

Neural Networks Applied to Financial Forecast

Matheus Rigoni Galvão - 54185

Dissertation presented to the School of Technology and Management of Bragança to obtain the Master's Degree in Informatics within the scope of the double degree program with the Federal Technological University of Paraná

Supervisors:

Rui Pedro Lopes

Simone Aires

This document does not include the suggestions made by the board.

Bragança

2023

Acknowledgements

I extend my heartfelt gratitude to my family for all the love and support that made my studies possible, as well to Professor Rui Pedro Lopes from IPB and Professor Simone Aires from UTFPR for their unwavering support and thoughtful guidance throughout the course of this thesis, and above all, throughout my constant pursuit of knowledge. Their expertise and mentorship have been instrumental in shaping and refining my work, and I am truly fortunate to have had the opportunity to learn from their wisdom, experience, and kindness.

Abstract

The realm of bond yield forecasting holds substantial importance, influencing critical economic decisions. Recent research has focused on harnessing the predictive capabilities of advanced deep neural network architectures for this purpose, yet the integration of macroeconomic data remains relatively unexplored. This study aims to evaluate the performance of LSTM models in bond yield forecasting using macroeconomic data and to discern the most pivotal features within the constructed macroeconomic dataset.

The experimental findings unveil five preeminent macroeconomic variables—namely interest rates, inflation, gross domestic product, total debt, money supply, and the intricate interrelation among bond yields of varying maturities—as key contributors to bond yield forecasting. The performance evaluation of the LSTM models demonstrates their aptitude for generating precise one-step predictions, particularly when fed solely with the target’s historical time-series data.

Nonetheless, a noteworthy limitation surfaces concerning the multi-step forecasting approach employed within this study. This methodology struggles to encapsulate forthcoming volatility within the target variables. Consequently, an alternative methodology is proposed as a stepping stone for future research, aiming to overcome this limitation and enhance predictive accuracy.

Keywords: Bond Yields, Forecasting, LSTM, Macroeconomics

Contents

1	Introduction	1
1.1	Objective and Research Questions	2
1.2	Document Structure	2
2	Related Work	4
2.1	Background	5
2.1.1	Bond	5
2.1.2	U.S. Treasury Bonds	5
2.1.3	Yield Curve	6
2.1.4	Importance of the Yield Curve	7
2.1.5	Naive Random Walk Forecast	7
2.1.6	Recurrent Neural Networks	8
2.1.7	LSTM	9
2.2	Forecasting Yield Curves	10
2.3	Financial Time Series Forecasting using Macroeconomic Data	11
2.4	Neural Network models for Financial Time Series Forecasting	12
2.5	Neural Network models for Yield Curve Forecasting	13
3	Methods	15
3.1	Data	16
3.1.1	Collection	16
3.1.2	Processing	17

3.1.3	Feature selection methods	21
3.2	Forecasting Models	22
3.2.1	Hyperparameter Tuning	23
3.2.2	Training	24
3.2.3	Univariate LSTM	24
3.2.4	Multi-Uni LSTM	25
3.2.5	Multivariate LSTM	26
3.3	Evaluation	27
3.3.1	Root Mean Squared Error	27
3.3.2	One-Step Forecasting	28
3.3.3	Multi-Step Forecasting	28
4	Results	31
4.1	Feature Selection	32
4.1.1	Correlation Coefficient	32
4.1.2	Mutual Information	34
4.1.3	Cross Correlation Function	35
4.1.4	Final Interpretation of Features	36
4.2	Hyperparameter Tuning	36
4.2.1	Univariate LSTM	37
4.2.2	Multi-Uni LSTM	37
4.2.3	Multivariate LSTM	38
4.3	Models Performance Evaluation	38
5	Conclusion and Discussion	43

Acronyms

CNN Convolutional Neural Network. 12, 13

GDP Gross Domestic Product. 11, 32, 34–36

LSTM Long Short-Term Memory. 2, 4, 8–10, 12, 14, 15, 22–26, 29, 31, 37, 39–41, 43–47

ML Machine Learning. 1, 13, 23

MLP Multilayer Perceptron. 12, 14

NN Neural Network. 1, 4, 5, 12–14, 23, 43, 47

NRW Naive Random Walk. 7, 8, 14, 24, 38, 41, 45, 46

RMSE Root Mean Squared Error. 27, 28, 32, 38

RNN Recurrent Neural Network. 1, 8–10, 12, 13

TCN Temporal Convolutional Network. 14

U.S. United States. 1, 2, 5–7

USD United States Dollar. 1

Chapter 1

Introduction

The United States (U.S.) Treasury bond market is considered a safe haven for investors and a reference point for the American economy and its future. With an overall size of around United States Dollar (USD) 23 trillion in 2022 [1], it is the biggest fixed-income market in the world. The yield curve of the U.S. Treasury bond market has been closely examined as a powerful indicator of economic recessions [2], therefore yields are constantly observed by investors as a measure of risk and expectations for the markets, being a piece of information of great interest by institutions, companies and governments.

Modeling and forecasting bond yields, as any other asset class, is an arduous task previously attempted by many statistical and Machine Learning (ML) models, in which the use of Neural Networks (NNs) has shown great results, the field has seen numerous advances proven successful in many real-world and financial applications, benefiting from their capability of learning complex non-linear relationships from raw numerical data [3]. A particular architecture called Recurrent Neural Network (RNN) presents remarkable results in the use of sequential data, and because of that, it has been used and seen as a promising tool for many time-series forecasting studies.

Studies such as [4] find strong evidence of macroeconomic effects on the future yield curve, even though most forecasting models consider solely the yield time-series data, possibly limiting its potential. This work proposes the application of a type of RNN

model, called Long Short-Term Memory (LSTM) Network, taking a macroeconomic approach to forecast different maturities of U.S. Treasury bond yields by building and using a dataset from multiple macroeconomic indicators of the American economy, such as inflation rate, unemployment, economic growth, stock market indexes, and currency value, among others.

1.1 Objective and Research Questions

The main objectives of this work are the evaluation of a LSTM model fed by macroeconomic data to forecast yields of bonds issued by the U.S. Treasury with different maturities, closely examine the advantages and limitations, alongside their causes considering the model and data used. By fulfilling the objectives determined, this work aims to answer the following research questions:

- *RQ1: What is the best selection found of macroeconomic features?*
- *RQ2: Is it possible for a LSTM model using macroeconomic data to make better forecasts than one using only the time-series data from the bond yields?*
- *RQ3: How well does a model using macroeconomic data perform in comparison to the benchmark model used in the literature?*

1.2 Document Structure

The present thesis is structured in the following manner. Chapter 2 provides a theoretical background with relevant information related to finance, economy and artificial intelligence in order to better clarify the elements present in this study, it also encapsulates the works related to the problem from the literature, since the application of classic statistical models to recent deep architectures, alongside the type of data utilized. Chapter 3 describes and justifies the methods applied in this study, concerning the development, tuning and evaluation of the chosen model, along with the steps taken to collect, prepare

and process the data used. Chapter 4 displays the results obtained from the series of tests. Finally, Chapter 5 presents the conclusions, reflects upon the methods, and discusses the validity of the study and possible improvements to be made in further research.

Chapter 2

Related Work

The study of NN applications to forecast yield curves has been growing in the last few years, with the development of new models focused on time series prediction and a better understanding of their limitations. A considerable amount of the work on yield curve forecasting was mainly done by researchers in the field of economy, therefore most of the applications are based on classic statistical models. More recent studies in financial time series forecasting explore the use of NN in order to take advantage of their ability to find non-linear patterns and produce accurate solutions in environments characterized by complex information. Most applications of NN to forecasting in finance focus on stock market and asset price prediction, some of those also applying macroeconomic data as input. Following this approach, a few studies develop similar applications to forecast yield curves, most of them choosing the LSTM architecture. This chapter accounts for a background section with the purpose of introducing the reader to the theoretical terms used in the paper, and a summary of previous research related to this study, grouped into four categories, regarding the forecast of yield curves, the use of macroeconomic data, NNs in financial forecasting and yields in specific.

2.1 Background

This section provides a brief introduction to relevant financial, economic, and technical terms present in the following chapters, such as what is a bond, the specific type of bond studied, an introduction to the yield curve, shapes, and importance. Furthermore, specific NN architectural terms are explained.

2.1.1 Bond

A bond is a debt security, issued by borrowers to raise money from investors willing to lend them money for a certain amount of time [5]. Bonds are issued with a predetermined maturity date and a coupon rate, the former determines when the investor will receive the total amount paid for the bond back and the latter is the interest paid over a specified amount of time to the lender. The issuer promises to pay the lender a specified rate of interest during the bond's life and to return the amount paid by the investor when the specified maturity is reached [6]. One of the largest and most influential entities in the global bond market is the U.S. Department of the Treasury, responsible for issuing debt securities commonly referred to as U.S. Treasury bonds.

2.1.2 U.S. Treasury Bonds

U.S. Treasury Bonds are issued by the U.S. Department of the Treasury on behalf of the federal government. Because this kind of security relies on the government, it is considered a safe and popular investment. Investors can acquire bonds directly from the issuer or from other investors in the market, which determines the bond's price by comparing its coupon rate to other bonds with similar maturities. Due to the influential role played by the American Economy on financial markets, U.S. Treasury Bonds set a benchmark for yield and risk assessments through mechanisms like the yield curve.

2.1.3 Yield Curve

The yield curve is a graphical representation that plots the interest rates of bonds with different maturities but similar risk levels. The x-axis shows the time to maturity, while the y-axis displays the yield, also known as current yield, which is a measure of return an investor realizes on a bond, calculated by dividing the annual coupon payment of the bond by its current price [6]. Bonds with different maturities and the same credibility can be compared by their yields using the yield curve graph, usually referring to bonds issued by the U.S. Treasury. Therefore, to predict the yield curve in a given time, one must predict the yields for each bond maturity. As illustrated in Figure 2.1, the curve is considered normal when bonds with longer maturities have bigger yields than the ones with shorter maturities, flat if they present the same yields and inverted if the long-term bonds have smaller yields than short-term bonds. These different patterns in the curve have been observed by governments and other institutions over the years, as it is believed that the curve contains valuable information on the economic health of a country, proving the importance of such a graph.

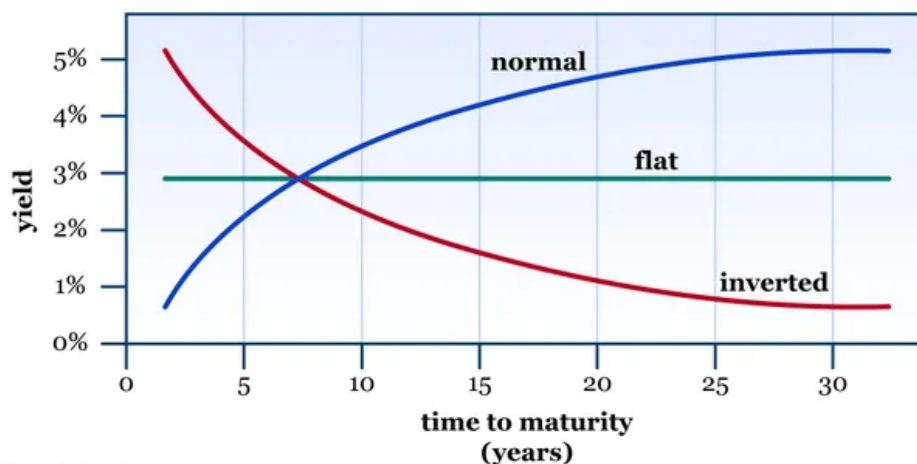


Figure 2.1: Yield curves with different shapes. (Source: Britannica)

2.1.4 Importance of the Yield Curve

A wide range of investors closely watch the yield curve of U.S. Treasury Bonds, the curve is considered to be an indicator of economic strength and a powerful model to predict the future of America's economic health [7]. A normal curve represents good expectations for the future of the economy, on the other hand, an inverted curve is a sign of weak economic expectations. Previous recessions in the USA followed yield curve inversions, for that reason, the curve is often attributed with possible predicting powers for economic recessions, [8] points that an inverted yield curve does not forecast economic recessions, but in fact, it may forecast the conditions that make a recession more likely.

