Predicting Business Cycles with Linear and Non-linear Filters

Beatriz Abrantes

Dissertation written under the supervision of Professor Pedro Afonso Fernandes.

Dissertation submitted in partial fulfilment of requirements for the MSc in Business Analytics, at the Universidade Católica Portuguesa, 03-01-2023.

Predicting Business Cycles with Linear and Non-linear Filters

Beatriz Loureiro Abrantes

January 3, 2023

Abstract

Business cycles represent the short-run fluctuations in economies and have a non-recurring periodic character that makes them difficult to forecast. This dissertation focuses on the cycle-trend decomposition techniques that are used to remove the long-run component and thus obtain the cyclical component of macroeconomic series. Statistical filters can be used for this purpose, and through them, this work aims to clarify and visualize the cycle-trend decomposition. The primary objective of this dissertation is to evaluate the performance of two types of filters, linear and non-linear. At the end, it is also expected that conclusions will be drawn about the tool used throughout this work, Power BI. After comparing the linear filter developed by Hodrick and Prescott (1997) with two non-linear filters, MR filter and median filter developed by Mosheiov and Raveh (1997) and Wen and Zeng (1999), respectively, the results obtained were favorable compared to the non-linear filter. The MR filter proved to be able to produce a more robust trend than the others and to identify economic periods in a natural way. The MED filter proved to be able to produce less volatile and noisy cyclical components than the others; this is due to its ability to capture sharp changes in the trend and suppress them in the cyclical component. This concluded that the nonlinear filters performed well against the linear filter under study. Power BI demonstrated throughout the work several capabilities that characterize it as a good Business Intelligence tool, however, with room for improvement.

Keywords: business cycles, linear filters, non-linear filters, trend-cycle decomposition, time series

Resumo

Os ciclos económicos representam as flutuações de curto prazo nas economias e têm um carácter periódico não recorrente que os torna difíceis de prever. Esta dissertação centra-se nas técnicas de decomposição ciclo-tendência, vulgarmente utilizadas para remover a componente de longo-prazo e assim obter a componente cíclica das séries macroeconómicas. Podem ser utilizados filtros estatísticos para este fim, e através dos mesmos, este trabalho visa clarificar e visualizar a decomposição ciclo-tendência. O principal objetivo desta dissertação é avaliar o desempenho de dois tipos de filtros, lineares e não lineares. No final, espera-se também que sejam tiradas conclusões sobre a ferramenta utilizada ao longo deste trabalho, o Power BI. Após comparar o filtro linear desenvolvido por Hodrick and Prescott (1997) com dois filtros não lineares, o filtro MR e o filtro mediano desenvolvido por Mosheiov and Raveh (1997) e por Wen and Zeng (1999), respetivamente, os resultados foram favoráveis para os filtros não lineares. O filtro MR provou ser capaz de produzir uma tendência mais robusta que os outros e de identificar períodos económicos de forma natural. O filtro mediano provou ser capaz de produzir componentes cíclicas menos voláteis e ruidosas do que os restantes; isto deve-se à sua capacidade de captar mudanças bruscas na tendência e de as suprimir da componente cíclica. Concluiu-se que os filtros não lineares tiveram um bom desempenho face ao filtro linear em estudo. O Power BI demonstrou várias capacidades que o caracterizam como uma boa ferramenta de Business Intelligence, no entanto, com espaço para melhorias.

Palavras-chave: ciclos económicos, filtros lineares, filtros não lineares, decomposição ciclo-tendência, séries temporais

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Pedro Afonso Fernandes, for his invaluable advice and support throughout the research and writing processes. His expertise and encouragement were essential in helping me complete this dissertation. I would also like to thank my family for their constant love and guidance. Finally, I want to extend my sincere thanks to my friends for their constant support and to my boyfriend, in particular, for his unwavering belief in me and understanding during this demanding period.

Contents

List of Figures

List of Tables

Acronyms

BP Band-Pass. 10

GDP Gross Domestic Product. 5, 6, 10, 12–14, 24–26, 28, 31

HP Hodrick-Prescott. 2, 8–10, 16, 25–29

MA Moving Average. 10

MAE Mean Absolute Error. 20

MED Median. 24, 28, 29, I

MR Mosheiov-Raveh. 11, 24–29, I, II

NBER National Bureau of Economic Analysis. 10, 14, 28

RMSE Root Mean Squared Error. 20

1 Introduction

Economic cycle theories are concerned with the short-term fluctuations in the context of a long-term growth trajectory and intend to investigate the causes of these changes as well as measures to mitigate their negative repercussions. These fluctuations are generally referred to as cycles.

The image that comes to mind while discussing cycles is a sine wave with its repeated and regular shape. However, the concept of cycle in economics and other disciplines refers to a broader idea. It was described by Kydland and Prescott (1990) as a repetition of several stages of addition and subtraction of deviations, which are frequently amenable to precise measurement. The idea of a "cycle" is typically associated with the recurrence of certain events in a predetermined order, with a recurring but non-periodic nature and without fixed periodicity, amplitude, or length.

