in opposite direction compared to the realised portfolio returns. This inconsistency is most likely associated to the effect of the data from the period just before the predictions start. Nonetheless, from March onwards, the excess return pattern starts to resemble that of CAPM. This could be a solid support for the earlier hypothesis done hereabove that the models are in fact trailing the true excess returns of the portfolio by one month rather than signalling actual changes in trend.

Studying the lower panel of Figure 24, the pattern of the MASE follows exactly that of the excess returns. Once again, the model fails to overpass the naïve forecast for October where a major outlier is observed in the returns data. Perhaps the most interesting finding that emerges from these results is that adding more risk factors, like size and value, does not resolve the fundamental limitation introduced in our methodology, namely the one month lag in the explanatory variables. Despite this, the FF3 exhibits marginally lower MASE on average compared to CAPM but still can't be considered a sufficiently reliable indicator for tracking the portfolio returns movement.

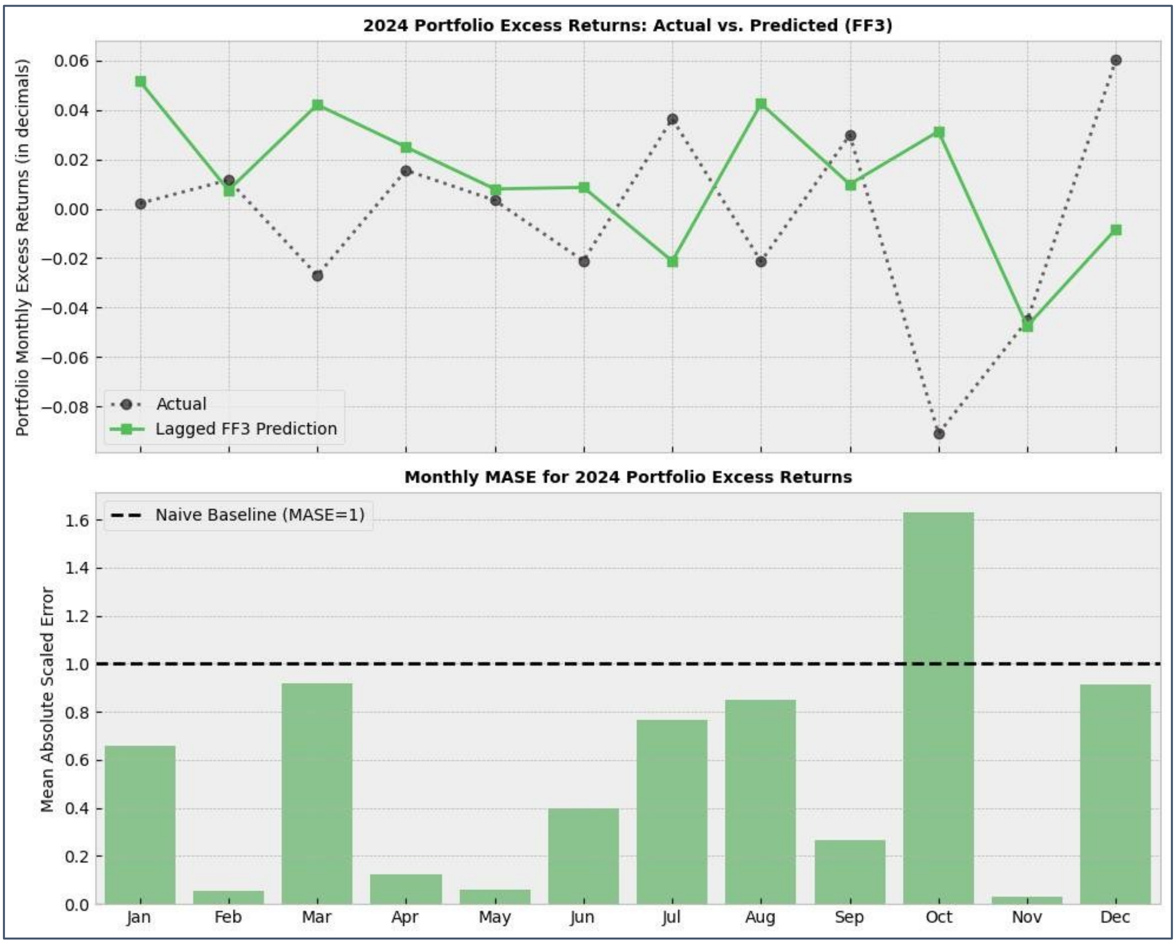


Figure 24: FF3 - Comparative analysis of portfolio excess returns and MASE for 2024

Following the analysis of the traditional regression models, we now examine the performance of the Prophet models. Employing a similar approach as previously, Figure 25 presents how Prophet forecasts progressed from January to December 2024. The forecast performs rather poorly for the first month, both in terms of direction and magnitude of predicted returns. Although the February forecast closely matches the one actually realised, a note of caution is required here. The returns predicted slope is declining whereas the true returns are in fact rising suggesting a high likelihood for this outcome to be coincidental.

From March to August 2024, the model is quite effective in identifying the direction of the portfolio returns albeit at a quite different magnitude. Specifically, it tends to overestimate them particularly on the higher positive end. Even so, the fact that the model manages to synchronise with the actual returns of the portfolio in terms of trend and seasonality is noteworthy.

The analysis reveals important limitations as well. Isolating the October forecast, a sharp decline occurs in reality but, instead, the model is prognosing an optimistic outcome. This is characteristics of the model's inability to account for unexpected market shocks, a result which is somewhat expected, given that the model is tuned to produce smoother, more conservative forecasts.