2.1.5 Naive Random Walk Forecast

The Naive Random Walk (NRW) algorithm is a forecasting model that assumes that the future values of a time series are equal to the current value, not considering any external factors or past observations. This is also known as a “no-change” forecast, where the predicted value is simply the current value of the time series [9]. The prediction logic used by the algorithm can be mathematically expressed in the equation 2.1.

$$Y_{t+1} = Y_t \tag{2.1}$$

Because it represents the simplest possible forecasting model, this model is often used as a baseline for comparison with more complex forecasting models. Any more complex model that cannot outperform the NRW model can be considered inadequate [10] or indicate that the time series is highly unpredictable and difficult to forecast accurately. Conversely, if a more complex model significantly outperforms the NRW model, this may suggest that there are clear trends and patterns in the time series that can be leveraged for improved forecasting accuracy [11].

Many financial indexes are widely believed to be random and unpredictable, which makes it difficult to develop accurate forecasting models for them. The NRW algorithm

provides a simple baseline model for such time series, if a complex model cannot outperform that baseline, then it is unlikely to be useful in practice, as it is not able to capture any predictable patterns in the data beyond the random fluctuations. Moreover, the use of this model as a benchmark allows researchers to compare the predictive accuracy of different forecasting models and to determine whether any improvement in accuracy is statistically significant [9].

As previously stated, the NRW algorithm serves as a benchmark for comparing its performance against more complex forecasting models. In the literature, one widely used neural model is the RNN architecture, defined in the following section, which is commonly evaluated in relation to the NRW.

2.1.6 Recurrent Neural Networks

A Recurrent Neural Network is an artificial neural network architecture applied to a wide variety of problems involving time sequences of events and ordered data. Among the most common applications are word prediction, motion detection, music synthesis and financial forecasting [12]. The major difference this kind of model has from common feed-forward networks is the presence of hidden states that work as a memory to represent past states of the network, arranged in cycles it allows output from some nodes to affect subsequent input to the same nodes, giving it the capacity to represent temporal behavior. The standard models suffer from a problem called vanishing gradient, which happens when the gradient is constantly reduced by back-propagating through the layers and steps to a point where the network stops learning, for that reason RNNs struggles to learn long-term relations. To address the limitations of standard RNNs, particularly in learning long-term dependencies, other enhanced RNN architectures were introduced, such as LSTM networks.

2.1.7 LSTM

Long Short-Term Memory (LSTM) is a type of RNN architecture that is particularly effective at modeling sequential data. The LSTM network was introduced by Hochreiter and Schmidhuber in 1997 [13], and since then, it has been widely used in a variety of applications, including natural language processing, speech recognition, and time series analysis.

The key feature of the LSTM models are their ability to capture long-term dependencies in sequential data. Traditional RNNs suffer from the vanishing gradient problem [14], which makes it difficult to learn long-term dependencies. LSTMs overcome this problem by introducing memory cells that can retain information over extended periods [13], and gating mechanisms that control the flow of information into and out of the cells [15]. This makes them particularly effective for working with sequential data that has long-term dependencies, including application for financial data analysis, as financial data often exhibits complex temporal patterns.

The architecture of a LSTM cell, as shown in Figure 2.2, consists of several interconnected components, including input, output, and forget gates, which regulate the information flow within the network [13]. The input gate allows new information to enter the memory cell, while the forget gate removes information that is no longer relevant. Finally, the output gate controls the information flow from the memory cell to the next time step in the sequence.

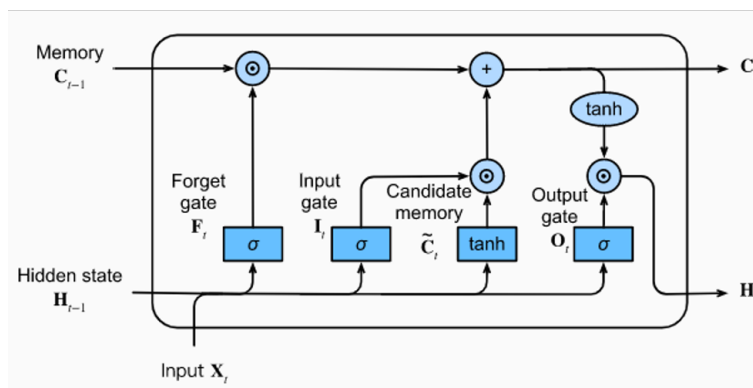


Figure 2.2: Architecture of a LSTM Unit (Source: Dive into Deep Learning)

This kind of model is considered well-suited for financial data analysis due to its ability to handle multiple inputs and outputs [16]. For example, when predicting stock prices, a LSTM can take into account a wide range of input features, including stock prices, volumes, and other relevant financial indicators, alongside the multiple features the model receives as inputs, it can also deal with multiple past values to take into account for its prediction, with a much larger capacity than common RNN, which suffer more from vanishing and exploding gradients. Moreover, LSTMs can generate multiple outputs, taking into account what has been fed to the model as input, and being able to predict many different values with one single model.

2.2 Forecasting Yield Curves

Several studies aimed to model and forecast yield curves through econometric and statistical approaches, one of the most cited and relevant in the related literature is the Nelson-Siegel parsimonious model [17]. It was later improved by Diebold and Li [18], who interpreted the yield curve as a dynamic model that achieves dimensionality reduction via factor structure. Their model has shown superior results at long horizons compared to various standard benchmark forecasts.

Further research investigating the results of the Diebold and Li [18] and Duffee [19] models was conducted by [20]. The study concluded that the results produced by these models do not differ significantly from those obtained by a Random Walk, the suggested reason behind the poor performance of the forecasting models was their inability to predict the short-term rate, and that longer-term rates incorporate the markets' expectation for the future short-term rate.

The Random Walk is the most common benchmark model used to compare the accuracy of yield curve forecasting models such as [18], [19], [21]. Evidences from [19], [20] support the difficulty to outperform such model, especially in short-term horizons. This work also adopts a Random Walk approach as the baseline for accuracy comparison.

2.3 Financial Time Series Forecasting using Macroeconomic Data

The forecasting of financial time series is a popular field of study that aims to predict several financial assets and economic indicators such as the stock market, gold prices, economic growth, inflation rate, and bond yields. In order to achieve more accurate results, many researchers approach the forecasting problem by adopting economic theories of cause and effect relating macroeconomic factors and the outcome of a certain value, not only using the time series of that value as the model's input. The role of macroeconomic variables cannot be ignored, as they play an important role in shaping the financial systems of both developed and developing countries [22].

Stock market prediction is a common interest of numerous studies, many of them applying macroeconomic variables to forecasting models. For instance, macroeconomic indicators such as unemployment rate, oil price, interest rate, consumer price index, exchange rate, exports, imports, debt, Gross Domestic Product (GDP) were used to forecast stock markets returns, prices and volatility by [22], [23], [24] respectively.

The forecasting of macroeconomic indicators themselves are also a field of great interest in the literature, as the variations of one indicator may affect others and vice versa. An agent-based model was developed by [25] to forecast GDP, inflation, household consumption and investment using several other macroeconomic variables, including tax rates. Other works such as [26], [27] study prediction models for GDP and economic growth using government data like total expenditure, exports, and imports.

The relation between the yield curve and the macroeconomy is an object of study from several works, many attempted to better understand the effects they have on each other, for instance, [28] finds relations between the bond market and how its investors observe the GDP growth, inflation and federal funds rate. Macroeconomic variables can be used to model and to forecast yields, a model using such variables was developed by [4] having the one from [18] as its base, which resulted in strong evidences of macroeconomic effects on the yield curve. The use of macro factors has also improved the yield forecasting

performance of a model developed by [21], as it compares to models exclusively using unobservable factors. On the strength of such evidences, this work chooses to compose a macroeconomic dataset for the forecasting of yields.

2.4 Neural Network models for Financial Time Series Forecasting

Following the recent and significant advances in the field of artificial NNs, diverse studies have been applying deep model architectures to forecast financial time series. Reasons behind this interest include the high complexity of financial markets, with complex, high-dimensional, and noisy time series, which makes the use of deep learning techniques an attractive approach, as NN models are particularly well suited to find accurate solutions in environments characterized by complex, noisy, irrelevant, or partial information [29].

As stated before, stock market prediction is one of the most popular subjects present in the literature of financial forecasting using NN models. Different architectures have been applied to stock price forecasting, from the more classic networks, for instance, Multilayer Perceptron (MLP) [29], [30], to more advanced architectures designed for sequential problems, such as Transformers [31], Convolutional Neural Networks (CNNs) [32] and RNNs [23] [24], [16], other approaches include the use of Generative Adversarial Networks [33] and reinforcement learning [34]. Among the literature reviewed, most of the time RNNs have presented the best performances, especially the sub-category of Long Short-Term Memory networks [24], [16], [35]. Another study conducted by [36] compared models based on CNN, LSTM, and a common RNN, the most accurate results were found to be from the CNN model, the reason concluded was the CNN's non-dependence on any previous information for prediction, as it uses only the current window of time, making it a good fit for highly dynamical time-series, where the patterns and dynamics existing will not always be the same, such as the stock market.

Although CNNs have produced better results for dynamic markets, applications forecasting other markets and time series can find better results with RNNs, especially when there is a dependence on many previous windows of information, which requires networks capable of remembering past patterns.

2.5 Neural Network models for Yield Curve Forecasting

The bond market, unlike more dynamic markets, shows correlations with other macroeconomic time series, which may have prediction patterns from way back the current time window.

The literature related to the application of NNs to the forecasting of bond yields is recent and limited, the works found are still studying the best deep architectures to use and how to target the data. Most studies apply simple NN models, which are not the best suited for sequential problems, many times the reason behind that choice is the comparison with other ML models, not paying much attention to the deep architecture used. Works from [3] and [37] are examples that use an MLP network to forecast U.S. Treasury bond yields, although the architecture chosen is not the preferred choice when it comes to time-series problems, the results obtained were promising and superior to the ones from other ML models tested. The work of [37] also compares a multivariate approach that forecasts many bond yields of different maturities at once with the single-task approach, which forecasts one at a time through the model, the conclusions drawn from the results show no significant difference between the two approaches, making it an interesting and time-saving technique to forecast many yields all at once, as most studies forecast the yield curve in a single-task manner. A study comparing ML models and their capacity to model the yield curve was conducted by [38], in which a feed-forward NN is used to predict the yield curve of BRICs countries, the empirical evidence obtained has clearly shown the superior capability of the NN in interpolating the behaviour of the yield

curve as compared to the other models.

Hybrid models are also present in the literature, by combining NN models with statistical techniques, [39] and [40] have developed hybrid models using feed-forward NNs, the former combining with cointegration analysis techniques to forecast French and German bond yields, and the latter using a Gaussian linear state-space model applied to Brazilian bond yields, both studies were able to obtain better overall out-of-sample forecasts than traditional approaches, such as the dynamic Nelson and Siegel model and its extensions.

Regarding deep architectures more suited for time-series problems, the only available study applied to yield curve found was the work of [41], which compares the performance of two models based on Temporal Convolutional Network (TCN) and LSTM, having the NRW algorithm as a benchmark, both models overcame the performance of the NRW in longer forecast horizons. Two different datasets were used in the study, and each one of the two models obtained the best performance on one of the datasets, making it impossible to determine what architecture is the best for the problem, although the conclusion shows the great potential of deep models on the forecasting of yields.

The only work found in the related literature using NNs and macroeconomic data for bond yields forecasting was the work from [3], where the authors use four macroeconomic indicators representing economic growth, interest rates, inflation and stock market volatility, applying them to a MLP network, which showed a better performance than the econometric models compared.

No work in the literature applying a LSTM model has been found using macroeconomic variables for the prediction of yield curve, therefore the present work proposes a different approach by building a dataset composed of macroeconomic factors and implementing it with a deep LSTM network, which is considered in the literature to be a better suited NN architecture for time-series prediction, as restated in the works of [16], [24], [41].