One of the most important references for the definition of business cycles comes from Burns and Mitchell (1946). The central idea is that cycles are formed by recurring movements, which include periods of deceleration that occur more or less simultaneously in various economic aggregates, followed by periods of acceleration that connect to the decline phase of the subsequent cycle.

On the other hand, as is shown in important references in the literature such as the work of Lucas (1997), current approaches to business cycle theories leave the classical definition behind and focus on the observation of deviations from the long-term trend component. Thus, the growth-cycle notion is given a lot of weight in the minds of contemporary authors.(Altissimo et al., 2000)

Although it has become conventional to see cycles as movements around a trend, the challenge of estimating that trend has never been fully overcome. Testing the veracity of one hypothesis or another will always be challenging since these components are unobserved. One of the goals of this work is to help clarify the separation of cycles and trends through observation of the same using cycle-trend decomposition methods.

The methods of cycle-trend decomposition are the ones most frequently used in monitoring and measuring business cycles. They are quite simple to deploy, always relying on a few hypotheses about how each of the economic activity's components would behave. Contemporary empirical macroeconomists do cycle-trend decomposition using a range of smoothing and detrending approaches. (Baxter and King, 1995)

Nowadays, great emphasis is placed on statistical filters, which try to separate specific frequencies from a series in order to determine the cyclical component. These filters are the decomposition method chosen for the focus of this work. Many filtering methods have been proposed to filter the trend component, including linear and nonlinear filters.

The filter developed by Hodrick and Prescott (1997) is unquestionably the most well-known and popular statistical filter for time series decomposition in macroeconomics, finance, and business analytics, and therefore it has been extensively researched, discussed, and criticized. Therefore, another goal set for this work is to compare the performance of this linear filter to that of two non-linear filters so that conclusions may be drawn regarding the performance of linear vs. non-linear filters.

The HP filter is presented in **Chapter 2**, along with its theoretical basis, underlying logic, and recent academic criticism. Also presented in this chapter are the two nonlinear filters that may be used as an alternative to the problems found in the HP filter: the Linear Programming Approach proposed by Mosheiov and Raveh (1997) and the Median Filter proposed by Wen and Zeng (1999). At the beginning of this chapter, a more mathematical and analytical introduction to the time series decomposition is given, also based on the existing literature.

Chapter 3 describes the procedures and tools used throughout the work to get the required results for the analysis. The data used for the study is provided firstly, followed by a benchmark approach for dating business cycles. The methodologies used to compute the different filters under discussion, as well as the statistical tools that will be used to evaluate their performance, are then described. Finally, Power BI, the platform on which all of the work's methodology was developed, is presented. Although it is not the primary goal of this study, it is expected that by the end of it, conclusions regarding the capabilities of Power BI, a platform on the rise in the field of statistical analysis, would be attainable.

The application of the methodologies and tools described in the prior chapter, together with the provided data, is done in **chapter 4**. As a consequence, in this chapter, readers may find the results obtained after applying the filters and statistical measures to an aggregated series of the Portuguese economy. A comparative analysis of the results obtained is presented so that an analysis of the different filters can be made. Thus, the advantages and disadvantages of each filter will be presented. However, without intending to find optimal performance for any of them because, in fact, such does not exist in the sense that the economic cycles and trends that are intended to be represented are not observable.

In **chapter 5**, towards a conclusion, some comments will be made addressing the implications and limitations of the results obtained, with the hope of contributing to the clarification of the objectives that were set.

2 Theoretical Discussion

2.1 Business Cycles

The study of economic cycles thrived and gained prominence with the release of several critical works, such as Wesley C. Mitchell's work in 1931. This author's perspective on business cycles entailed breaking down series into cycles, which were then divided into four distinct phases. This work was continued with the publication of Measuring Business Cycles by Burns and Mitchell (1946), that became a fundamental work in the study of business cycles. The authors proposed a definition of business cycles that has been largely accepted for many years:

> ... a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; this sequence of changes is recurrent but not periodic; in duration business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar character with amplitudes approximating their own.

(Kydland and Prescott, 1990)

Koopmans (1947) created a sort of taboo on the reporting of facts about business cycles by criticizing the aforementioned study. This critique was based on two grounds. The first criticism was that there is no systematic review of the theoretical reasons that justify the inclusion of the variables that were chosen. The second point is that this study lacks explicit hypotheses about the probability distribution of variables. (Kydland and Prescott, 1990)

For several years after that, the study of business cycles ceased to be an active subject of economic studies. Then, the study of business cycles as recurring variations came back to life. In the major research centers, economists and analysts are increasingly concerned with providing explanations for the factors that cause aggregate production and other series to fluctuate repeatedly over a trend. (Kydland and Prescott, 1990)

Another research project of significant importance was presented by Lucas (1977) who revived the business cycle theories. Contrary to the previous vision on business cycles, Burns and Mitchell (1946), Lucas does not recognize the occurrence of cycle sequences as inherent variations in economic activity, nor does he see the necessity to differentiate between the various stages of the cycle. What is crucial to Lucas are the comovements throughout time of the cyclical components of economic series, classifying business cycle regularities as "comovements of the deviations from trend in different aggregative time series.". Although Lucas does not define trend, it is presumed by Kydland and Prescott (1990) that he shared their viewpoint and that the idea of trend is tied to the steady state growth theory [1]. This long-term growth depends directly on the rate at which technology advances. In theory, it was expected that the trend would be approximately the curve that would be drawn on the graph of the time series being decomposed. It should also be a linear transformation of the series, and the lengthening of the series shouldn't significantly change the value of the deviations. However, passing these requirements onto the time series decomposition procedure has been largely debated.