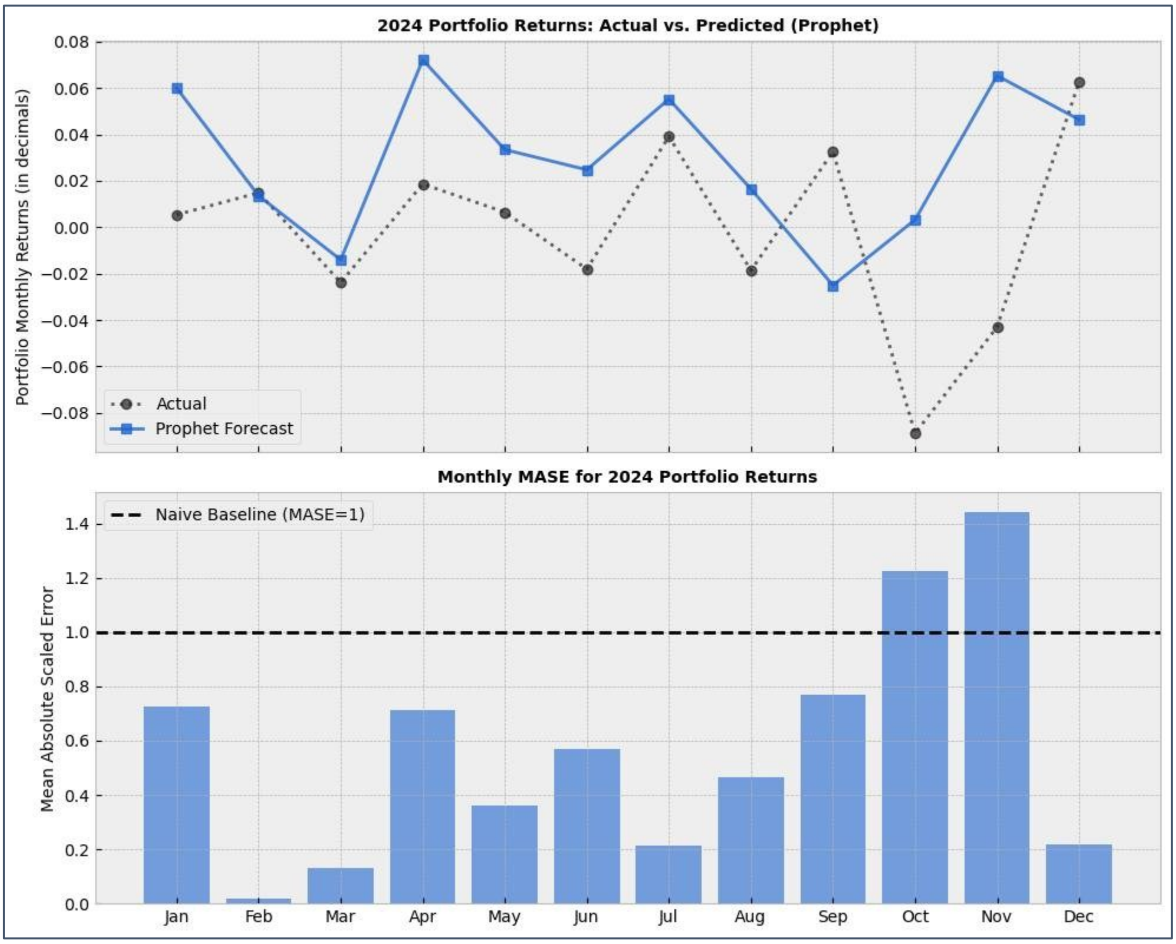


Figure 25: Prophet - Comparative analysis of portfolio returns and MASE for 2024

The lower panel of Figure 25 provides a detailed look of the MASE values relative to the naïve benchmark. The error analysis further supports our findings so far. For most months, the MASE remains well below one indicating that the forecasts are performing better than a random walk. The question remains though for the MASE of February that is exceptionally low and should be viewed carefully if not discarded in real life situations. Also, MASE results confirm that the model is most likely speculating rather than predicting during highly volatile periods as those recorded in October and November.

We now proceed to examine whether incorporating the Fama-French factors into our Prophet model has had a favourable influence in its predictive accuracy. By directly comparing Figure 26 with Figure 25 it is deduced, to our disappointment, that they are virtually identical. Applying the Fama-French factors as regressors appears to have had little to no measurable effect, at least by visual testing both charts, on the refined Prophet model prognosis.

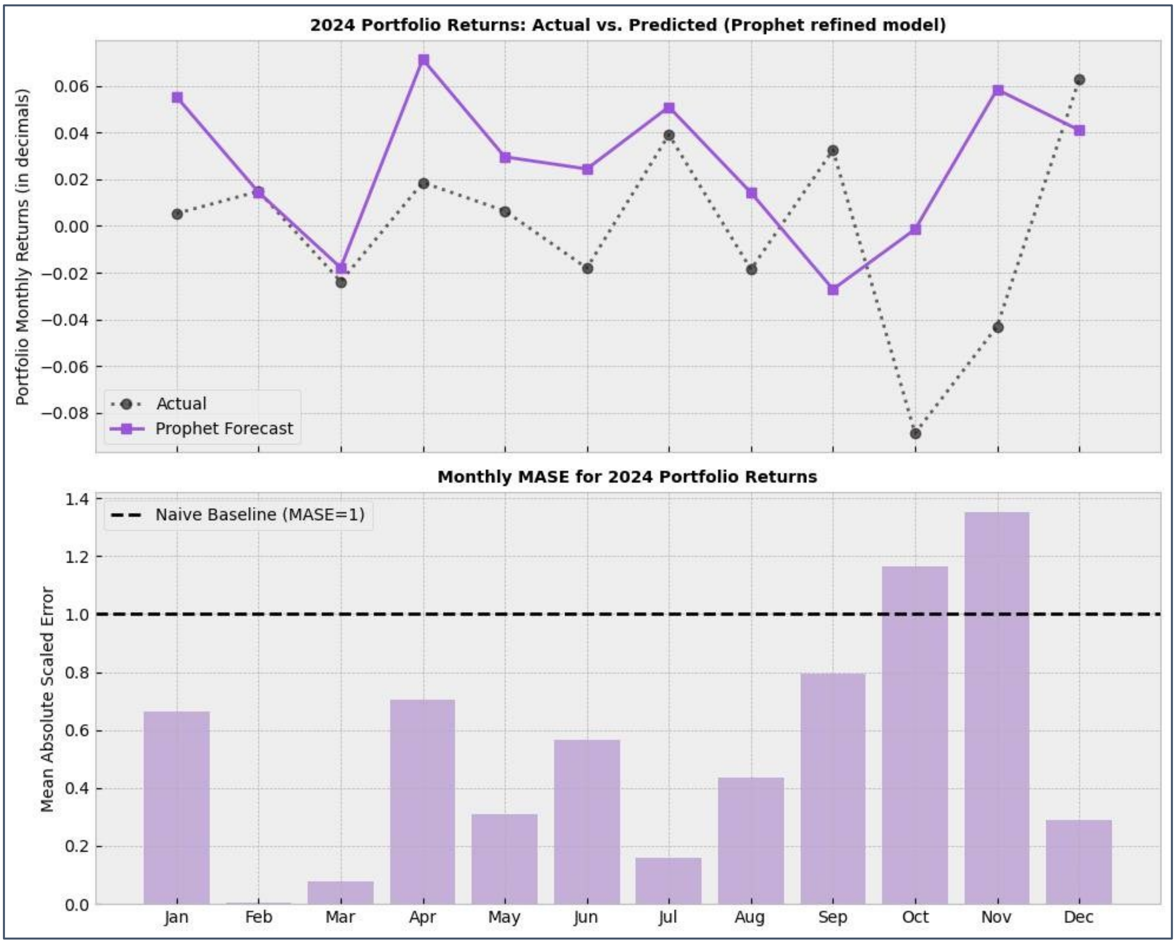


Figure 26: Prophet with regressors - Comparative analysis of portfolio returns and MASE for 2024

Having discussed all the individual models thus far, the following chart (Figure 27) compares the monthly MASE of all four of them against the naïve benchmark. Apparently, most forecasts succeed to surpass the naïve benchmark. What is also interesting from this table is that the Prophet models predictions are remarkably close in every month. If the first two months of forecasting are ignored, CAPM and FF3 also converge significantly particularly in terms of trend but the month to month variation seems greater than the Prophet models.