Chapter 3

Methods

This chapter regards the methodology applied during the development of the project, the step-by-step process, employed to create the desired dataset, develop different models suited for the data input and output format, and properly evaluate their performances.

The study includes an evaluation of LSTM models on the forecasting of bond yields employing macroeconomic data, in which it was necessary to research, collect, ensemble, and process multiple different datasets, in order to create a robust and usable one, subsequently used to train the models.

Different types of LSTMs were developed, differing in the format of the input and output, due to factors related to the data and longer-term forecasting necessities. For instance, an univariate model was used to evaluate the performance using solely the time series data of one yield. A second one was designed to receive multiple inputs, including the macroeconomic variables, and output the prediction of a single yield. Lastly, a third one was built to receive multiple inputs and output the same number of inputs, such a model was necessary to overcome the challenge of predicting multiple steps into the future, due to the necessity of refeeding the model with each output. The architecture of each model is described more deeply in the chapter, alongside the hyperparameter tuning and training process.

The evaluation methods are also presented, justified and explained in this chapter, defining the error metric used, as well as the possible approaches to forecasting time

series, such as one or more steps into the future.

3.1 Data

The current section focuses on the whole process of working with the data introduced to the models afterward. It describes many steps, including researching the best macroeconomic datasets for the problem, identifying reliable sources, processing multiple datasets to create a usable one for the model, and selecting the best features within the processed dataset.

3.1.1 Collection

In order to input a robust macroeconomic dataset to the model developed, a custom dataset built upon the integration of other datasets was necessary. Following the research made on the main macroeconomic factors that may affect US bond yields, a process of data collection has been implemented, taking into account only reliable sources of information.

US bond yields are influenced by a range of macroeconomic factors, including economic growth, inflation, monetary policy, and global economic and social conditions. Economic growth is a key factor affecting bond yields since higher growth prospects increase the demand for credit, leading to higher yields [42]. Inflation also affects bond yields as higher inflation expectations can cause investors to demand higher yields to compensate for the loss in purchasing power over time [43]. Monetary policy decisions made by the Federal Reserve can also impact bond yields, as changes in interest rates and the money supply can affect borrowing costs and overall economic activity [44]. Finally, global economic conditions, such as changes in exchange rates, trade policies, and geopolitical risks, can also affect bond yields, as they can impact the overall demand for US Treasury bonds and the perceived risk associated with them [45].

The macroeconomic datasets collected were chosen in order to represent some of the previously mentioned factors. However, not all of the best-representing datasets are publicly available and free, which made it challenging to find reliable data on some factors.

Most of the collected datasets were found on the official website of the Federal Reserve of Saint Louis, also known as FRED, which is an online database that provides access to economic data and research tools for students, academics, and policymakers. Table 3.1 contains information about each macroeconomic dataset collected.

Factor	Dataset	Frequency	Range	Source
Inflation	CPIAUCSL	Monthly	1947-2022	FRED
Interest Rates	DFF	Daily	1954-2022	FRED
Unemployment	UNRATE	Monthly	1948-2022	FRED
Stock Market Volatility	VIXCLS	Daily (Week Days)	1990-2022	FRED
Money Supply	M1SL	Monthly	1959-2022	FRED
Economic Growth	GDP	Quarterly	1947-2022	FRED
Public Debt	GFDEBTN	Quarterly	1966-2022	FRED
Dollar Index	DXY	Monthly	1973-2020	Macrotrends

Table 3.1: Macroeconomic Datasets

The data regarding the bond yields was collected from the official website of the U.S. Department of the Treasury (<https://home.treasury.gov/interest-rates-data-csv-archive>), using the Daily Treasury Par Yield Curve Rates, which has a maximum date range from 1990 to 2021, containing values with a daily frequency, exclusively including work days. The mentioned dataset contains data from 12 different maturities, but not all of those have a date range from 1990, for that reason, only the following 5 maturities were used in this study: 3 months, 1 year, 3 years, 5 years and 10 years. Also considered as some of the most relevant for the yield curve, due to the fact they can represent short, medium and long-term classes of bonds.

3.1.2 Processing

One of the main difficulties encountered when working with multiple macroeconomic datasets is to arrange them in an usable format, each set has a different date range and frequency, therefore it is not simple to joint them all into one. A series of steps were taken in order to create a complete and robust dataset, from correctly arranging the

dates, filtering the range, dealing with missing values, and dividing them into separate sets to be later input to the model.

The first step was to put the datasets together by joining them by their dates, the result of such operation is displayed in Figure 3.1, where several missing values can be seen, due to the difference of date ranges of each dataset.

DATE	Inflation	Interest-rates	Unemployment	VIX	Money-supply	Dollar-index	GDP	Total-debt
1947-01-01	21.48	NaN	NaN	NaN	NaN	NaN	243.164	NaN
1947-02-01	21.62	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1947-03-01	22.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1947-04-01	22.00	NaN	NaN	NaN	NaN	NaN	245.968	NaN
1947-05-01	21.95	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
2022-11-14	NaN	NaN	NaN	23.730000	NaN	NaN	NaN	NaN
2022-11-15	NaN	NaN	NaN	24.540001	NaN	NaN	NaN	NaN
2022-11-16	NaN	NaN	NaN	24.110001	NaN	NaN	NaN	NaN
2022-11-17	NaN	NaN	NaN	23.930000	NaN	NaN	NaN	NaN
2022-11-18	NaN	NaN	NaN	23.930000	NaN	NaN	NaN	NaN

25062 rows × 8 columns

Figure 3.1: Joined datasets

With the purpose of reducing missing values, a date range present in all sets had to be chosen and applied in the filtering. The range used was from 02/01/1990 to 02/01/2020, being the bottom limit chosen because the bond yield dataset contains data initiated from that date, and the upper limit because it gives a margin of about one year to the end of most datasets and the dollar index set goes as far as 2020. Moreover, the COVID-19 pandemic period is excluded in that range, such a period is often ignored in financial forecasting studies, because of its temporal distinctiveness, marked by sudden fluctuations and extreme volatility, making it a difficult period to predict [46]. The resulting set can be seen in Figure 3.2, still containing a lot of missing values, caused by the different frequencies of each set, as some are daily, weekly, monthly, or quarterly distributed, posing a challenge for the goal of building a daily set.

The following step was to apply a functional method to fill the missing values between the real data, based on the literature, a highly effective method is to use interpolation, which consists of linearly estimating values between two points. Studies such as [47] show

DATE	Inflation	Interest-rates	Unemployment	VIX	Money-supply	Dollar-index	GDP	Total-debt
1990-01-02	NaN	8.54	NaN	17.240000	NaN	NaN	NaN	NaN
1990-01-03	NaN	8.37	NaN	18.190001	NaN	NaN	NaN	NaN
1990-01-04	NaN	8.29	NaN	19.219999	NaN	NaN	NaN	NaN
1990-01-05	NaN	8.20	NaN	20.110001	NaN	NaN	NaN	NaN
1990-01-06	NaN	8.20	NaN	NaN	NaN	NaN	NaN	NaN
...
2019-12-29	NaN	1.55	NaN	NaN	NaN	NaN	NaN	NaN
2019-12-30	NaN	1.55	NaN	14.820000	NaN	NaN	NaN	NaN
2019-12-31	NaN	1.55	NaN	13.780000	NaN	NaN	NaN	NaN
2020-01-01	258.682	1.55	3.5	NaN	3993.6	NaN	21538.032	23223813.0
2020-01-02	NaN	1.55	NaN	12.470000	NaN	NaN	NaN	NaN

10958 rows × 8 columns

Figure 3.2: Dataset filtered by date

the effectiveness of this method in filling in missing values, outperforming other methods in terms of accuracy, and reducing errors in financial time series forecasting. The Figure 3.3 presents the resulting dataset.

DATE	Inflation	Interest-rates	Unemployment	VIX	Money-supply	Dollar-index	GDP	Total-debt
1990-01-02	128.000000	8.54	5.300000	17.240000	798.100000	92.967	5960.028000	3.143754e+06
1990-01-03	128.000000	8.37	5.300000	18.190001	798.100000	92.967	5960.028000	3.143754e+06
1990-01-04	128.000000	8.29	5.300000	19.219999	798.100000	92.967	5960.028000	3.143754e+06
1990-01-05	128.000000	8.20	5.300000	20.110001	798.100000	92.967	5960.028000	3.143754e+06
1990-01-06	128.000000	8.20	5.300000	20.160001	798.100000	92.967	5960.028000	3.143754e+06
...
2019-12-29	258.641452	1.55	3.509677	14.356667	3994.906452	101.464	21543.526565	2.322308e+07
2019-12-30	258.654968	1.55	3.506452	14.820000	3994.470968	101.464	21541.695043	2.322333e+07
2019-12-31	258.668484	1.55	3.503226	13.780000	3994.035484	101.464	21539.863522	2.322357e+07
2020-01-01	258.682000	1.55	3.500000	13.125000	3993.600000	101.464	21538.032000	2.322381e+07
2020-01-02	258.682000	1.55	3.500000	12.470000	3993.600000	101.464	21538.032000	2.322381e+07

10958 rows × 8 columns

Figure 3.3: Interpolated dataset

Succeeding the processing of the macroeconomic datasets, a separate set containing solely yield data was filtered, leaving exclusively the maturities the present study is interested in. As shown in Figure 3.4.

A final dataset resulted from the joining of the previously cited macroeconomic dataset and the dataset containing the yields of the 5 maturities was constructed, as seen in Figure 3.5, which can be later used to simultaneously forecast all yields in a multivariate form. In order to simplify the prediction input used when predicting one yield at a time, the

	3 MO	1 YR	3 YR	5 YR	10 YR
DATE					
1990-01-02	7.83	7.81	7.90	7.87	7.94
1990-01-03	7.89	7.85	7.96	7.92	7.99
1990-01-04	7.84	7.82	7.93	7.91	7.98
1990-01-05	7.79	7.79	7.94	7.92	7.99
1990-01-08	7.79	7.81	7.95	7.92	8.02
...
2022-10-31	4.22	4.66	4.45	4.27	4.10
2022-11-01	4.23	4.75	4.48	4.27	4.07
2022-11-02	4.22	4.76	4.54	4.30	4.10
2022-11-03	4.25	4.78	4.63	4.36	4.14
2022-11-04	4.21	4.76	4.58	4.33	4.17

8219 rows × 5 columns

Figure 3.4: Dataset of yields

previous dataset was separated into 5 other sets, each one containing all macroeconomic factors and one different maturity. The Figure 3.6 shows an example of such a dataset regarding the 10 years bond.

	Inflation	Interest-rates	Unemployment	VIX	Money-supply	Dollar-index	GDP	Total-debt	3 MO	1 YR	3 YR	5 YR	10 YR
DATE													
1990-01-02	128.000000	8.54	5.300000	17.240000	798.100000	92.967	5960.028000	3.143754e+06	7.83	7.81	7.90	7.87	7.94
1990-01-03	128.000000	8.37	5.300000	18.190001	798.100000	92.967	5960.028000	3.143754e+06	7.89	7.85	7.96	7.92	7.99
1990-01-04	128.000000	8.29	5.300000	19.219999	798.100000	92.967	5960.028000	3.143754e+06	7.84	7.82	7.93	7.91	7.98
1990-01-05	128.000000	8.20	5.300000	20.110001	798.100000	92.967	5960.028000	3.143754e+06	7.79	7.79	7.94	7.92	7.99
1990-01-08	128.000000	8.27	5.300000	20.260000	798.100000	92.967	5960.028000	3.143754e+06	7.79	7.81	7.95	7.92	8.02
...
2019-12-26	258.600903	1.55	3.519355	12.650000	3996.212903	101.464	21549.021130	2.322235e+07	1.58	1.53	1.65	1.72	1.90
2019-12-27	258.614419	1.55	3.516129	13.430000	3995.777419	101.464	21547.189609	2.322259e+07	1.57	1.51	1.60	1.68	1.88
2019-12-30	258.654968	1.55	3.506452	14.820000	3994.470968	101.464	21541.695043	2.322333e+07	1.57	1.57	1.59	1.68	1.90
2019-12-31	258.668484	1.55	3.503226	13.780000	3994.035484	101.464	21539.863522	2.322357e+07	1.55	1.59	1.62	1.69	1.92
2020-01-02	258.682000	1.55	3.500000	12.470000	3993.600000	101.464	21538.032000	2.322381e+07	1.54	1.56	1.59	1.67	1.88

7503 rows × 13 columns

Figure 3.5: Joined dataset with macroeconomic datasets and yields

Once the data has been processed into the desired format, a selection of best features is necessary in order to reduce dimensionality, eliminate irrelevant or redundant features and improve models' performances. The process employed for identifying the most relevant and informative features is described as follows.