2.2 Time Series Decomposition

Before delving into the topic of series decomposition and its mechanisms, a few fundamental and introductory time series considerations must be made. A time series differs from cross-sectional data because its variables are measured over time. It can be written as:

$$\{x_1, x_2, \ldots x_T\} \text{ or } \{x_t\}, t = 1, 2, \ldots T \tag{1}$$

As indicated by Wooldridge (2012), economic time series are outcomes of random variables (variables that are not known before), and these sequences of random variables indexed by time are conventionally known as stochastic processes or time series processes. The sample size of a time series corresponds to the number of time intervals over which we observe the variables in the study.

Usually, economic time series are not independent and identically distributed. This is the case for GDP, where the values are related over time. As a result, as Cochrane (1997) pointed out, a day with a very low value is unlikely to be followed by one with a very high value, or vice versa. Wooldridge (2012) noted that when studying the relationship

[1] Per capita income, consumption, investment, and others must all rise at the same pace as innovation and technology in order for there to be steady state growth.

between variables using regression analysis, we must assume that there is some stability; otherwise, when the variables change arbitrarily at each time interval, it is very difficult to know how the variables affect each other.

Numerous economic time series, such as the GDP, show steady growth throughout time. This movement may be easily explained as a linear model of the data made up of a deterministic temporal trend and a noise component. When the trend is eliminated from these trend stationary series, the outcome is a stationary process, meaning that the mean and autocovariances are time independent. Detrending and differencing are prominent macroeconometric techniques for converting a series into a stationary process. However, as Hamilton (1994) pointed out, both techniques have significant limitations and may be insufficient to represent the business cycle component found in most macroeconomic indicators.

Time series (y_t) can be decomposed into trend or growth (x_t, low frequency), seasonal (s_t), cyclical (c_t, medium frequency), and irregular (ϵ_t, high frequency) components (Enders, 2014).

$$y_t = x_t + s_t + c_t + \epsilon_t, t = 1, \ldots, n. \tag{2}$$

Other frequencies than c_t are undesirable because the purpose is to focus on business cycles. Therefore, there are some mechanisms to remove seasonality, but macroeconomic aggregates are usually offered seasonally adjusted, so this component was not treated in the data preparation phase.

Thus, forecasting practice is typically supported by time series (y_t) decomposition into trend or growth (x_t) and cyclical (c_t) components, which Mosheiov and Raveh (1997) refers to as the main factors that contribute to data variability. When studying macroeconomic data, the trend component determines the slowly varying growth of aggregate economic activity, and the cyclical component stands for the rapidly varying fluctuations known as business cycles. (Phillips and Jin, 2021). As Enders (2014) pointed out, the slope of the trend line was thought to be determined by slowly varying factors, such as technological evolution, and business cycles were shown to be one reason for the deviations from the trend. Thus, trend estimation is one of the main tasks when forecasting time series. The trend is commonly perceived as a steady movement over time that may be extracted using suitable filtering techniques, and the cycle is approximated by the residual, $\epsilon_t^* = y_t - x_t^*$.

Linear and non linear filters can be applied to perform this task. Commonly, linear filters are the most widely used ones, and their popularity is due to their theoretical basis and their known elegance. Non-linear filters are less used, and the amount of underlying theory is reduced. Wen and Zeng (1999) proposed that non-linear filters are better at capturing sharp changes in an economic series' trend stochastic component. The authors also noted that linear filters tend to produce smooth trends that can fail to capture sharp changes, underperform in the presence of signal-dependent noise, and cannot effectively remove heavy-tailed distribution noise. Recently, nonlinear filters have been considered successful because they preserve sharp changes in the data and perform robust noise mitigation. This is in agreement with the conclusions of the research performed by Assunção and Fernandes (2017), which says that nonlinear robust filters could have better performance than linear filters while caching discrete changes in the trend growth component of economic series. They pointed out that linear filters produce "artificial" decompositions of trend and cycle components that might not reflect the real data generation process.

In certain circumstances, moving averages may be used to extract or remove trends in economic time series. We generate a new time series, y_t^*, by using a moving average on a time series y_t:

$$y_t^* = \sum a_k y_{t-k}. \tag{3}$$

Moving averages are commonly employed in two different ways: to smooth the time series in order to determine the underlying trend (two-sided moving average, (4)) or as basic time series forecasting approaches (one-sided moving average, (5)).

$$z_t = \frac{1}{2k+1} \sum_{j=-k}^{k} y_{t+j}, \quad t = k+1, k+2, \ldots, n \tag{4}$$

$$z_t = \frac{1}{k+1} \sum_{j=0}^{k} y_{t-j}, \quad t = k+1, k+2, \ldots, n \tag{5}$$

Baxter and King (1995) demonstrated that a symmetric moving average ($a_k = a_{-k}$ for $k = 1, \ldots K$) with zero weights yields a stationary series with quadratic deterministic trends. Moving averages can also render stochastic trends stationary when a time series is a representation of an integrated stochastic process.