Of interest here are also the observations that all models fail in October, where the actual returns present a sharp decline, but not in November where an extreme offset is also noticed. This specific forecast should not be discarded from our evaluation as this would introduce bias in our study. Instead, it could be interpreted as evidence that the models either generally fail to account for outliers in the data and when they appear to do so it is only by coincidence.

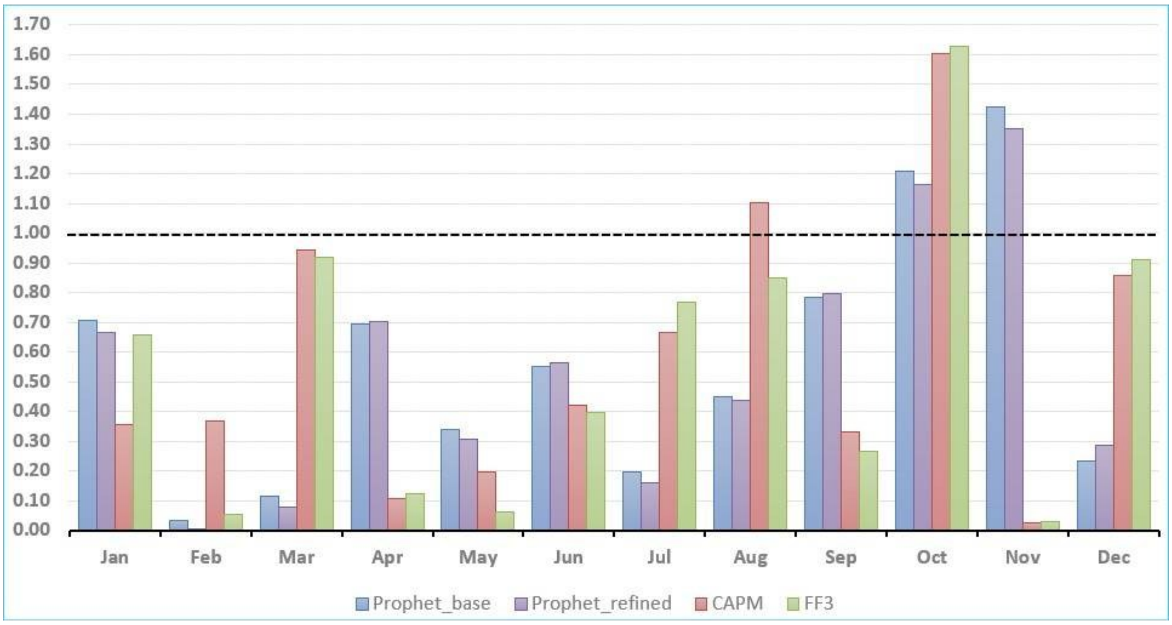


Figure 27: Comparative summary of monthly MASE results for all models – Values in decimals

A comparative evaluation of all four models, based on their average MASE, RMSE and MAE for 2024, is presented in Table 3. Note that the number in the parenthesis refers to the ranking of each model with (1) denoting the lowest error hence the best performing model in the category.

	Prophet_Base	Prophet_Refined	Lagged_CAPM	Lagged_FF3
RMSE	0.0523 (2)	0.0507 (1)	0.0551 (4)	0.0546 (3)
MAE	0.0422 (3)	0.0408 (1)	0.0437 (4)	0.0418 (2)
MASE	0.5617 (3)	0.5434 (1)	0.5816 (4)	0.5562 (2)

Table 3: Summary of evaluation metrics for the test dataset – Values in decimals

From the table above it is evident that there is little difference in the forecasting performance of the models in question. Even so, the refined Prophet model performed best overall, against all other competitors, in all evaluation metrics. Scoring the lowest RMSE highlights its effectiveness in minimising large prediction errors. The improvement from the base Prophet model is closely linked with the addition of FF3 regressors that had a positive, albeit minimal, impact on accuracy. The base Prophet model takes the second place in the RMSE metric that confirms our earlier observation that Prophet models are more efficient in controlling excessively high prediction errors.

By contrast, the FF3 model has overpassed both the base Prophet and CAPM in terms of MAE and MASE. From a careful review of the data, a likely explanation is the poor result of the Prophet models for October taken together with the increased likelihood that the traditional models low forecast error for this month have occurred by chance. As expected, the CAPM demonstrates the weakest performance in all metrics but not substantially worst than the other models. This remark provides enough evidence to argue that the overall differences in forecasting accuracy remain limited.

4.3 Discussion of findings in the context of the Greek market

It is now necessary to closely examine our findings and deduce all valuable insights from our work so far. This shall further assist us in fulfilling the principal aim of our thesis, which is to propose a practical solution for forecasting a portfolio's directional shifts within the Greek stock market.

The first step is to review what useful insights can be drawn in relation to the nature of our portfolio and how can we exploit these to our benefit to predict its future course. Figure 28 plots our portfolio and ATHEX Composite index monthly returns for our entire sample period, beginning of 2018 to end of 2024. The chart data tell us that the portfolio returns fluctuate mainly around 15% and to -10% with few extreme datapoints reaching close to 25% to almost -20% in a monthly level.

The main issue identified from this chart is that our portfolio seems to follow the market quite closely and usually with a slightly lower magnitude. What comes out of this is that it behaves more like a defensive collection of assets rather than one that is highly independent of the market. This interpretation is consistent with the beta coefficients derived from the CAPM (≈ 0.8) and FF3 (≈ 1) models, which is a typical characteristic of assets with low systematic risk. So, despite our efforts to select stocks having the lowest correlation possible, the portfolio did not reach the degree of independence from the market index that we had initially targeted.

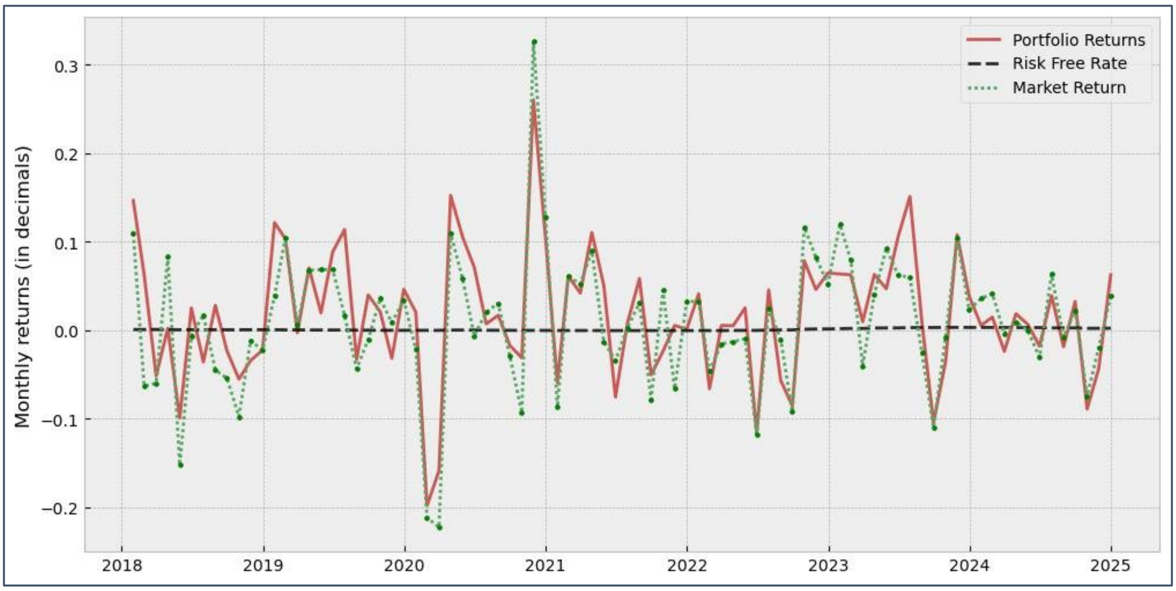


Figure 28: Stock portfolio returns vs ATHEX Composite Index return in month frequency