DATE	Inflation	Interest-rates	Unemployment	SAP500	Money-supply	Dollar-index	GDP	Total-debt	10 YR
1990-01-02	128.0000	8.54	5.3000	17.24	798.1000	92.967	5960.0280	3.143754e+06	7.94
1990-01-03	128.0000	8.37	5.3000	18.19	798.1000	92.967	5960.0280	3.143754e+06	7.99
1990-01-04	128.0000	8.29	5.3000	19.22	798.1000	92.967	5960.0280	3.143754e+06	7.98
1990-01-05	128.0000	8.20	5.3000	20.11	798.1000	92.967	5960.0280	3.143754e+06	7.99
1990-01-08	128.0000	8.27	5.3000	20.26	798.1000	92.967	5960.0280	3.143754e+06	8.02
...
2019-12-26	258.6009	1.55	3.5194	12.65	3996.2129	101.464	21549.0211	2.322235e+07	1.90
2019-12-27	258.6144	1.55	3.5161	13.43	3995.7774	101.464	21547.1896	2.322259e+07	1.88
2019-12-30	258.6550	1.55	3.5065	14.82	3994.4710	101.464	21541.6950	2.322333e+07	1.90
2019-12-31	258.6685	1.55	3.5032	13.78	3994.0355	101.464	21539.8635	2.322357e+07	1.92
2020-01-02	258.6820	1.55	3.5000	12.47	3993.6000	101.464	21538.0320	2.322381e+07	1.88

7503 rows × 9 columns

Figure 3.6: Example of final dataset for one yield

3.1.3 Feature selection methods

Feature selection is a critical step in forecasting studies as it involves identifying the most relevant and informative variables that are likely to impact the outcome being predicted. In essence, feature selection algorithms help identify the subset of input features most relevant to the target variable, while filtering out redundant or irrelevant ones. By doing so, they reduce the dimensionality of the data and improve the accuracy and efficiency of the forecast model. First and foremost, feature selection helps to mitigate the issue of overfitting. When there are too many input features, the forecast model can become overly complex and start to fit the noise in the data instead of the underlying patterns. This can lead to poor generalization and inaccurate predictions. It can also help to improve the interpretability of the forecast model, when there are too many input features, it can be challenging to interpret the significance and impact of each feature on the forecast output. The possible reduction of the computational burden of training a model is another reason to apply feature selection methods, saving time and resources.

Three different feature selection methods were employed in this study to analyze time-series data, applied to each yield maturity as a single target, namely: Correlation Coefficient, Mutual Information, and Cross-Correlation Function. These methods were chosen due to their simplicity and wide adoption [48], as well as for the different insights each one provides on the possible relations contained in the data.

The first method, known as Correlation Coefficient, is a statistical metric that quantifies the degree of linear dependence between two variables. In feature selection, variables that show a strong correlation, either positive or negative, with the target variable, are considered to be important [49].

Methods based on Mutual Information, such as the second method used, quantify the dependency between two variables, by measuring the reduction in uncertainty for one variable given a known value of the other variable, capturing linear or non-linear relationships [50]. Such a method provides insight into possibly relevant variables, which have high mutual information with respect to the target.

The third method employed is the Cross Correlation Function. Which is widely used for time-series data due to its capacity to measure the similarity between two series in different shifts of time, through a function of the displacement of one relative to the other, often referred to as “lag” [51]. This mechanism helps to identify temporal relations, such as if one series affects the other many days later, which can be of crucial importance in time series data where the position of data in time matters.

These three methods can provide different comprehensive approaches to identify relevant features in time-series data, considering both linear and nonlinear dependencies, as well as temporal relationships, which present three different perspectives about the relation between the data. By combining the conclusions from these methods, a more robust and accurate feature subset can be obtained, possibly reducing unnecessary complexity and enhancing the performance of the forecasting models, presented in the following section.

3.2 Forecasting Models

In the course of this study’s development, three variations of the LSTM network were built and employed, due to the data characteristics aimed to be tested on the input and output. The first model was built following a univariate format, where the network receives one input and outputs a prediction of the same input. The second model, in

contrast, has a multivariate format, meaning it takes multiple inputs and produces the same number of outputs as predictions of those inputs. Finally, the last model developed has a mixed approach based on the two other models, it receives multiple inputs and generates a single output, such an architecture aims to attempt to enhance the network's focus on that single target. In order to find the best hyperparameters to be applied in each approach, aiming to maximize the training performance, all three models passed through a carefully designed process of hyperparameter tuning based on the targeted data, better explained in the following section.

3.2.1 Hyperparameter Tuning

Hyperparameters in ML and NN models are manually configured settings that influence the behavior and performance of the models. They are not learned from the data but rather determined by the researcher or developer. Inadequate or suboptimal hyperparameter settings can lead to issues such as underfitting or overfitting, impacting the model's predictive capabilities [52].

All the models implemented in this study passed through such a tuning process, by using specialized libraries to simulate several combinations of different settings for each type of model and targeted output variable, such as each yield maturity, it was possible to determine the best configurations based on the lowest validation loss.

The hyperparameters tuned in the process include training epochs, layers, hidden dimensions, learning rate, and finally, the look-back used for the input data, which refers to the number of previous time steps that the LSTM model considers when making predictions. With such optimal hyperparameters found, the training process can leverage them to possibly achieve better predictions.

3.2.2 Training

The training process differed in epochs for each model, however the size of the training, validation and test sets remained constant throughout all processes, divided in the following proportion, 70% of the data was used for training, 10% for validation, and lastly, 20% for testing. All the models were trained using a one-step approach, receiving a determined sequence of data as input to predict the immediate next value. Later on, a multi-step approach was also employed to evaluate the models' performances on longer time horizons.

3.2.3 Univariate LSTM

A univariate LSTM model was developed and employed in this study with the purpose of being used as a benchmark model, alongside the NRW, to make a comparison with the models applying macroeconomic variables. In the case of the univariate model, it is used to compare the performance of forecasts using solely the data from the time series we want to predict, in comparison to other models employing more variables to forecast the same desired variable.

The format of such a model has a very simple design, in which a single variable is received as input, and a single output is produced, as shown in Figure 3.7. It also holds some benefits, such as a lighter learning process and the possibility to produce multi-step predictions, due to the possibility of using the single output to feed the single input cell.



Figure 3.7: Diagram of Univariate Model

3.2.4 Multi-Uni LSTM

With the aim of employing the macroeconomic dataset constructed to make yield predictions, a LSTM model capable of receiving multiple inputs was also developed. Integrating a mixed approach, with a multivariate input layer, and a univariate output layer, such a model receives many variables as input and outputs a single variable that it aims to predict, as represented in Figure 3.8.

The choice of this model's design is based on the idea of using the LSTM to capture multiple relations from the inputs with the single target, in an attempt to leverage the focus of the model by predicting a single variable, instead of aiming to predict others that are of no interest for the present study.

However, such an architecture has an impactful disadvantage when it comes to long-term forecasting, in order to make multiple predictions into the future, the model needs to use its predictions back as input, but as it does not predict all the input variables, it can not refeed itself with its predictions, creating a challenge for the effective use of the macroeconomic variables on long-term forecasting.

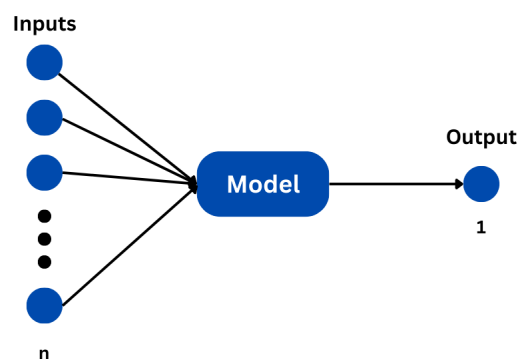


Figure 3.8: Diagram of Multivariate Input Univariate Output Model

3.2.5 Multivariate LSTM

In order to use the multiple macroeconomic factors as input to the network, but also be able to make reliable multi-step forecasting, a fully multivariate model was designed and employed, in which the input and output layers have the same number of multiple cells as portrayed in the diagram of Figure 3.9. In a multivariate LSTM model, where multiple variables are used to make predictions, it is essential to have the same number of inputs and outputs when performing multi-step forecasting. This requirement arises from the way the model utilizes an iterative process in which its predictions at each time step are fed back as inputs for the subsequent time steps, by doing so, the predicted values become part of the input sequence for the next prediction.

If the number of inputs and outputs in the model is not the same, it becomes problematic to propagate the predictions accurately over multiple time steps. Mismatched dimensions would lead to inconsistencies in the input-output relationship, which could result in unstable or unreliable predictions.

By maintaining the same number of inputs and outputs, the model ensures that the dimensions of the input and output sequences align perfectly. This alignment guarantees the proper flow of information throughout the forecasting process, allowing the model to leverage its own predictions effectively.

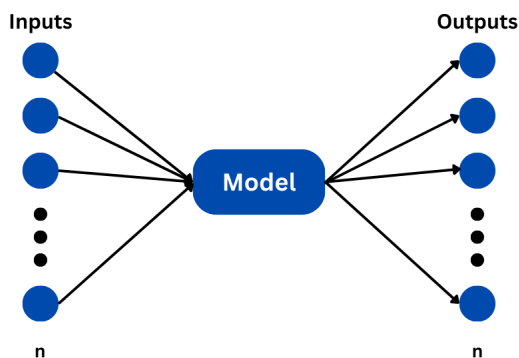


Figure 3.9: Diagram of Multivariate Model

3.3 Evaluation

The process of evaluating the models' performances was chosen based on the literature researched, most of the works related use the Root Mean Squared Error (RMSE) as the primary metric to measure the error contained in the predictions. Time-series problems are commonly evaluated using one-step and multi-step approaches. One-step forecasting involves predicting the immediate next value in the time series based on the previous data. Multi-step forecasting, on the other hand, aims to forecast multiple future values beyond the next step. By using the multi-step approach, models can provide insights into the future behavior of the time series over a longer time horizon and compare the accuracy of models on different time horizons.

3.3.1 Root Mean Squared Error

The RMSE is a commonly used metric to evaluate the accuracy of predictions in time-series problems. It measures the accuracy of the prediction by averaging the square error value of each observed error [53].

To calculate the RMSE, the differences between the predicted and actual values are squared, then the mean of these squared differences is taken, and finally, the square root of this mean is computed (Equation 3.1). In this equation, n is the number of observations in the dataset, $y_{\text{pred},i}$ represents the predicted values, and $y_{\text{actual},i}$ represents the actual values.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{\text{pred},i} - y_{\text{actual},i})^2} \quad (3.1)$$

where:

- n is the number of observations in the dataset.
- $y_{\text{pred},i}$ represents the predicted values.
- $y_{\text{actual},i}$ represents the actual values.

In time-series problems, where the goal is often to forecast future values based on historical data, the RMSE is a valuable method for evaluating the performance of predictive models. Making it possible to understand how well a model is capturing the underlying patterns and dynamics of the time series based on the error obtained.

3.3.2 One-Step Forecasting

One-Step forecasting is a technique used in time-series analysis to predict the future value of a time series based on its historical data up to a certain point. It involves making predictions for one time step ahead, meaning that the forecasted value is only for the next point in the series [54].