2.3 Statistical Filters

2.3.1 Hodrick-Prescot filter

The Hodrick-Prescott filter represents a moving average linear filtering technique that is widely used to decompose a time series into trend and cyclical components. When applied to stationary time series, this filter approximates an ideal high-pass filter. A good high-pass filter suppresses low frequencies while allowing higher ones through. (Guay and St-Amant, 1996)

The smoothness of the growth component is measured by the squares of the second difference, and the cyclical component represents the deviations from the trend. Under this assumption, the use of the HP filter entails reducing the variability of the cyclic component and is subject to a penalty for the variance in the growth component's second difference. The trend component is thus the solution to the following problem:

$$\min \left[\sum_{t=1}^{N} (y_t - x_t)^2 + \lambda \sum_{t=3}^{N} (g_t - g_{t-1})^2 \right], \tag{6}$$

The larger the value of λ, known as the smoothness penalty, the smoother the estimated trend will be. Hodrick and Prescott (1997) citing prior beliefs proposed that $\lambda = 1600$ for quarterly data. The baseline theory says that a larger change in the cyclical component would be around 5% and a larger change in the trend component would be around 1/8% (per quarter). As $\lambda = \sigma_c^2/\sigma_v^2$ it was suggested that $\lambda = (5/(1/8))^2 = 1600$. As Ravn and Uhlig (2002) highlighted, the majority of researchers followed this value for quarterly data, however, there is no agreement when the data comes in other frequencies and therefore they proposed a methodology to choose the smoothness penalty for nonquarterly data.

In its critique, Hamilton (2017) stressed that the HP filter produces spurious dynamic behaviors created by the filter rather than returning cycles and cross-correlations of the real data generation process. Harvey and Jaeger (1993) and Cogley and Nason (1995) make similar claims, claiming that this filter can generate non-existent cycles and equations. It becomes unclear whether the outcomes are facts or artifacts, which might lead to misleading conclusions. These observations are based on Nelson and Kang (1981) demonstration. Implementing the HP filter in an integrated process is identical to detrending a random walk process. The authors demonstrated that a detrended random walk contains spurious cycles with average lengths of nearly two-thirds of the sample length. As a result, the detrended series has significant transitory variations whereas the original data contains none (Cogley and Nason, 1995).

St-Amant and van Norden (1997) highlighted another issue with this filter. The HP filter produces a two-sided moving average, and its weights change from the mid-sample to the end of the sample; the filter becomes a one-sided moving average at the extremes of the series, and it may be necessary to ignore extreme data. As a realistic two-sided filter, in the middle of the sample, no observation is given more than 6% of the weight. However, at the extremes of the sample, the filter is one-sided, and the final observation contributes up to 20%. As a result, the HP trend becomes more erratic near the last observations. This issue may also be seen symmetrically at the beginning of the sample.

Mohr (2005) demonstrated that the HP filter cyclical component for quarterly data (estimated with $\lambda = 1600$) must contain some residual seasonality as well as other high frequencies with less than six quarters blended with genuine business cycle frequencies.

2.3.2 Median filter

The group of nonlinear filters known as "median filters" was developed by Wen and Zeng (1999). These filters have proven to be extremely useful in removing trend and noise components from time series, particularly in the field of electrical engineering. The authors intended to demonstrate that this type of filter may also be highly effective in economics and proposed the median filter, arguing that it performs better than linear filters such as the HP filter. They pointed out that linear filters automatically attribute all acute changes in time series to non-fundamental shifts, ignoring the possibility that the growth of the series may also undergo abrupt shifts and peaks (Wen and Zeng, 1999). The median filter uses a median measure X as a smooth operator. An odd number of sample values are sorted, and the middle or median value is selected as the output, that is, the trend component (Wen and Zeng, 1999).

$$\hat{x}_t = \mathrm{X}\left[y_{t-k}, \ldots, y_t, \ldots, y_{t+k}\right] \tag{7}$$

where $t = 1,...,N$, length of l=$2K$+1, and for quarterly data, the chosen value for k is 11. To filter the outmost input samples, when parts of the filter window go beyond the input signal, the outmost input samples (y_1 and y_N values) are reproduced as many times as necessary. This approach is known as "First and Last Values Carry-on Appending".

Comparing the median filter, with a length of 23 quarters (l=2*11+1=23), with other filters, Wen and Zeng (1999) found that the linear filters are excessively noisy for business cycles compared with their method. The proposed filter was compared with the linear HP, MA, and BP filters, and the cycles produced by the median filter for GDP matched almost perfectly the NBER dated times of recession in the USA. This suggests that the linear filters tested, such as the HP filter, are inefficient in preserving GDP growth and successfully catching rapid changes in the trend component.