At this point, we shall further explore the output of the best performing model, namely Prophet refined. Our efforts will focus in extracting essential information specifically on the direction of the portfolio trend. In pursue of this, the actual and forecasted monthly returns along with their corresponding trendlines are plotted for our entire dataset in Figure 29. The trendline of the actual returns is estimated using the OLS method. What is interesting about this chart is that the trendlines have little but distinct difference in slope. The actual returns trendline is flat to slightly negative in contrast with that of Prophet which has a more clearly upward slope.

It is apparent that Prophet is somewhat more optimistic in its returns expectations than what is taking place in reality. However, it is essential to acknowledge that Prophet's trend component is defined solely within the training dataset. The model studies the training data (spanning from 2018 to 2023) and extracts the underlying smooth directional pattern from that period. This is then projected forward in time, in our case one year ahead. There could be several possible explanations to explain this result. It seems, though, more possible that it is related to the prolonged training data period of six years, which in fact could have distorted the slope calculated by Prophet by including older observations. Perhaps a considerable improvement would have been to narrow the training period, for example by excluding the first two or three years, and then compare the new trendlines.

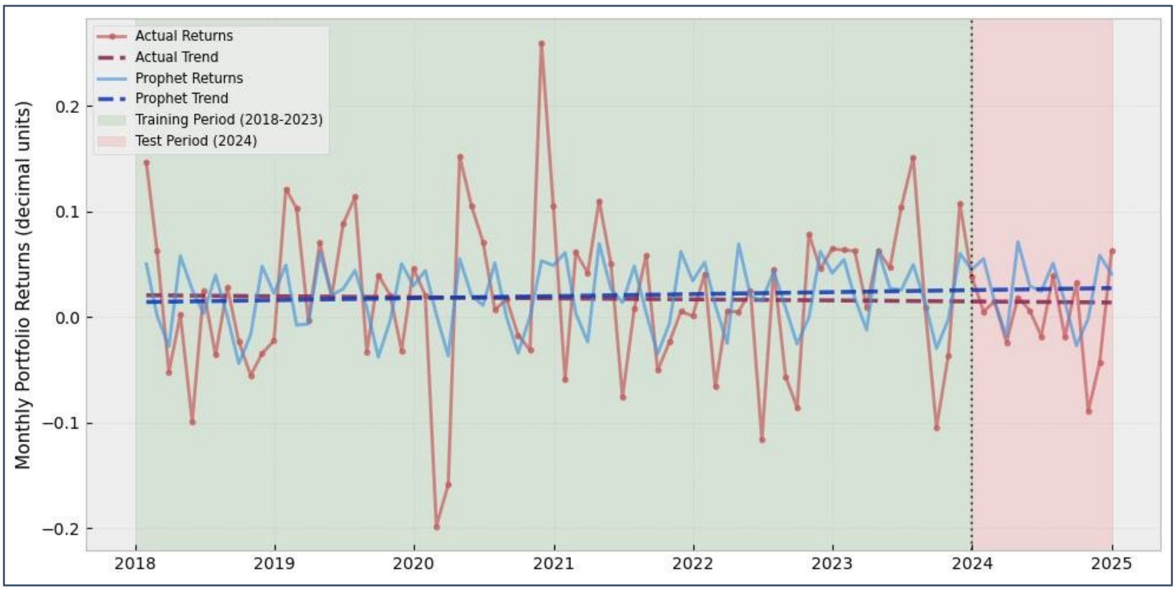


Figure 29: Actual vs. Forecasted monthly portfolio returns and trendlines for the full dataset

We now focus especially to the testing period (2024). From Figure 30, the offset between both trends (actual and forecasted) trend is about 2% for January and increasing at almost 3% at December 2024. For a small-scale portfolio this gap could be acceptable. But for larger scale portfolio it is a considerable deviation with great financial impact. Another troubling issue is that the gap widens over time weakening the forecast reliability.

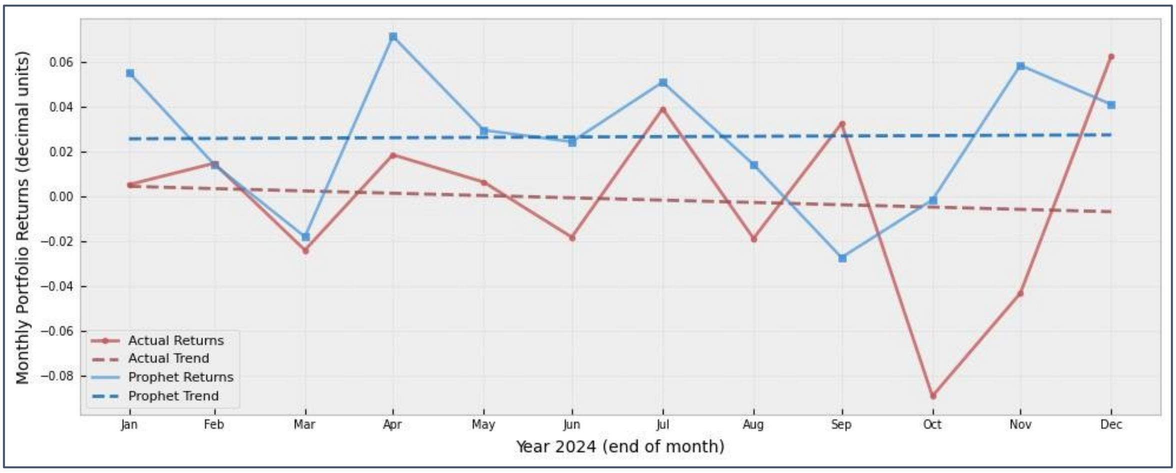


Figure 30: Actual vs. Forecasted monthly portfolio returns and trendlines for Year 2024

In the final part of this discussion, we shall explore the yearly seasonal component of the portfolio. Prophet has a build-in function, applied on the training data, that enables us to monitor how the portfolio returns repeatedly vary across the calendar year. Figure 31 shows that we should expect a strong reduction in our portfolio returns particularly on months March to April and September to October. Conversely, the highest returns are usually earned from April to May and to a less extent from middle of January to February and November to December. As such, the model detects a recurring seasonal pattern that is neither linear nor consistently stable throughout the year. The variation in returns is approximated slightly 8%.