When applied this technique, the model is trained on the available historical data up until a specific point in time. It then uses this information to predict the value at the next time step. Once the true value at that time step becomes available, the model is updated with this new observation, and the process is repeated for the next time step.

It's important to note that one-step forecasting does not provide a long-term forecast of the time series. It focuses on predicting the immediate future, and to make predictions for multiple steps ahead, a rolling or iterative forecasting approach should be employed, where each predicted value is incorporated as an input for subsequent predictions.

3.3.3 Multi-Step Forecasting

The forecasting of multiple values into the future requires a different approach, one that should be able to reuse previous predictions, performing a series of consecutive predictions, each using the previous ones as inputs. This approach is commonly referred to as multi-step forecasting, it involves generating predictions for multiple future time steps, beyond just the next point in the time series. It requires a model that can leverage the previous predictions to generate subsequent ones, often using a recursive or iterative procedure.

Using the given historical data, the initial forecast is created for the following time step in multi-step forecasting. Once this prediction has been produced, it is added to

the sequence of inputs and used as input for the subsequent prediction. Up until the required number of steps into the future is achieved, this procedure is continued repeatedly, including the prior forecasts as part of the input into each new prediction [55].

An example of this approach is portrayed in Figure 3.10, where a univariate LSTM is forecasting a future horizon of 2 days, for the first prediction it uses the real values from the sequence of the past 3 days, therefore predicting the value for the 4th day, which is later used to predict the 5th day, alongside with day 2 and day 3, which are real values, resulting in the prediction of the next 2 days.

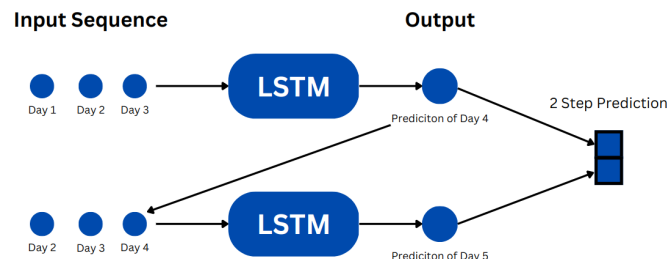


Figure 3.10: Diagram of Multi-Step Forecasting

Errors in forecasting with multiple steps have the potential to build up over time. Any errors or uncertainties in the first predictions may spread and intensify with each subsequent step because each prediction depends on the previous steps. To safeguard the reliability and accuracy of long-term forecasts, it is crucial to carefully assess potential error accumulation when performing multi-step forecasting.

Chapter 4

Results

This chapter presents the results obtained from the methods employed throughout this research, alongside an evaluation and initial discussion of such. Starting with the results from the feature selection process, followed by the Hyperparameters obtained from the tuning of each model, and finally, an assessment of the three LSTM models' performances. Each of these critical steps is addressed in separate sections to allow for a clear understanding of the methods, the rationale behind their use, and their individual contributions to the ultimate objective of the research.

The initial section, "Feature Selection" provides a meticulous review of the results stemming from the feature selection methods utilized in the research. It assesses the contribution of each feature to the models, elaborating on the reasoning behind the inclusion or exclusion of these features. The section further discusses the impact of feature selection on model complexity, performance, and interpretability.

The second section, "Hyperparameter Tuning" encapsulates the outputs of hyperparameter optimization. This section focuses on the identification and presentation of the optimal hyperparameters for each LSTM model. The section details how the hyperparameters were selected and adjusted to yield the most effective LSTM models, paving the way for improved prediction capabilities.

Finally, the third section, "Model Performance Evaluation" conducts an in-depth analysis of the performance of the LSTM models developed during the research. This study

concentrates solely on RMSE across different time horizons for each model. Through this approach, we delve into an extensive comparative analysis, establishing a comprehensive understanding of the models' effectiveness across various prediction horizons for each bond maturity and the data used.

4.1 Feature Selection

Across the process of selecting the most relevant features, three different methods were employed: correlation coefficient, mutual information, and cross-correlation function. Each method offers a unique perspective on the relationship between the features and the target variable, aiding in the identification of the most informative predictors. This section presents a detailed account of the results and interpretations of each method. The analysis in this section also unravels the importance and influence of each feature on each yield maturity time-series, guiding the decision-making process for feature inclusion or exclusion.

4.1.1 Correlation Coefficient

The first method applied had the objective of calculating the correlation coefficient between all 5 yield maturities and all the macroeconomic variables, plus the yields with themselves.

The results are displayed in the form of a heat map in the Figure 4.1. Notably, inflation shows a strong negative correlation, indicating that as inflation increases, yields tend to decrease. Conversely, interest rates exhibit a strong positive correlation, suggesting that higher interest rates values coincide with increased yields. Additionally, GDP and total debt display strong negative correlations, implying that lower GDP and higher total debt correspond to higher yields. Money supply shows a negative correlation, indicating that increased money supply leads to lower yields. These macroeconomic features, inflation, interest rates, GDP, total debt, and money supply, present the most relevance with the fluctuation of yields.

A strong positive correlation is also observed among the yields themselves as shown in the Figure 4.2. The correlations range from 0.867619 to 1.000000, indicating a high degree of association and similar trends among the different maturities. The correlation strengthens as the maturity period lengthens.

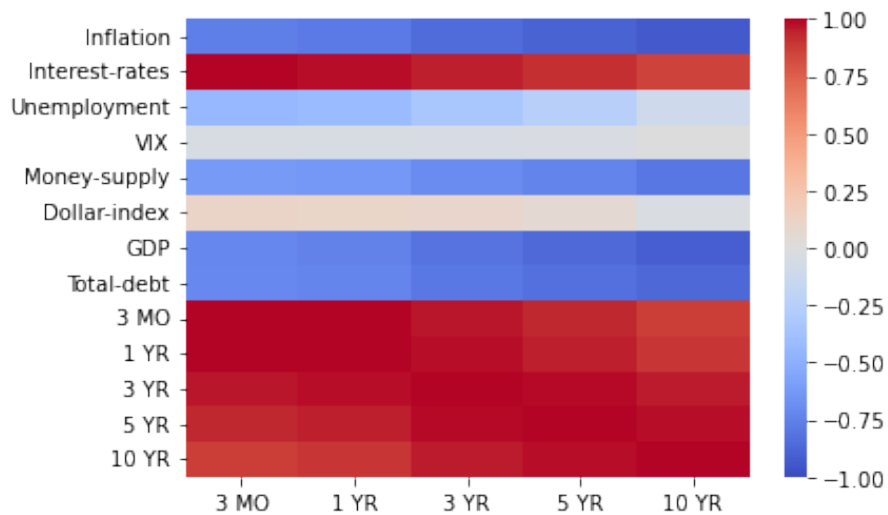


Figure 4.1: Heat-Map of all features

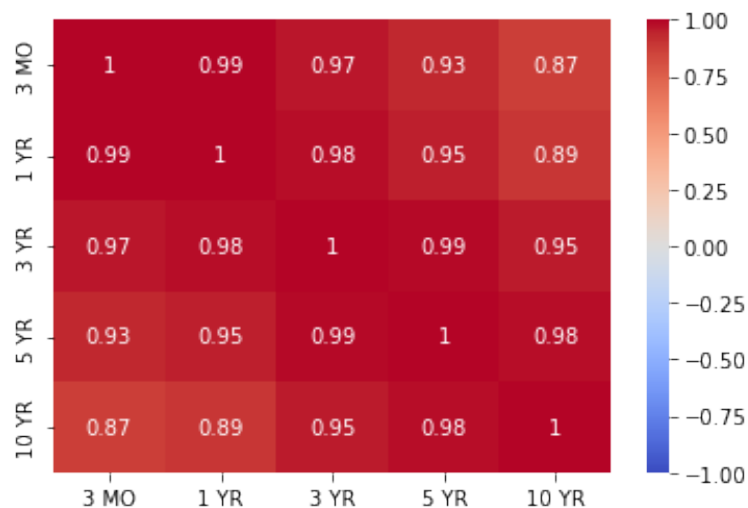


Figure 4.2: Heat-Map of Yields

4.1.2 Mutual Information

The second feature selection method was employed with the purpose of calculating the mutual information present between each macroeconomic variable and each yield maturity. The results obtained from this method are present in the Table 4.1 as follows.

Feature	3 Months	1 Year	3 Years	5 Years	10 Years
GDP	3.320000	3.278174	2.854899	2.757981	2.720223
Inflation	3.241260	3.213484	2.798508	2.701464	2.648677
Total-debt	3.222401	3.161591	2.695603	2.597220	2.590668
Money-supply	2.809193	2.710230	2.380594	2.321649	2.348859
Interest-rates	2.321670	2.087192	1.653991	1.556650	1.519810
Unemployment	2.183609	2.161362	1.865604	1.787940	1.729936
Dollar-index	1.828818	1.833650	1.589061	1.511511	1.513498
VIX	0.464405	0.429078	0.367734	0.361990	0.378420

Table 4.1: Mutual Information Results

For each yield maturity, several features show significant predictive power based on their mutual information scores. Commonly appearing among the top features are GDP, inflation, total debt, and money supply. These economic indicators consistently demonstrate a strong relationship with yield maturities.

Furthermore, other features such as unemployment, interest rates, dollar index and VIX (a stock market volatility index) also contribute to the prediction, although their impact may vary across different yield maturities. These additional factors provide complementary insights into the potential shifts in yield levels.

It is important to note that the specific rankings of features and their mutual information scores may differ across yield maturities. Nonetheless, the consistent presence of GDP, inflation, total debt, and money supply among the top features underscores their relevance in predicting all yields.

The results obtained from the feature selection methods so far show strong evidence of relations between some macroeconomic factors and yields. However, when dealing with time-series, it is important to conduct an analysis of temporal effects between series, due to the inherent sequential nature of the data. Temporal methods help identify and select those features that best predict the future based on past patterns, accounting for possible

time lags. The following method provides insight into the time-dependent relationships among series.

4.1.3 Cross Correlation Function

The last method used for selecting features was the Cross Correlation Function, Table 4.2 contains the maximum absolute correlation obtained for all macroeconomic variables and yields.

Feature	3 Months	1 Year	3 Years	5 Years	10 Years
Interest-rates	0.9929	0.9832	0.9491	0.9131	0.8515
Inflation	0.7628	0.7837	0.8560	0.8982	0.9362
Total-debt	0.7586	0.7708	0.8067	0.8337	0.8741
GDP	0.7313	0.7448	0.8181	0.8659	0.9156
Money-supply	0.7253	0.7386	0.7616	0.7704	0.7986
Unemployment	0.4488	0.4370	0.3581	0.2925	0.2256
VIX	0.3671	0.3562	0.2786	0.2172	0.1497
Dollar-index	0.2335	0.2195	0.1789	0.1275	0.1152

Table 4.2: Cross-Correlation Results for Maximum Absolute Correlation

Across all maturities, interest rates have a strong impact, signaling a fundamental influence on yield behavior. This correlation, however, seems to fade as the maturity term increases from 3 months to 10 years. As the yield lengthens, the correlation with interest rates drops from nearly perfect for 3 months at 0.9929 to a significantly lesser although substantial 0.8515 for the 10 year term.

Inflation also exhibits a critical relationship across all maturities. Interestingly, the pattern for inflation is the inverse of interest rates. With the shortest maturity of 3 months, the correlation is strong but not dominant, at 0.7628. Yet, as we move to longer maturities, the correlation intensifies, reaching a very high correlation of 0.9362 for the 10-year term. This underlines the more profound impact inflation has over longer maturities.

Total debt follows a similar pattern to inflation, with the correlation increasing with maturity length, hovering around a robust correlation of 0.75 to 0.87 across all terms.

The GDP correlation with the yield displays an intriguing shift in its behavior. Its relevance increases as we move from shorter to longer maturities, becoming second only

to inflation in the 10-year term with a correlation of 0.9156. The increasing importance of GDP for longer-term yields implies that overall economic growth impacts long-term investments more profoundly than short-term ones.

Money supply appears to be a relevant factor across all terms, with its correlation figures remaining quite stable in the 0.72-0.80 range. This illustrates the central role of monetary policy on yields regardless of their maturity.