2.3.3 Mosheiov-Raveh filter

The Mosheiov-Raveh filter is a non linear filtering technique alternative to the Hodrick-Prescott filter, producing a piecewise trend. A decent estimate of the trend should have two key characteristics: fidelity (closeness to the data) and smoothness. Mosheiov and Raveh (1997) proposed a linear programming approach to estimate the trend, using the sum of absolute values instead of the sum of squares to measure the smoothness and fidelity of the trend (Zhao and Wei, 2003). Mosheiov and Raveh (1997) set the goal of minimizing a linear combination of these two properties.

In order to measure the fidelity property it is used the sum of absolute deviations. That is:

$$\sum_{t=1}^{N} |Y_t - T_t| \tag{8}$$

The measure of smoothness is given by the sum of the second order differences:

$$\sum_{t=1}^{N-2} \Delta^{(2)} T_t \tag{9}$$

It is assumed that the time series, Y_t, is approximately monotone. However, the estimated trend is required to be perfectly monotone. Therefore, the estimated trend is the vector that minimizes the linear combination between fidelity and smoothness:

$$\min \left\{ \alpha \sum_{t=1}^{N} |Y_t - T_t| + (1-\alpha) \sum_{t=1}^{N-2} |T_{t+2} - 2T_{t+1} + T_t| \right\} \tag{10}$$

For the highest value of α, $\alpha = 1$, the trend is the best estimate concerning fidelity. On the contrary, $\alpha = 0$ would produce a straight line because it would be the best estimate regarding smoothness. (Mosheiov and Raveh, 1997) The α value is user-dependent and can totally change the estimation result. Mosheiov and Raveh (1997) proposed a value of

0.1 based on their research. However, given that $\theta = (1-\alpha)/\alpha$ and the fact that Hodrick and Prescott (1997) assumed $\lambda = 1600$, then Assunção and Fernandes (2017) proposed that $\theta = \sqrt{1600} = 40$ would be a better choice for quarterly data.

It is also possible to use this filter in polytone curves without additional computational effort as long as the points at which the function changes between decreasing and increasing are previously determined.

The monotonicity required on the estimates is considered the biggest limitation of this method for the authors, although they defend the fact that this limitation is weaker than the ones that have been proven to exist on linear filters. If the requirements are respected, it is possible to compute a good trend estimate with reasonable values for both fidelity and smoothness. An advantage of this filter is that it does not force the use of fixed coefficients or weights (Mosheiov and Raveh, 1997) (Assunção and Fernandes, 2017). The monotonicity condition can be relaxed by computing the MR trend with least absolute squares (quantile regression) instead of the linear programming approach (Assunção and Fernandes, 2022).

3 Methodology and Data Collection

3.1 Data Collection

The time series studied and to which the filters under analysis were applied correspond to macroeconomic data belonging to the Portuguese national accounts. This data is available from the *Instituto Nacional de Estatística* (INE) [2]. The European System of National and Regional Accounts (ESA 2010) serves as the methodological guideline for the establishment of the National Accounts, creating a uniform, systematic, and precise approach their compilation, thereby ensuring worldwide comparability of outcomes.

The time series selected for analysis and incorporated into the study correspond to Gross Domestic Product (GDP), Private Consumption, Investment and Exports. The time interval of these observations ranges from the first quarter of 1997 to the second quarter of 2022, inclusive. All data is quarterly distributed, so there are 182 observations for each of these four series in the used dataset. The use of quarterly data allows the fixed parameters needed to compute the filters to be the most widely used, most studied, and most corroborated in the literature. It is also important to note that all these series are already seasonally adjusted, which prevents interference from unwanted frequencies in the study of the trend and cyclical component.

GDP is a broad indicator of a country's economic output and can be influenced by private consumption, investments, exports, and other indicators. It is the total value of all products and services generated in an economy over a period of time, usually a year. It is often used as an indicator of a country's economic success and is seen as a significant measure of the overall state and performance of an economy. Private consumption is also an important driver of economic growth because it increases the demand for goods and services. It corresponds to household spending on goods and services and is a crucial element of GDP since it accounts for a considerable share of overall economic activity. Investment can be an important engine of economic growth because it increases the production capacity of an economy. It refers to spending by firms and households on capital goods, helps boost the productive capacity of an economy, and can also result in economic growth. Finally, efficient export management can help increase an economy's revenue and promote growth. Exports are the sales of products and services made in a country to

[2]This data can be accessed here.

customers in other countries. They are a significant source of foreign money and, by increasing demand for local products and services, can promote economic expansion.