Altogether, we should be aware that timing matters for portfolio performance but it does not explain everything on its own. If the portfolio encounters increased market volatility, as seen for 2024, the seasonal pattern derived from earlier years data may not fully correspond to the realized one. Forecasting the future using data from the past is an inherent flaw of any time series process that should be systematically accounted for during the analysis.

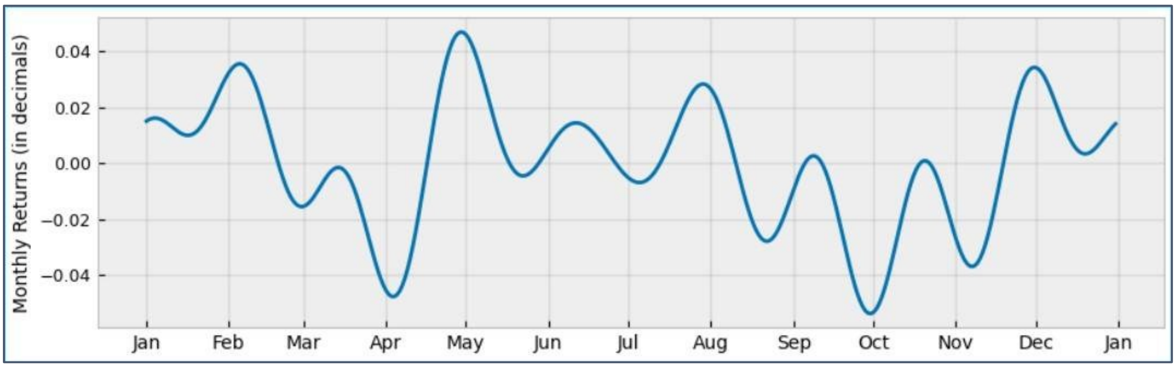


Figure 31: Yearly seasonality component for the Prophet refined model

4.4 Limitations and sensitivity analysis

The principal limitations of our study are acknowledged below:

- Our empirical analysis relies on a test dataset of 72 monthly readings. This is a rather small sample, especially for regression analysis done in the context of the CAPM and FF3 model, as it can increase the margin of error.
- Although Prophet is a flexible forecasting tool, it is primarily designed for business time series data with clear trend, systematic seasonality patterns and holiday effects. By removing the holidays layer as well as the weekly and daily seasonality, as they

had no meaning in our monthly dataset, we may have weakened the model's predictive consistency and reliability.

- The FF3 factors integrated as regressors into Prophet model do not account for exogenous indicators. Factors such as market sentiment, momentum or other macroeconomic ratios could possibly have a greater impact in the refined Prophet model predictions than SMB and HML alone.
- The price currency posed another limitation in our research. The Fama-French factors are expressed in USD currency while our stock prices in EUR so a currency conversion was needed. The real effects from this currency conversion can't be quantified.
- As our portfolio consists only of Greek stocks, it was decided that a Greek Treasury bill was the most appropriate proxy for the risk-free rate. To avoid creating negative values during currency conversion, it was treated as equal to a USD risk free rate. Although this creates a currency discrepancy, it was a necessary practical choice that helped us to remove more significant ambiguities in the regression analysis.
- Finally, in our analysis it has been recognised that the portfolio stock weights are not rebalanced. This, however, is unrealistic as for the portfolio weights to remain the same it would require daily trading. In practice, this approach could not realistically be exercised as it would incur significant transaction costs.

A sensitivity analysis is important to strengthen the reliability of the models by showing whether the results remain stable under varied conditions. A first test would be to reevaluate the performance of all models over a different sample period. For example, directing the forecasting for Year 2023 would help us assess how the initially selected Prophet set up performs in another market phase. This approach could be further extended if we adjusted the Prophet settings to increase their fit on the data and record the performance by forecasting the portfolio return over years with high volatility e.g. Year 2019-2020-2021. Useful insights could then be derived on how the model reacts at such special circumstances.

A second test would be to train the exact same Prophet model of our study to a completely different stock mix. Portfolios with higher volatility, different sector concentration or even from assets from various markets could be examined. This exercise could show how sensitive our forecasting model is to a range of risk level that each portfolio has and assist us in taking more informed decisions. It also appears necessary to reduce the training time of the model as this has a significant effect in the derivation of the trend.

5 Conclusions

5.1 Consolidation of main findings and key insights

This study set out to investigate whether returns from a Greek based stock portfolio can be reliably forecasted using both traditional asset pricing models (CAPM, FF3) and data-driven approaches, and to what extent can these methods be useful to the financial community.

The findings in this study confirm that traditional models can indeed be applied on the Greek market. They can also extend our knowledge in the fact that their use lies predominantly in explaining historic returns consistently rather than forecasting them. CAPM and FF3 produced similar regression results both with high explanatory power. At the same time the positive intercept (alpha) should be treated cautiously as Modern Portfolio Theory suggests that it is more likely associated to chance or statistic noise as an investment can not perform better than the predicted risk undertaken in the long term. Although using the lagged factor method to forecast next month's returns managed to outperform the naïve benchmark for most monthly predictions, it is believed that they are not sufficiently reliable.

To our knowledge, this study is among the first to utilise Prophet in forecasting stock returns listed in the Greek stock exchange. The evidence from our empirical research indicate that Prophet is a stronger candidate in returns forecasting compared to traditional models. It consistently outperformed the naïve forecast during normal market conditions but was not able to handle sharp changes in the returns. Still, it can not be overlooked that the model was incapable to predict the slightly negative trendline of the true monthly returns albeit it this closely associated with the feature selection during model fitting. The minimal predictive value that was added to the model when integrating FF3 factors as external regressors is noteworthy.

It is unfortunate that the study did not include the testing of more than one portfolio or the same portfolio at different time periods. Also, the feature selection in Prophet were very conservative and reduced the model's ability to adapt to outliers. In this regard we did not manage to exploit to our advantage the full potential of the algorithm. Despite the fact that the method of lagged data applied on the traditional models has not produced positive feedback, this can still be of use in guiding future researchers to alternative modelling paths.

While investigating the feasibility of return forecasting, our work shed new light on the nature of Greek stocks and their strong attraction to the Athens General Index. The portfolio analysis showed that despite our effort to select stocks with negative (or at least low) correlation, the resulting basket of assets still moved closely with the market index. The results of this research, therefore, raise an important question as to whether it is feasible to construct a diversified portfolio consisting only of Greek equities. This new understanding should help future researchers in finding new methods to better address this limitation.

In total, this thesis has provided a deeper insight in forecasting stock portfolio returns and showed that this is feasible to a limited extent. The methods and processed presented are a step further in comprehending stock returns forecasting. Yet they require thoughtful consideration by analysts and should be used together with other scientific tools in order to provide a complete proposal in forecasting portfolio returns.

5.2 Practical implications for investment decision-making

The insights gained from this study can be of assistance to finance practitioners involved in investment decision-making in the Greek stock market.

Our empirical results suggest that forecasting models can be helpful as supporting tools in the evaluation and prediction of portfolio performance. Still, none of our examined models was able to predict monthly returns without a high margin of error especially on periods when true returns presented high variability. Thus, analysts should treat model forecasting primarily as indicators for the possible direction of future returns rather than as definitive estimates that should be followed inarguably.