The correlations for unemployment, VIX, and the dollar index, while present, are noticeably weaker in comparison to the other factors, with unemployment and VIX correlations decreasing as maturity lengthens. The dollar index correlation is consistently low across all maturities, hinting at a less direct impact on yields.

In summary, interest rates, inflation, total debt, GDP, and money supply emerge as the most relevant factors based on this method. Their impact varies across different maturities, revealing the nuanced landscape of yield dynamics.

4.1.4 Final Interpretation of Features

The results from the feature selection methods provide a compelling argument for the central importance of five primary features: interest rates, inflation, GDP, total debt, and money supply, across the five yields of different maturities. The consistency of these findings across three separate methods adds further robustness to the conclusion, underscoring the key role these variables may play in modeling and predicting yield behaviors.

Furthermore, a strong positive correlation was observed among the yields themselves, indicating additional relationships. Therefore different yield maturities may also be used as features to predict yields.

4.2 Hyperparameter Tuning

The hyperparameter tuning process was employed and produced results for all three types of models. Unlike the multivariate model, which predicts all five maturities at once, the univariate and multi-uni models have been designed to predict one maturity of yield at

a time, therefore five different models must be created for each one of these two types of LSTM to produce predictions for all five bond maturities considered in this study.

Each model developed (5 univariate, 5 multi-uni, 1 multivariate) has undergone hyperparameter tuning, aiming to achieve the best setting possible for each specific target, the hyperparameters considered were the number of hidden layers, dimensions of hidden layers, learning rate, number of epochs to be trained and the look-back used by the LSTM. The results obtained from this process are divided and presented throughout this section.

4.2.1 Univariate LSTM

The results obtained for the univariate models targeting different maturities are shown in Table 4.3, in which is important to outline how the best look-back values found are quite different for each maturity, for instance, when predicting the yield for 1 year bond, the best look-back found was 15 days, being a much shorter amount of time compared to the 100 days found for the 5 years bond.

Maturity	num_layers	hidden_dim	learning_rate	num_epochs	look_back
3 MO	2	32	0.0298	200	80
1 YR	2	16	0.0860	200	15
3 YR	2	32	0.0437	200	40
5 YR	1	16	0.0397	250	100
10 YR	1	64	0.0691	250	40

Table 4.3: Hyperparameters obtained for the Univariate Models

4.2.2 Multi-Uni LSTM

The hyperparameters found for the Multi-Uni models presented in Table 4.4 show more consistent values for all maturities than the ones found for univariate models, specially considering the look-back values.

Maturity	num_layers	hidden_dim	learning_rate	num_epochs	look_back
3 MO	1	64	0.0228	300	20
1 YR	2	32	0.0284	300	80
3 YR	1	32	0.0730	250	40
5 YR	2	16	0.0523	300	40
10 YR	2	32	0.0071	300	20

Table 4.4: Hyperparameters obtained for the Multi-Uni Models

4.2.3 Multivariate LSTM

The multivariate model, as mentioned before, has the capacity to predict the yields of all maturities at once, therefore it was necessary to find the best hyperparameters for just one model of its kind. The results obtained are displayed in Table 4.5, the best settings found represent a simple architecture with only one hidden layer with 64 nodes and a moderate value of look-back, accounting for 40 days.

num_layers	hidden_dim	learning_rate	num_epochs	look_back
1	64	0.0170	250	40

Table 4.5: Hyperparameters obtained for the Multivariate Model

4.3 Models Performance Evaluation

The final evaluation of the developed models was made using the best hyperparameters configuration found for each. The testing dataset used was produced by selecting the final 20% of the full dataset. Table 4.6 presents a comparison of performance, in the form of RMSE, obtained by each model, as well as the NRW benchmark, doing one-step predictions, therefore predicting at each step only one value into the future with real data from the previous values. As observed in Table 4.6 none of the developed models was able to outperform the NRW algorithm, commonly the case found in the literature when forecasting only one future value, as such benchmark predicts each future value as being the same as the previous one, which is usually very close.

Although the univariate model could not outperform the NRW predicting one future

value, it got very close to the benchmark, performing the best among the LSTM models, nearly perfectly fitting the testing dataset, as shown in Figure 4.3, in which a plot from the one-step forecasting of the yield from the 10 YR Bond maturity is presented.

Maturity	Uni	Multi	Multi_Uni	NRW
3 MO	0.0436	0.1074	0.2256	0.0192
1 YR	0.0362	0.8713	0.3115	0.0210
3 YR	0.0562	0.6816	0.0872	0.0374
5 YR	0.0640	0.2206	0.1435	0.0422
10 YR	0.0667	1.4291	0.4478	0.0423

Table 4.6: RMSE Results from One-step forecasting on the Testing Dataset

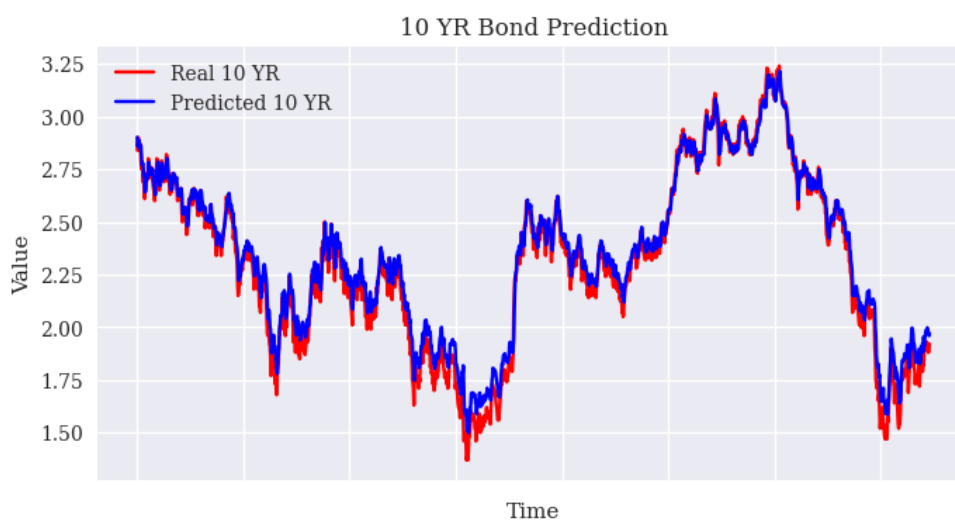


Figure 4.3: Univariate LSTM performance One-step forecasting the 10 YR Bond

The multivariate model has shown the capacity of predicting some targets quite precisely and others very far from the real curve, as shown in Figure 4.4 portraying the real curves in red and the predicted in blue, in which is possible to observe its good performance in one-step forecasting the 3 YR and 5 YR Bonds yield, and lacking performance for the other targets. Such a behavior could be explained by the large number of different target variables dealt with by the model, as well as the difference in scale between the variables, although all being normalized during the implementation, the ones with the largest scale difference, are also the ones with worst performance, such as the Money-supply and GDP.

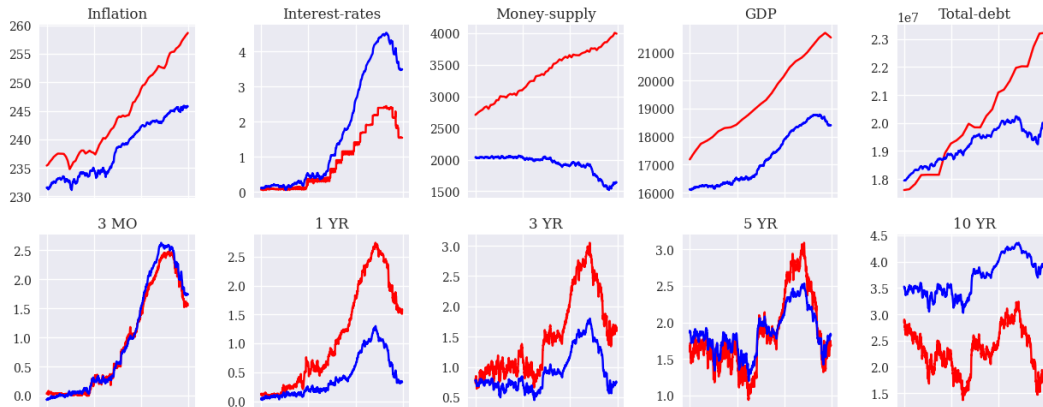


Figure 4.4: Multivariate LSTM performance One-step forecasting

The LSTM model with multivariate input and univariate output was developed with the purpose of better focusing the yields as output, instead of the macroeconomic variables. Such an implementation shows indeed a better performance for most maturities than the fully multivariate model as seen in Table 4.6. The best one-step forecasting performance by this model has been the 3 YR Bond yield, presented in Figure 4.5.



Figure 4.5: Multi-Uni LSTM performance One-step forecasting the 3 YR Bond

Following the one-step forecasting evaluation, the univariate and multivariate models, which are suited for multi-step forecasting, were evaluated by performing multi-step predictions of horizons of 5, 10, 20, 30, 60, 120 and 240 values into the future. The results

are divided into five different tables, 4.7a, 4.7b, 4.7c, 4.7d, 4.7e, each one regarding one bond maturity as follows.

Horizon	Uni	Multi	NRW
5	0.0382	0.0463	0.0288
10	0.0700	0.0527	0.0402
20	0.1111	0.0690	0.0550
30	0.1070	0.0801	0.0729
60	0.1243	0.0891	0.1205
120	0.1542	0.0978	0.1871
240	0.3550	0.1046	0.3316

(a) 3 Months Bond

Horizon	Uni	Multi	NRW
5	0.0194	0.0382	0.0358
10	0.0498	0.0528	0.0501
20	0.0489	0.0785	0.0635
30	0.1206	0.0941	0.0891
60	0.1491	0.1088	0.1241
120	0.2351	0.1257	0.2210
240	0.4699	0.1226	0.3971

(b) 1 Year Bond

Horizon	Uni	Multi	NRW
5	0.0574	0.1835	0.0632
10	0.0453	0.3045	0.0864
20	0.0764	0.4664	0.1115
30	0.1329	0.5042	0.1433
60	0.1183	0.4730	0.1839
120	0.2686	0.4127	0.3010
240	0.7436	0.3377	0.5144

(c) 3 Years Bond

Horizon	Uni	Multi	NRW
5	0.0578	0.4554	0.0723
10	0.1725	0.5770	0.1000
20	0.2213	0.7307	0.1300
30	0.1986	0.7477	0.1643
60	0.1710	0.6855	0.2053
120	0.2394	0.6146	0.3209
240	0.7170	0.5712	0.5084

(d) 5 Years Bond

Horizon	Uni	Multi	NRW
5	0.0399	0.7645	0.0729
10	0.0655	0.8406	0.1016
20	0.1300	0.9214	0.1302
30	0.1215	0.9072	0.1686
60	0.1161	0.8598	0.2158
120	0.1671	0.8583	0.3241
240	0.3273	0.9676	0.5057

(e) 10 Years Bond

Table 4.7: RMSE Results from Multi-Step Forecasting

The results from the multi-step forecasts show that the LSTM models can outperform the NRW benchmark model in many different horizons and different bond maturities. The LSTM results look good compared to the NRW, however, it does not reflect a very accurate model for multi-step predictions, as the NRW model is not a great benchmark for predicting multiple steps, although it is often used.

Both LSTM models showed a tendency of flattening multi-step predictions, outputting

very similar values at each iteration, as shown in Figure 4.6, where the univariate model seems to correctly predict the downward trend of the values, outperforming the benchmark model. The multivariate multi-step predictions for the 10 YR Bond yield, displayed in Figure 4.7, show the same effect of flattening its own predictions, however, this model is much more inaccurate than the previous.