Therefore, through these indicators, it is possible to make an evaluation of the country's economic situation over the last 45 years. Economic indicators like this one are possibly the most fundamental ways to track economic conditions. They provide short-term economic analysis and can be useful components in econometric forecasting models. Although GDP is widely acknowledged as the key measure to employ in the study of economic activity, the other series used for the same studies were also put to the test, since this entire study is based on a more analytical approach to the decomposition of economic series. One reason for using these three variables alongside GDP is that they exhibit pro-cyclical behavior, which means that their cyclical components move in the same direction as GDP.[3] The representation of the series' initial distributions is presented below:

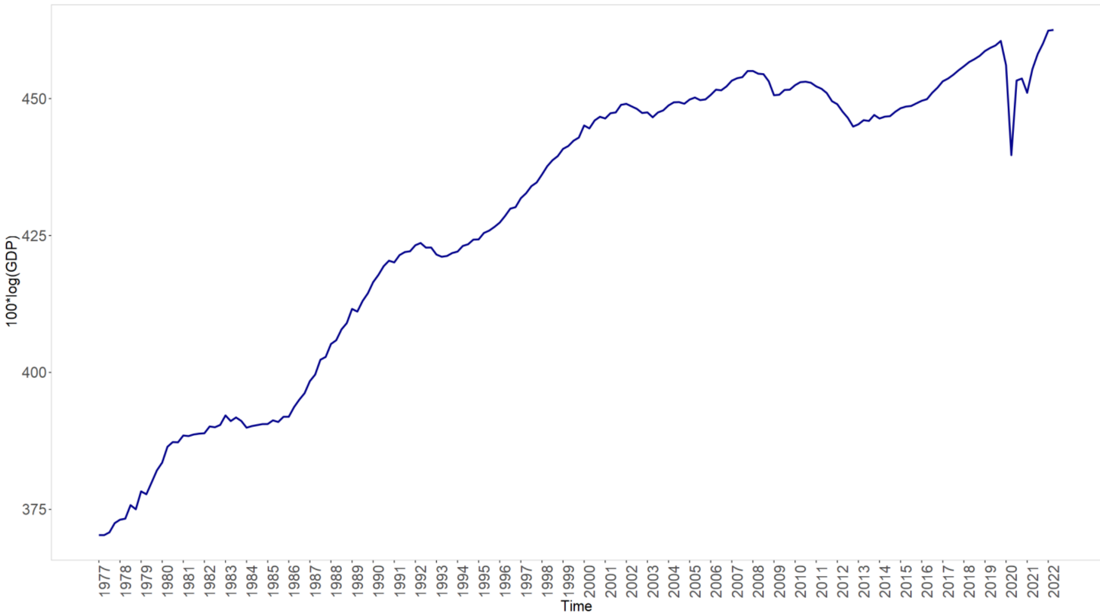

Figure 1: Portuguese GDP from 1977 to 2022

A word of caution must be added. Currently, the value that represented 100% of GDP in the first quarter of 1977 only accounts for 39.8% of the total GDP observed during the same period but for the year 2022. Thus, it makes sense that the value of the cyclical and trend components will change throughout time and space in accordance with development. Before applying the filters, the logarithm of the series was used to ensure that this did not affect the analysis; hence, the results that follow are expressed as percent

[3]This behavior was demonstrated on appendix, section 6.1

deviations.

3.2 Dating Business Cycles

Given that the primary goal of this work is to assess the performance of various filters in extracting existing frequencies on economic time series, namely, business cycles, it is critical to date them as a reference for comparison and analysis of results. To this end, it was employed one of the most widely used techniques worldwide for dating business cycles.

The most prominent and well-known case of dating the high and low points of economic activity (generally referred to as peaks and troughs, respectively) is the NBER, National Bureau of Economic Analysis, for the United States of America. To implement NBER definitions, Gerhard Bry and Charlotte Boschan created their algorithm in 1971. The principle of the algorithm relies on a combination of filters and criteria to find local minima and maxima at the series' level (or logarithm). A local minimum is a trough that occurs after a local maximum, which corresponds to a peak, and the period between a peak and a trough corresponds to an expansion, whereas the period between a peak and a trough corresponds to a recession. The algorithm suggests that each phase of the business cycle (from peak to trough or vice versa) must last at least five months and that peaks and troughs must alternate. (Rua, 2017)

Several economic series like GDP are usually available on a quarterly basis and the Bry and Boschan algorithm was designed for monthly frequency. As a result, the "BBQ algorithm" proposed by Harding and Pagan (2002) was used, which has the same properties as the original algorithm but is adapted to quarterly data.

The central bank of the Portuguese Republic, *Banco de Portugal*, published in 2017 an article that establishes a reference chronology for the business cycles comprising the period from the beginning of 1977 to the end of 2015 for Portuguese GDP. Rua (2017) used the Harding and Pagan (2002) method.

As is widely acknowledged, the GDP is a measure of economic activity. According to Burns and Mitchell (1946), a combined economic activity might get a more defined and conceptually quantifiable meaning through this indicator. As a result, and for the sake of brevity, the business cycle chronology generated using the methods outlined for the GDP indicator is shown on the table 1.