We can not stress out more how important it is for investors to be protected against unsystematic risk. In the present study both CAPM and FF3 models illustrated the substantial influence of the market on the performance of our portfolio and confirmed that traditional models remain pertinent today for understanding the risk profile of a portfolio. More specifically, the FF3 model also showed how sensitive a portfolio is to size and growth characteristics of a company based on preselected stock weights. The value of this information is important to investors as it helps explain how a portfolio is likely to react under different market conditions even when the forecasting models make predictions with limited accuracy.

As observed in the discussion section of this thesis, Prophet smoothed out historic returns and made moderate predictions for the future returns of the portfolio. In effect, this could serve as an ideal baseline on which finance practitioners can encompass risk mitigation mechanisms to protect against high volatility that models can not foresee. A step further could be to exploit the seasonal return patterns obtained from Prophet and use them for a close monitoring of the portfolio behaviour.

A limitation of our empirical study is that it does not fully consider the problems that can appear in real investing situations. Hence, caution should be taken when generalising these findings and applying them to practical real-world situations without critically analysing them. Besides what is common belief in our days is that the market complies with macroeconomics rather than seasonal patterns, so patience becomes the most precious and rare asset when it comes to stock investing.

One of the most significant contributions of this study is that it offers a practical framework that blends traditional econometric systems with a modern data-driven algorithm for forecasting in the Greek stock market. Although far from perfect, it could still serve as a valuable tool that may assist us in taking more informed decisions in the stock market arena.

5.3 Roadmap to further research

To better understand the implications of our results, future studies could address the following research fields.

Considerably more work will need to be done to determine the number and the type of assets a modern Greek portfolio should consist of to be as diversified as possible.

Further research should be carried out to establish the Fama-French factors in euro currency terms for Greece. Obtaining factor data for Greek stocks was proven to be a challenging task for this thesis since such datasets are generally scarce and, in most cases, not publicly available for free.

A natural progression of the present work is to use an autoregressive machine learning algorithm, such as Neural Prophet, to possibly improve predictive performance while employing risk proxies like bitcoin or gold prices as external regressor to the model.

References

Foreign

Angelo, M. D., Fadhiilrahman, I., & Purnama, Y. (2023). Comparative Analysis of ARIMA and Prophet Algorithms in Bitcoin Price Forecasting. *Procedia Computer science, 227*, 490-499. https://doi.org/10.1016/j.procs.2023.10.550

Angelopoulou, M. (2018). *Examining the factors that explain stock returns in the Greek stock market through the CAPM and the Fama and French Model.* Hellenic Open University.

Annapoorna, E., Sujil, S. V., Sreepriya, S., Abhishek, S., & Anjali, T. (2024). Revolutionizing Stock Price Prediction with Automated Facebook Prophet Analysis. *2024 International Conference on Inventive Computation Technologies (ICICT).* Lalitpur: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICICT60155.2024.10544766

Berenson, M. L., Levine, D. M., Szabat, K. A., & Stephan, D. F. (2020). *Basic Business Statistics: Concepts & Application* (14th ed.). Pearsons.

Brealey, R. A., Myers, S. C., & Marcus, A. J. (2015). In R. A. Brealey, S. C. Myers, & A. J. Marcus, *Fundamentals of Corporate Finance* (pp. 234-259, 270-289, 356-377, 386-405, 460-483). New York: McGraw-Hill.

Cheng, J., Tiwari, S., Djebbouri, K., Mandeep, M., & Shahzad, U. (2023). Forecasting Bitcoin prices using artificial intelligence: Combination of ML,SARIMA and Facebook Prophet models. *Technological Forecasting & Social Change*, 1-15. https://doi.org/https://doi.org/10.1016/j.techfore.2023.122938

Damodaran, A. (2014, August 22). *Session 4: Defining and Measuring Risk*. Retrieved from Aswath Damodaran on Valuation: https://www.youtube.com/watch?v=qofdyX0vudk

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance, 46*(5), 1575-1617. https://doi.org/https://doi.org/10.2307/2328565

Fama, E. F., & French, K. R. (1992). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56. https://doi.org/https://doi.org/10.1016/0304-405X(93)90023-5

Forjan, J. (2019, March 22). *Arbitrage Pricing Theory and Multifactor Models of Risk and Return (FRM P1 – Book 1 – Chapter 12)*. Retrieved from AnalystPrep: https://www.youtube.com/watch?v=FR3Fl-GUbqc&t=796s

Forjan, J. (2025, August 4). *MPT and CAPM (FRM Part 1 2025 – Book 1 – Chapter 5)*. Retrieved from AnalystPrep: https://www.youtube.com/watch?v=gOh3aJB5fRI&list=PLIYnk9FMYcksCoVznFEJArmFVR-R5KP0i&index=11

Ganti, A. (2025, 21 Sunday). *Understanding the Efficient Frontier: Maximize Returns, Minimize Risk.* Retrieved from Investopedia: https://www.investopedia.com/terms/e/efficientfrontier.asp#citation-9

Giouvritsa, T. (2016). *Greek capital market. Efficiency vs Inefficiency.* Hellenic Open University.

Hagmann, A. (2025, 10 1). *Complete 2-in-1 Python for Business and Finance Bootcamp*. Retrieved from Udemy: https://www.udemy.com/course/complete-python-for-business-and-finance-bootcamp/learn/lecture/16820564?start=270#overview

Hayes, A. (2025, August 6). *What Is the Risk-Free Rate of Return, and Does It Really Exist?* Retrieved from Investopedia: https://www.investopedia.com/terms/r/risk-freerate.asp

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems With Applications*, 226-251. https://doi.org/https://doi.org/10.1016/j.eswa.2019.01.012

Hyndman, R. J., & Koehler, A. B. (2006). Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting, 22*(4), 679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Indulkar, Y. (2021). 2021 International Conference on Emerging Smart Computing and Informatics (ESCI). *International Conference on Emerging Smart Computing and Informatics (ESCI).* Pune: AISSMS Institute of Information Technology. https://doi.org/10.1109/ESCI50559.2021.9397004

Jahangiry, P. (2025, Apr 5). *Module 8- part 1- Deep Dive into Facebook Prophet: Paper Review – Helpful but Overhyped?* Retrieved from pjalgotrader.github.io: https://www.youtube.com/watch?v=_6JW5qdOZG8

Khoa, B. T., & Huynh, T. T. (2023). A Comparison of CAPM and Fama-French Three-Factor Model Under Machine Learning Approaching. *JOURNAL OF EASTERN EUROPEAN AND CENTRAL ASIAN RESEARCH, 10*(7), 1100-1111. https://doi.org/10.15549/jeecar.v10i7.1402

Manolakis, V. (2012). *Fama and French Three-Factor model: Application to Greek Stock Market.* University of Macedonia.