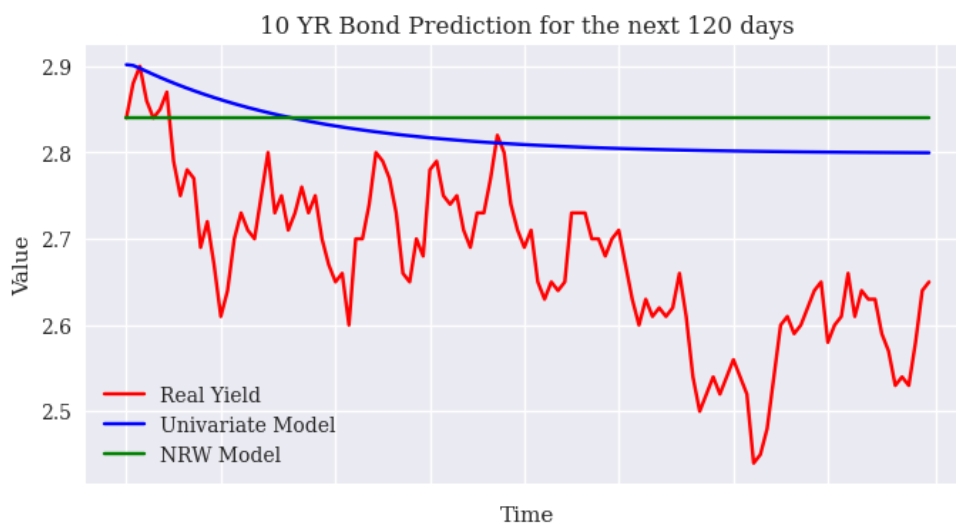


Figure 4.6: 120 days Multi-Step Prediction of the 10 YR Bond Yield by Univariate Model



Figure 4.7: 120 days Multi-Step Prediction of the 10 YR Bond Yield by Multivariate Model

Chapter 5

Conclusion and Discussion

The main objectives of this thesis included the construction of a usable and thorough macroeconomic dataset, the meticulous identification and curation of optimal features within this dataset, the design and application of tailored LSTM models suited to the dataset's characteristics, and the thorough evaluation of these models' predictive accuracy in the context of forecasting bond yields. As such, this thesis not only sought to contribute to the realm of macroeconomics but also to the broader understanding of predictive NN modeling in financial markets, alongside their capabilities and limitations. This work focused on answering three main research questions, which were answered in the context of the current methodology employed, although the answers could be different if other methods were applied, which will be discussed in this chapter as proposals for further work on the same problem.

The first research question regarded the selection of features within the macroeconomic dataset built, this question was addressed through the utilization of feature selection algorithms, ultimately identifying five most influential variables directly impacting the desired targets. The subsequent question aimed to ascertain whether the model proposed with the constructed dataset could outperform a model using solely the target's time-series data, the findings, with the applied data and methodology, indicate that the proposed model did not outperform the time-series-based counterpart. Finally, the third research question sought to conduct a comparative assessment of the performance of the proposed

model in relation to the benchmark utilized within the existing literature, the obtained results demonstrate a superior accuracy over certain time horizons. However, the outcomes unveil a pronounced limitation in conducting multi-step predictions utilizing LSTMs, as implemented within this thesis. This constraint has prompted the proposal of an alternative approach for forecasting multiple future values, thereby laying the groundwork for prospective investigations in addressing this particular challenge.

As the foundational work for the implementation of the predictive models, a custom dataset aiming for robustness was built, encompassing a diverse array of macroeconomic indicators. This dataset emerged as a fusion of disparate time series, each characterized by distinct date ranges, frequencies, and scales. The main challenge in handling this data pertained to the strategic determination of a uniform frequency and the rectification of missing values, a task accomplished through the use of interpolation techniques. It is noteworthy to underscore that the chosen methodology for data treatment could have led to divergent outcomes, thereby highlighting the significance of acknowledging the potential influence of data processing choices. Moreover, it is imperative to recognize that the nature of a custom-made dataset introduces the possibility that the derived patterns might not encapsulate the most relevant patterns to the modeling of financial time series, inherently recognized for their volatility and stochastic nature. Following the construction of the dataset, it passed by a process of feature selection to determine the most influential features across all the bond maturities. Leveraging the application of three distinct techniques, converging to coherent results that indicate the influence of five main features, which were selected from the interpretation of such an analysis. The analysis also highlighted the influence of the bond yields time series between each other, therefore bonds with different maturities demonstrate strong relations among them.

Taking into account the dataset format, three LSTM models were developed to suit the input and output formats required. The first one was a univariate model developed to produce predictions based only on the individual time series of each target. A second model with a multivariate input layer and a univariate output layer was developed to predict a single target based on multiple input variables, an architecture ultimately to

be found unsuitable for multi-step predictions as employed in this study, because of the different formats of input and output layers. Therefore the necessity of developing a third fully multivariate model that could handle its own outputs as part of subsequent inputs. The implementation of all three models underwent a process of hyperparameter tuning within a configuration space constrained by the available computational resources. This constraint, while ensuring operational feasibility, also limited the exploration of more intricate and complex model architectures. The results from such a process determined the best configuration found to be used for each type of model with each one of the five targeted bond maturities. The training processes were done using the hyperparameters found, including the look back sequence length used by the LSTMs, as so to produce the prediction of one next value.

The final evaluation of all models as in comparison with the NRW benchmark was done by applying each model to a test set comprising 20% of the full dataset, the evaluation was divided into two parts, the first regarding the performances of one-step predictions, and the second taking account the multi-step predictions of a series of values by refeeding the model with its own predictions. The findings obtained from the results based on the evaluation of one-step predictions show that none of the models was able to ultimately outperform the benchmark in any of the bond maturities, which is known to have a low error when predicting the immediate next value, considering it is often very close to the last one, favoring the benchmark performance when dealing with very short future horizons, such as one value. However, the developed models did perform well in some instances, coming very close to the benchmark, especially the univariate LSTM, with the best performance obtained among the LSTMs, showing that it still performs better when using only the target's time series data than the ones using the macroeconomic dataset. The model with a multivariate input and univariate output has proven to be overall superior to the multivariate model, quite correctly fitting some bond yields curves, such as the 3 Months Bond, which supports the theory that it could focus better on the desired target, instead of computing predictions of undesired targets as the multivariate model does.

The multi-step predictions as implemented by the method employed in this study, were not possible to be computed for one of the LSTM models, the model referred to in the work as Multi-Uni, due to the nature of the mechanism of resuing the outputs for subsequent inputs, which prevented the analysis of such a model's performance on longer time horizons. However, the other two LSTM models were evaluated predicting a series of multiple values. The results show that both the univariate and multivariate models were able to outperform the error obtained by the NRW benchmark model in multiple different instances of time horizons and bond maturities, although the univariate model has obtained better results in most instances. While the diminished errors when compared to the benchmark model might suggest an adept predictive model for multiple values, this outcome does not necessarily indicate a precise alignment with the underlying curve when visualized graphically. Notably, the LSTM models tend to generate outputs that exhibit a higher degree of resemblance to one another at each iteration. Over time, this results in the formation of a progressively smooth prediction curve. This apparent smoothness could signify either a challenge in the model's capacity to effectively learn patterns from the data or its potential to forecast a specific underlying trend. The models' inability to capture potential future volatility undermines their efficacy for long-term forecasting. This observed limitation may stem from inherent characteristics of the utilized data, potentially arising from the scarcity of examples that were insufficient to adequately train the models. Additionally, the intricate challenge of identifying pertinent patterns in financial time series data, renowned for their stochastic tendencies, further compounds the complexity of the challenge.

Building upon the foundation established within this thesis and considering the insights from the achieved results, it becomes evident that alternative methodologies could have been employed, potentially yielding divergent outcomes. This pertains to every stage of the study, spanning from data preprocessing to model training, as well as the evaluation of multi-step predictions. The findings unveiled through this investigation highlight both the promise and the limitations of applying LSTMs to the realm of finance. While the study underscores the LSTMs' potential in generating accurate one-step predictions, it

equally underscores their inherent inadequacy in forecasting longer-term financial series. Considering the forecasting of multiple values as the hardest challenge found in this thesis, the main proposal for further work on this problem is the implementation of a different training methodology for the models, which were trained to predict one value at a time, requiring the reuse of outputs for multi-step predictions. Conversely, if the models are trained to predict multiple values at once for a given target by employing a multivariate output layer to represent various future values of that same target, they might yield distinct outcomes capable of encapsulating the inherent volatility of the time series. This approach would also make possible the forecasting of multiple future values utilizing a model structured with a range of macroeconomic variables as inputs and a single target in the output, which was not possible using the method applied in the present study.

In summary, this thesis has provided valuable insights at the intersection of predictive NN models and macroeconomics. The conducted evaluation highlights the potential of LSTMs in short-term bond yield forecasting, while also revealing limitations in long-term predictions. Notably, this study has generated a robust macroeconomic dataset and devised well-suited LSTM models tailored to it. All the algorithms, trained models, and data used are publicly accessible for further exploration at https://gitlab.estig.ipb.pt/mi_matheus/nn-forecast. The assessment of predictive models on this constructed dataset could potentially pave the way for future advancements in financial forecasting models, particularly in the convergence of macroeconomics and deep neural networks.

Bibliography

- [1] SIFMA. “Us treasury securities statistics - sifma.” (), [Online]. Available: <https://www.sifma.org/resources/research/us-treasury-securities-statistics/>. accessed: 30.10.2022.
- [2] J.-B. Hasse and Q. Lajaunie, “Does the yield curve signal recessions? new evidence from an international panel data analysis,” *The Quarterly Review of Economics and Finance*, vol. 84, pp. 9–22, 2022, ISSN: 1062-9769. DOI: <https://doi.org/10.1016/j.qref.2022.01.001>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1062976922000011>.
- [3] M. Castellani and E. A. d. Santos, “Forecasting long-term government bond yields: An application of statistical and ai models,” 2006. [Online]. Available: <https://depeco.iseg.ulisboa.pt/wp/wp042006.pdf>.
- [4] F. Diebold, G. Rudebusch, and S. Aruoba, “The macroeconomy and the yield curve: A dynamic latent factor approach,” *SSRN Electronic Journal*, Jan. 2004. DOI: 10.2139/ssrn.562503.
- [5] Investor.gov. “Us securities and exchange commission.” (), [Online]. Available: <https://www.investor.gov/introduction-investing/investing-basics/investment-products/bonds-or-fixed-income-products/bonds>. accessed: 05.12.2022.
- [6] D. R. McKay and D. A. Peters, “Bond basics and the yield curve,” *Plastic Surgery*, vol. 27, no. 1, pp. 83–85, 2019. DOI: 10.1177/2292550319830964. eprint: <https://doi.org/10.1177/2292550319830964>. [Online]. Available: <https://doi.org/10.1177/2292550319830964>.

- [7] C. Zaloom, “How to Read the Future: The Yield Curve, Affect, and Financial Prediction,” *Public Culture*, vol. 21, no. 2, pp. 245–268, May 2009, ISSN: 0899-2363. DOI: 10.1215/08992363-2008-028. eprint: https://read.dukeupress.edu/public-culture/article-pdf/21/2/245/488573/PC21.2\05_Zaloom_sh.pdf. [Online]. Available: <https://doi.org/10.1215/08992363-2008-028>.
- [8] D. Andolfatto and A. Spewak, “Does the yield curve really forecast recession?” *Economic Synopses*, no. 30, 2018. DOI: 10.20955/es.2018.30. [Online]. Available: <https://doi.org/10.20955/es.2018.30>.
- [9] R. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, English, 2nd. Australia: OTexts, 2018.
- [10] F. X. Diebold and R. S. Mariano, “Comparing predictive accuracy,” *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 253–263, 1995. DOI: 10.1080/07350015.1995.10524599. [Online]. Available: <https://www.tandfonline.com/doi/abs/10.1080/07350015.1995.10524599>.
- [11] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The m4 competition: 100,000 time series and 61 forecasting methods,” *International Journal of Forecasting*, vol. 36, no. 1, pp. 54–74, 2020, M4 Competition, ISSN: 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2019.04.014>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169207019301128>.
- [12] L. Medsker and L. C. Jain, Eds., *Recurrent Neural Networks: Design and Applications*, 1st. CRC Press, 1999. DOI: 10.1201/9781003040620. [Online]. Available: <https://www.taylorfrancis.com/books/edit/10.1201/9781003040620/recurrent-neural-networks-lakhmi-jain-larry-medsker>.
- [13] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735. eprint: <https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf>. [Online]. Available: <https://doi.org/10.1162/neco.1997.9.8.1735>.