Since 1977, six recessionary periods have been identified, with the shortest and most recent lasting only 2 quarters in 2020 and the longest lasting 9 quarters between 2010 and 2013. The asymmetry between expansions and recessions is an economic cycle phenomenon that is also evident in the Portuguese case. The average length of a recession is 4.7 quarters, whereas the average length of an expansion is 24.2 quarters. This translates to the average length of the Portuguese economic cycle of 28.9 quarters. The dated business cycles have shown an amplitude of 17.6 for expansions and 6.8 for recessions.[4]

Phase	Start date	End date	Duration
Expansion	-	1983Q1	-
Recession	1983Q1	1984Q1	4
Expansion	1984Q1	1992Q2	33
Recession	1992Q2	1993Q2	4
Expansion	1993Q2	2002Q1	35
Recession	2002Q1	2003Q2	5
Expansion	2003Q2	2008Q1	19
Recession	2008Q1	2009Q1	4
Expansion	2009Q1	2010Q3	6
Recession	2010Q3	2012Q4	9
Expansion	2012Q4	2019Q4	28
Recession	2019Q4	2020Q2	2
Expansion	2020Q2	-	-

Table 1: Chronology of recessions and expansions

[4]The vertical distance of cyclical components between trough and peak is used to calculate amplitude.

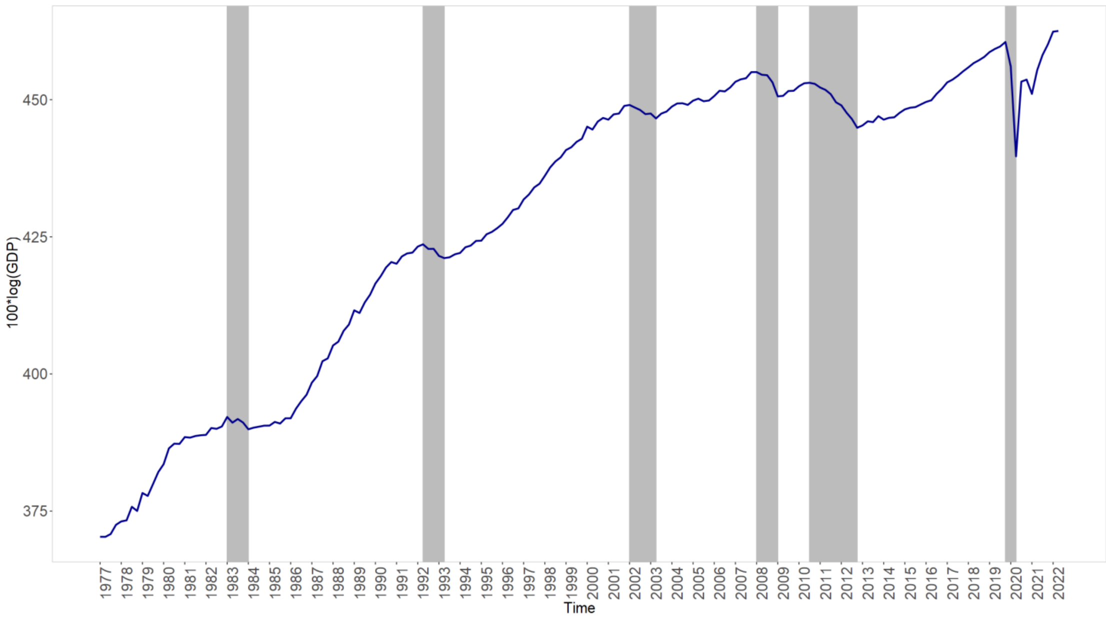

Figure 2: Portuguese GDP and recessions dates shaded

3.3 Filters Computation

The three filters mentioned in the previous chapter were programmed in two programming languages, R and Phyton, through R and Phyton Script Visuals available in Power BI. The main code is available in the appendix, 7.4.

3.3.1 Hodrick-Prescot filter

The Hodrick-Prescott filter was computed following the computational strategy described by Kim et al. (2009). As described in section 2.3.1, when applying the HP filter, the trend component is chosen to reduce either the smoothness of the trend or the sum of the square residuals, that is, to minimize the equation (6). Where the first term in the equation represents the residual size and the second term represents the smoothness of the predicted trend. The regularization value $\lambda > 0$ is used to account for the trend component's second difference. As a result, the bigger the value of λ, the smoother the estimated trend. This value was set to 1600 following the previous literature that uses this value for quarterly data.

The equation (6) may be conveniently expressed as a matrix:

$$\min \left[(y-x)^T (y-x) + \lambda x^T D^T D x \right], \tag{11}$$

with $D \in \mathbb{R}^{(N-2)\times N}$, an upper triangular Toeplitz matrix where the first row corresponds to [1 -2 1 0 ... 0],

$$D = \begin{bmatrix} 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 & \\ & & & & 1 & -2 & 1 \end{bmatrix}. \tag{12}$$

According to Kim et al. (2009), the function (11) has a unique minimizer since it is strictly convex in x and it can be written as:

$$x^* = \left(I + \lambda D^T D\right)^{-1} y. \tag{13}$$

We could also derive the optimal fitting error, i.e. the cyclical component of the filter, from the optimal condition $y - x^* = \lambda D^T D x^*$:

$$c^* \equiv y - x^* = \lambda D^T D \left(I + \lambda D^T D\right)^{-1} y. \tag{14}$$

Following the strategy described, a function that computes the trend component $x*$ given a time series y and a smoothness parameter λ was created. The developed function was easily supported by the languages used and required no significant computational effort.