Maris, G. (2009). *Application of the Fama and French.* Thessaloniki: University of Macedonia.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance, 7*(1), 77-91. https://doi.org/https://doi.org/10.2307/2975974

Mougiou, A. (2020). *An Empirical Investigation of Momentum Effect on the Greek Stock Market.* Hellenic Open University.

NeuralNine. (2024, November 18). *Evaluation Metrics For Regression - When & Why To Use What*. Retrieved from NeuralNine: https://www.youtube.com/watch?v=TrzUlo4BImM&t=2s

PrepNuggets. (2022, July 26). *CFA® Level I Portfolio Management - Minimum Variance Portfolios and Efficient Frontier*. Retrieved from PrepNuggets: https://www.youtube.com/watch?v=x45D7sIb9Mw

Reilly, F. K., & Brown, K. C. (2000). *Investment Analysis and Portfolio Management.* Dryden Press.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance, 19*(3), 425-442. https://doi.org/https://doi.org/10.2307/2977928

Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business, 39*(1), 119-138. https://doi.org/10.1086/29846

Taylor, S. J., & Letham, B. (2017). *Forecasting at scale.* PeerJ Preprints 5:e3190v2. https://doi.org/10.7287/peerj.preprints.3190v2

Thanh, B. K., & Huynh, T. T. (2023). Comparison of CAPM and Fama-French three-factor model under machine leanrning approaching. *JOURNAL OF EASTERN EUROPEAN AND CENTRAL ASIAN RESEARCH, 10*(7), 1100-1111. https://doi.org/https://doi.org/10.15549/jeecar.v10i7.1402

The Friendly Statistician. (2025, October 29). *What Are MAE, RMSE, And MAPE In Time Series Forecasting?* Retrieved from The Friendly Statistician: https://www.youtube.com/watch?v=2QuqzBsKkMw

Zhang, Y. (2025). A Comparative Study of Markowitz, CAPM and Fama-. *3rd International Conference on Management Research and Economic Development* (pp. 15-20). Macau, China: University of Macau. https://doi.org/10.54254/2754-1169/2025.23725

Greek

Ioannidis, A. K. (2024). *Stock prediction using the algorythm XGBOOST.* University of Macedonia.

Kousiou, E. (2005). *Portfolio Performance Measures on the Athens Stock Exchange*. Hellenic Open University.

Koutsianikouli, E. (2022). *Machine Learning in Portfolio Construction*. Hellenic Open University.

Michailidis, G. M. (2009). *Analysis of the Greek stock market from the perspective of modern portfolio theory*. University of Macedonia.

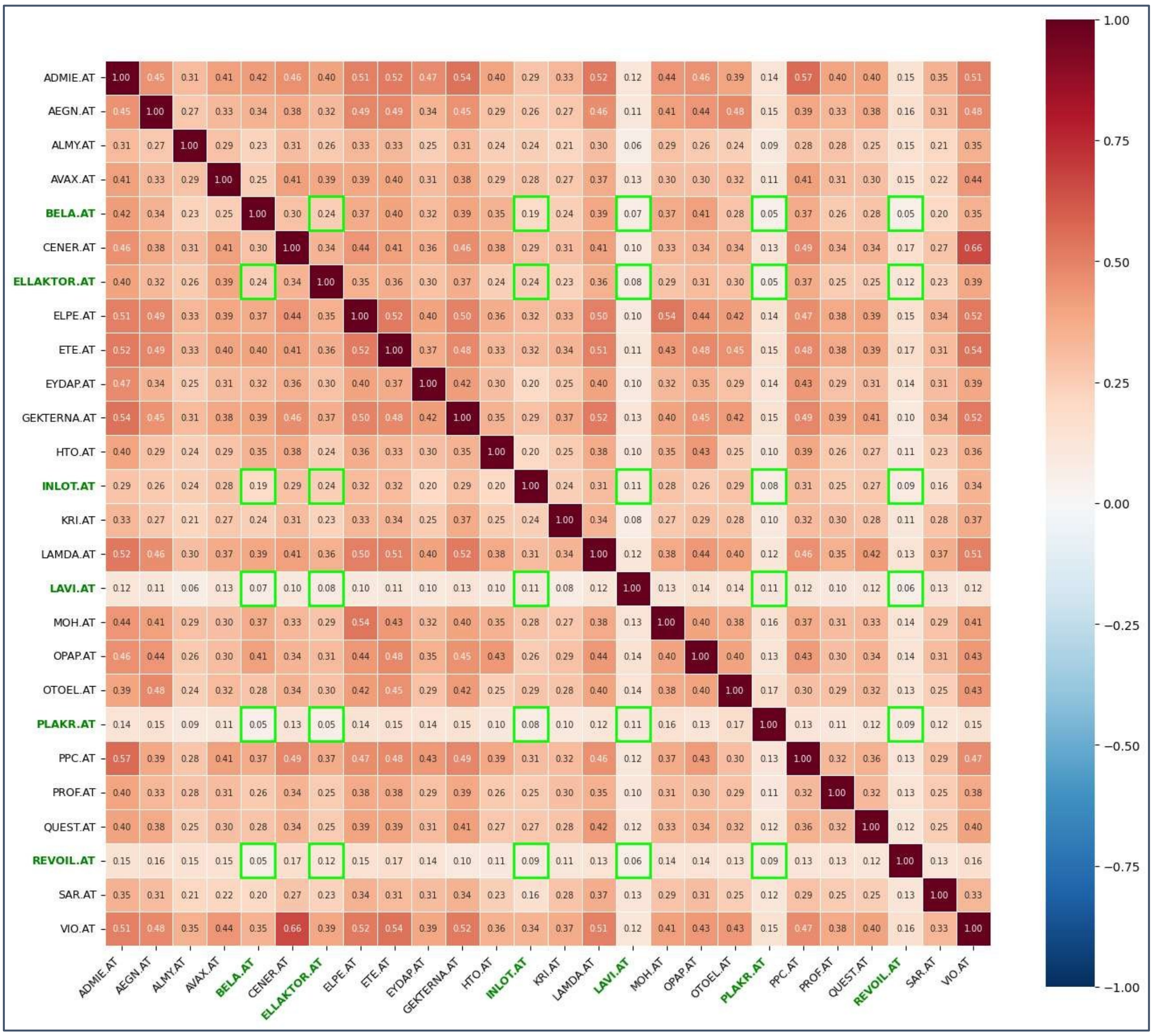
Papadimitriou, M. (2019). *Effective portfolio management through risk assessment factors and expected return*. Hellenic Open University.

Tzanetis, R. N. (2024). *Stock return predictions in the Greek stock market using machine learning.* University of Piraeus.