- [14] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, Mar. 1994. DOI: 10.1109/72.279181.
- [15] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to Forget: Continual Prediction with LSTM,” *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000, ISSN: 0899-7667. DOI: 10.1162/089976600300015015. eprint: <https://direct.mit.edu/neco/article-pdf/12/10/2451/814643/089976600300015015.pdf>. [Online]. Available: <https://doi.org/10.1162/089976600300015015>.
- [16] T. Fischer and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions,” *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018, ISSN: 0377-2217. DOI: <https://doi.org/10.1016/j.ejor.2017.11.054>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0377221717310652>.
- [17] C. R. Nelson and A. F. Siegel, “Parsimonious Modeling of Yield Curves,” *The Journal of Business*, vol. 60, no. 4, pp. 473–489, Oct. 1987. DOI: 10.1086/296409. [Online]. Available: <https://www.jstor.org/stable/2352957>.
- [18] F. X. Diebold and C. Li, “Forecasting the term structure of government bond yields,” *Journal of Econometrics*, vol. 130, no. 2, pp. 337–364, 2006, ISSN: 0304-4076. DOI: <https://doi.org/10.1016/j.jeconom.2005.03.005>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0304407605000795>.
- [19] G. R. Duffee, “Term premia and interest rate forecasts in affine models,” *The Journal of Finance*, vol. 57, no. 1, pp. 405–443, 2002. DOI: <https://doi.org/10.1111/1540-6261.00426>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6261.00426>.
- [20] M. Guidolin and D. L. Thornton, “Predictions of short-term rates and the expectations hypothesis,” *International Journal of Forecasting*, vol. 34, no. 4, pp. 636–664, 2018, ISSN: 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2018.03>

- .006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169207018300517>.
- [21] A. Ang and M. Piazzesi, “A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables,” *Journal of Monetary Economics*, vol. 50, no. 4, pp. 745–787, 2003, ISSN: 0304-3932. DOI: [https://doi.org/10.1016/S0304-3932\(03\)00032-1](https://doi.org/10.1016/S0304-3932(03)00032-1). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0304393203000321>.
- [22] W. Lv and J. Qi, “Stock market return predictability: A combination forecast perspective,” *International Review of Financial Analysis*, vol. 84, p. 102376, 2022, ISSN: 1057-5219. DOI: <https://doi.org/10.1016/j.irfa.2022.102376>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S105752192200326X>.
- [23] S. J. Binoy and J. Jos, “Financial market forecasting using macro-economic variables and rnn,” in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Apr. 2022, pp. 1366–1371. DOI: 10.1109/ICACITE53722.2022.9823905.
- [24] Y. Song, X. Tang, H. Wang, and Z. Ma, “Volatility forecasting for stock market incorporating macroeconomic variables based on garch-midas and deep learning models,” *Journal of Forecasting*, vol. 42, no. 1, pp. 51–59, 2023. DOI: <https://doi.org/10.1002/for.2899>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2899>.
- [25] S. Poledna, M. G. Miess, C. Hommes, and K. Rabitsch, “Economic forecasting with an agent-based model,” *European Economic Review*, vol. 151, p. 104306, 2023, ISSN: 0014-2921. DOI: <https://doi.org/10.1016/j.euroecorev.2022.104306>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0014292122001891>.

- [26] C. Glocker and S. Kaniovski, “Macroeconometric forecasting using a cluster of dynamic factor models,” *Empirical Economics*, vol. 63, no. 1, pp. 43–91, Jul. 2022, ISSN: 1435-8921. DOI: 10.1007/s00181-021-02129-w. [Online]. Available: <https://doi.org/10.1007/s00181-021-02129-w>.
- [27] A. Alamsyah and M. F. Permana, “Artificial neural network for predicting indonesian economic growth using macroeconomics indicators,” in *2018 International Symposium on Advanced Intelligent Informatics (SAIN)*, Aug. 2018, pp. 15–19. DOI: 10.1109/SAIN.2018.8673347.
- [28] K. L. Seip and D. Zhang, “The GDP, the US treasury yield and the federal funds rate: who follows whom, when and why?” *Journal of Financial Economic Policy*, vol. 14, no. 2, pp. 187–206, Oct. 2021. DOI: 10.1108/JFEP-11-2020-0241.
- [29] G. Grudnitski and L. Osburn, “Forecasting s&p and gold futures prices: An application of neural networks,” *Journal of Futures Markets*, vol. 13, no. 6, pp. 631–643, 1993. DOI: <https://doi.org/10.1002/fut.3990130605>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/fut.3990130605>.
- [30] A. Chudziak, “Predictability of stock returns using neural networks: Elusive in the long term,” *Expert Systems with Applications*, vol. 213, p. 119 203, 2023, ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2022.119203>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417422022217>.
- [31] C. Wang, Y. Chen, S. Zhang, and Q. Zhang, “Stock market index prediction using deep transformer model,” *Expert Systems with Applications*, vol. 208, p. 118 128, 2022, ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2022.118128>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417422013100>.
- [32] L. Di Persio and O. Honchar, “Artificial neural networks architectures for stock price prediction: Comparisons and applications,” *International journal of circuits*,

- systems and signal processing*, vol. 10, no. 2016, pp. 403–413, 2016. [Online]. Available: <https://www.naun.org/main/NAUN/circuitssystemssignal/2016/b482005-303.pdf>.
- [33] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, “Stock market prediction based on generative adversarial network,” *Procedia Computer Science*, vol. 147, pp. 400–406, 2019, 2018 International Conference on Identification, Information and Knowledge in the Internet of Things, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2019.01.256>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050919302789>.
- [34] Y. Li, P. Ni, and V. Chang, “Application of deep reinforcement learning in stock trading strategies and stock forecasting,” *Computing*, vol. 102, no. 6, pp. 1305–1322, Jun. 2020, ISSN: 1436-5057. DOI: [10.1007/s00607-019-00773-w](https://doi.org/10.1007/s00607-019-00773-w). [Online]. Available: <https://doi.org/10.1007/s00607-019-00773-w>.
- [35] Y. Lin, Z. Lin, Y. Liao, Y. Li, J. Xu, and Y. Yan, “Forecasting the realized volatility of stock price index: A hybrid model integrating ceemdan and lstm,” *Expert Systems with Applications*, vol. 206, p. 117736, 2022, ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2022.117736>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S095741742201017X>.
- [36] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, “Stock price prediction using lstm, rnn and cnn-sliding window model,” in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 1643–1647. DOI: [10.1109/ICACCI.2017.8126078](https://doi.org/10.1109/ICACCI.2017.8126078).
- [37] M. Nunes, E. Gerding, F. McGroarty, and M. Niranjana, “A comparison of multitask and single task learning with artificial neural networks for yield curve forecasting,” *Expert Systems with Applications*, vol. 119, pp. 362–375, Apr. 2019, ISSN: 0957-4174. DOI: [10.1016/j.eswa.2018.11.012](https://doi.org/10.1016/j.eswa.2018.11.012). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417418307322> (visited on 11/15/2022).

- [38] O. Castello and M. Resta, “Modeling the yield curve of brics countries: Parametric vs. machine learning techniques,” *Risks*, vol. 10, no. 2, p. 36, Feb. 2022, ISSN: 2227-9091. DOI: 10.3390/risks10020036. [Online]. Available: <http://dx.doi.org/10.3390/risks10020036>.
- [39] C. Wegener, C. von Spreckelsen, T. Basse, and H.-J. von Mettenheim, “Forecasting government bond yields with neural networks considering cointegration,” *Journal of Forecasting*, vol. 35, no. 1, pp. 86–92, 2016. DOI: <https://doi.org/10.1002/for.2385>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2385>.
- [40] P. C. Kauffmann, H. H. Takada, A. T. Terada, and J. M. Stern, “Learning forecast-efficient yield curve factor decompositions with neural networks,” *Econometrics*, vol. 10, no. 2, p. 15, Mar. 2022, ISSN: 2225-1146. DOI: 10.3390/econometrics10020015. [Online]. Available: <http://dx.doi.org/10.3390/econometrics10020015>.
- [41] A. Eklind, “Deep neural networks to ensure the quality of calculated yield curves in banking,” M.S. thesis, KTH, School of Electrical Engineering and Computer Science (EECS), 2020, p. 60.
- [42] P. Veronesi, *Handbook of fixed-income securities*. John Wiley & Sons, 2016.
- [43] N. G. Mankiw, *Principles of macroeconomics*. Cengage learning, 2020.
- [44] G. Peersman, “Macroeconomic effects of unconventional monetary policy in the euro area,” *ECB Working Paper*, no. 1397, Oct. 2011. DOI: 10.2139/ssrn.1950039. [Online]. Available: <https://ssrn.com/abstract=1950039>.
- [45] B. Eichengreen, *Exorbitant privilege: The rise and fall of the dollar and the future of the international monetary system*. Oxford University Press, 2011.
- [46] C. C. Ngwakwe *et al.*, “Effect of covid-19 pandemic on global stock market values: A differential analysis,” *Acta Universitatis Danubius. (Economica)*, vol. 16, no. 2, pp. 255–269, 2020. [Online]. Available: <https://dj.univ-danubius.ro/index.php/AUDOE/article/view/258>.

- [47] C. John, E. Ekpenyong, and C. Nworu, “Imputation of missing values in economic and financial time series data using five principal component analysis (pca) approaches,” *Central Bank of Nigeria Journal of Applied Statistics*, pp. 51–73, Aug. 2019. DOI: 10.33429/Cjas.10119.3/6.
- [48] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014, 40th-year commemorative issue, ISSN: 0045-7906. DOI: <https://doi.org/10.1016/j.compeleceng.2013.11.024>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0045790613003066>.
- [49] A. G. Asuero, A. Sayago, and A. G. González, “The correlation coefficient: An overview,” *Critical Reviews in Analytical Chemistry*, vol. 36, no. 1, pp. 41–59, 2006. DOI: 10.1080/10408340500526766. [Online]. Available: <https://doi.org/10.1080/10408340500526766>.
- [50] J. R. Vergara and P. A. Estévez, “A review of feature selection methods based on mutual information,” *Neural Computing and Applications*, vol. 24, no. 1, pp. 175–186, Jan. 2014. DOI: 10.1007/s00521-013-1368-0. [Online]. Available: <https://doi.org/10.1007/s00521-013-1368-0>.
- [51] A. E. Usoro *et al.*, “Some basic properties of cross-correlation functions of n-dimensional vector time series,” *Journal of Statistical and Econometric Methods*, vol. 4, no. 1, pp. 63–71, 2015. [Online]. Available: http://www.scienpress.com/Upload/JSEM%20FVol%204_1_3.pdf.
- [52] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, “Algorithms for hyper-parameter optimization,” in *Advances in Neural Information Processing Systems*, J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, Eds., vol. 24, Curran Associates, Inc., 2011. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf.
- [53] C. M. Castilho, “Forecasting time series based on features,” 2020. [Online]. Available: <https://hdl.handle.net/10216/141197>.

- [54] T. Limouni, R. Yaagoubi, K. Bouziane, K. Guissi, and E. H. Baali, “Accurate one step and multistep forecasting of very short-term pv power using lstm-tcn model,” *Renewable Energy*, vol. 205, pp. 1010–1024, 2023, ISSN: 0960-1481. DOI: <https://doi.org/10.1016/j.renene.2023.01.118>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S096014812300143X>.
- [55] S. Suradhaniwar, S. Kar, S. S. Durbha, and A. Jagarlapudi, “Time series forecasting of univariate agrometeorological data: A comparative performance evaluation via one-step and multi-step ahead forecasting strategies,” *Sensors*, vol. 21, no. 7, 2021, ISSN: 1424-8220. DOI: 10.3390/s21072430. [Online]. Available: <https://www.mdpi.com/1424-8220/21/7/2430>.