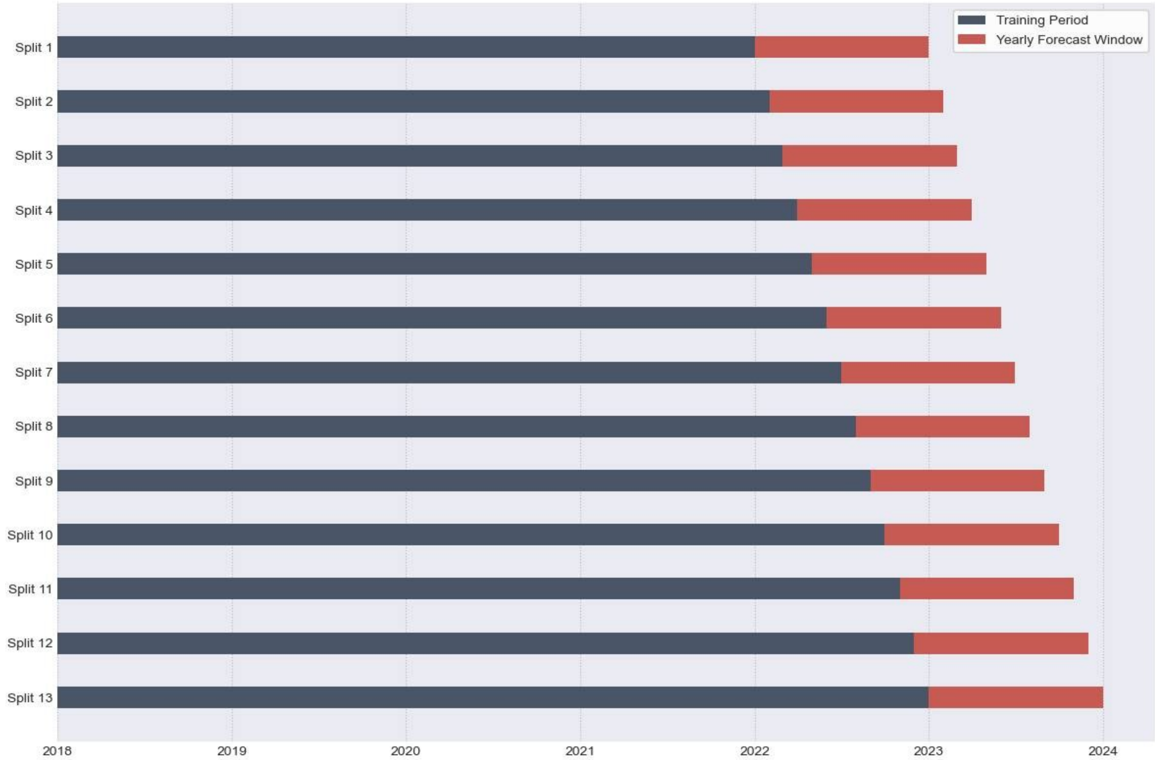
Appendix A: "Athens Stock Exchange: Eligible Companies"

Yahoo Fin. Ticker	Greek Company Name	Sector/Industry	Capitalisation	Brief Description
ADMIE.AT	Independent Power Transmission Operator	Utilities	Mid cap	Owner and operator of the Greek High Voltage Electricity Transmission System
AEGN.AT	Aegean Airlines S.A.	Aviation	Large cap	The flag carrier and largest airline of Greece, providing extensive domestic and international flight services as a Star Alliance member
ALMY. AT	Alumil Aluminium Industry S.A.	Aluminium/Construction	Small cap	An international group specializing in the design and production of advanced architectural aluminum systems
AVAX.AT	Avax S.A.	Construction	Mid cap	One of Greece's largest construction firms, managing major infrastructure, energy, and industrial projects
BELA.AT	Jumbo S.A.	Retail	Large cap	A major retail group specializing in toys, baby products and seasonal items with a strong presence in Southeastern Europe
CENER.AT	Cenergy Holdings S.A.	Industrial Goods & Services	Mid cap	Specializes in energy and data transmission through high-voltage cables and steel pipe manufacturing
ELLAKTOR.AT	Ellaktor S.A.	Construction	Large cap	A major infrastructure group with operations in concessions, real estate and environmental services
ELPE.AT	HELLENIQ ENERGY Holdings S.A. (formerly Hellenic Petroleum)	Oil & Gas	Large cap	A leading integrated energy group in SE Europe active in refining and fuel marketing
ETE.AT	National Bank of Greece S.A.	Banking	Large cap	The largest financial institution in Greece
EYDAP.AT	ATHENS WATER SUPPLY & SEWERAGE S.A.	Utilities	Large cap	The National water supply and sewerage company in Greece.
GEKTERNA.AT	GEK Terna S.A.	Construction & Energy	Large cap	A diversified group leading in construction, infrastructure concessions and renewable energy production
HTO.AT	Hellenic Telecommunications Organization S.A.	Telecoms	Large cap	One of the largest telecommunications provider in Greece, offering fixed-line, mobile, broadband, and pay-TV services
INLOT.AT	Intralot S.A. Integrated Lottery Systems and Services	Gaming Technology	Large cap	A global technology supplier and operator in the lottery and gaming sector
KRI.AT	Kri-Kri Milk Industry S.A.	Food & Beverage	Mid cap	A leading dairy company and great exporter specializing in yogurt and ice cream
LAMDA.AT	Lamda Development S.A.	Real Estate/Development	Large cap	A real estate leader specializing in large-scale commercial projects, most notably *The Ellinikon* complex in Athens
LAVI.AT	Lavipharm S.A	Healthcare	Small cap	Integrated pharmaceutical group specializing in medicine and distribution.
MOH.AT	Motor Oil (Hellas) Refineries S.A.	Oil & Gas	Large cap	A leading oil refinery and marketing company operating one of Europe's most advanced refineries and a vast gas station network
OPAP.AT	OPAP S.A. (Organization of Football Prognostics S.A.)	Gaming	Large cap	Greece's leading sports betting operator and one of Europe's largest managing lotteries
OTOEL.AT	Autohellas S.A.	Leisure/Car Rental	Mid cap	One of the largest car rental and fleet management company in Greece
PLAKR.AT	Plastika Kritis S.A.	Plastics Industry	Mid cap	A global manufacturer of plastics and agricultural films with multiple international production sites
PPC.AT	Public Power Corporation S.A. (DEH)	Utilities	Large cap	The largest electricity producer and supplier in Greece
PROF.AT	Profile Systems & Software S.A	Technology	Small cap	Software provider for or Banking, Treasury, Risk, FinTech and Investment Management
QUEST.AT	Quest Holdings S.A.	Technology	Mid cap	A diverse group active in telecommunications, freight services and green energy
REVOIL.AT	Revoil S.A	Oil & Gas	Small cap	Energy company focused on storage and distribution of petroleum products
SAR.AT	Sarantis S.A	Household/Personal	Large cap	Manufacturer and distributor of consumer and health care products
VIO.AT	Viohalco S.A.	Diversified Industrials	Large cap	An industrial holding company of leading groups in aluminum, copper and steel pipes
TOTAL			26	

Appendix B: "Pearson Correlation Matrix for 26 stocks"



Appendix C: “Rolling Window Cross Validation Chart”



Author's Statement:

I hereby expressly declare that, according to the article 8 of Law 1559/1986, this dissertation is solely the product of my personal work, does not infringe any intellectual property, personality and personal data rights of third parties, does not contain works/contributions from third parties for which the permission of the authors/beneficiaries is required, is not the product of partial or total plagiarism, and that the sources used are limited to the literature references alone and meet the rules of scientific citations.