3.3.2 Median filter

The Median filter was computed under the 'first and last values carry-on appending strategy' proposed by Wen and Zeng (1999), described before. The trend component x_t is the solution of the equation (7). Where k=11 following the previous literature that uses this value for quarterly data.

The developed function for median filtering replicates the 11 first and last values of the series, and after that the median values are computed, the output is a new series that represents the trend. This strategy has a fairly simple basic logic, and as a result, this function has been well integrated in the programming languages used.

3.3.3 Mosheiov-Raveh Filter

As described on section 2.3.3 the Mosheiov-Raveh filter aims to solve the equation that origins the minimal linear combination between the sum of the absolute deviations and the second order differences. This is represented by equation (10), where the first term in the equation represents the fidelity measure, while the second term represents the smoothness of the predicted trend. α was set to 40 following the previous literature that uses this value for quarterly data.

On this case it is also useful to use the matrix form of problem (10) in order to solve it:

$$\min\left\{\|y - I_n h\|_1 + \theta\|Dh\|_1\right\} \tag{15}$$

where $D \in \mathbb{R}^{(N-2)\times N}$, is an upper triangular Toeplitz matrix where the first row corresponds to [1 -2 1 0 ... 0], the same as (12).

To solve this problem and find the trend component, the modification to Karmarkar's original algorithm proposed by Meketon et al. (1986) was applied. This technique, formulated to solve linear problems, was implemented by Koenker (2008). This approach is simple to implement, especially if a weighted least squares subroutine is accessible. In reality, it is an iterative reweighted least squares method. This approach uses median regression and focuses on minimizing the sum of the absolute residuals, a problem known as Least Absolute Deviations (LAD). It is simpler to use and more adaptable than the conventional linear programming technique; another benefit is the relaxation of the monotonicity condition in the piecewise linear trend, as noted by Assunção and Fernandes (2022).

3.4 Statistical Tools

3.4.1 Cross-Validation

As James et al. (2021) described, cross-validation may be employed to determine the test error associated with a particular statistical learning approach to assess its performance (model evaluation) or to choose the suitable amount of flexibility (model selection).

K-fold cross-validation is one of the most commonly used standard procedures. This method divides the set of observations at random into K groups, called folds, of about similar size. The first fold is used as a validation set, and the method is applied on the other K-1 folds. Accuracy measures are determined based on the data in the held-out fold. This process is performed k times, and each time a new set of observations is used as the validation set. The figure below illustrates the procedure and the split of the series into training and test sets.

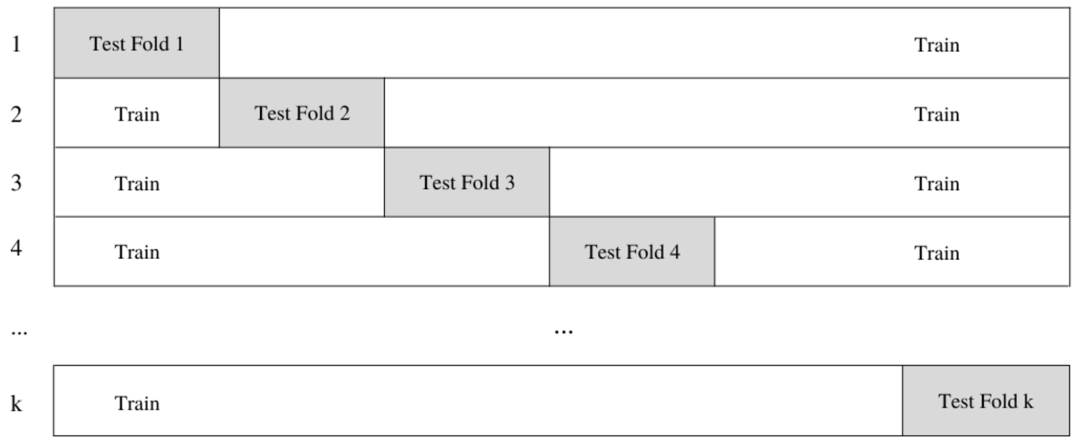

Figure 3: K-Fold cross-validation schema

When it comes to time series prediction, however, it is not possible to take random values from the series and allocate them to either the test or training sets because data from the future should not be utilized to predict data from the past. Therefore, the cross-validation approach mentioned does not work on dependent data, such as time series. When working with time-related and rapidly evolving situations where the properties of the environment change over time, time-based splitting is the best method to create a statistically robust model assessment and a better representation of real events. It was employed time-based cross-validation to compare the different filters used to predict the trend component of economic time series. This method uses a sort of sliding window practice; it begins with a small sample of data for train, then forecasts future values, and after that verifies the projected data points' accuracy. This process can be illustrated by the following figure.

It is necessary to assess how well the predictions correlate with the data observed in order to evaluate the performance of each method. To that end, there are many accuracy measures, each with its own peculiarities, that quantify how close the predicted value is to the actual value for a given observation. The K-fold cross-validation procedure yields K estimates of the test error. Averaging these values yields the k-fold cross-validation estimate.

One of the best-known measures is the Mean Squared Error, which measures the av-

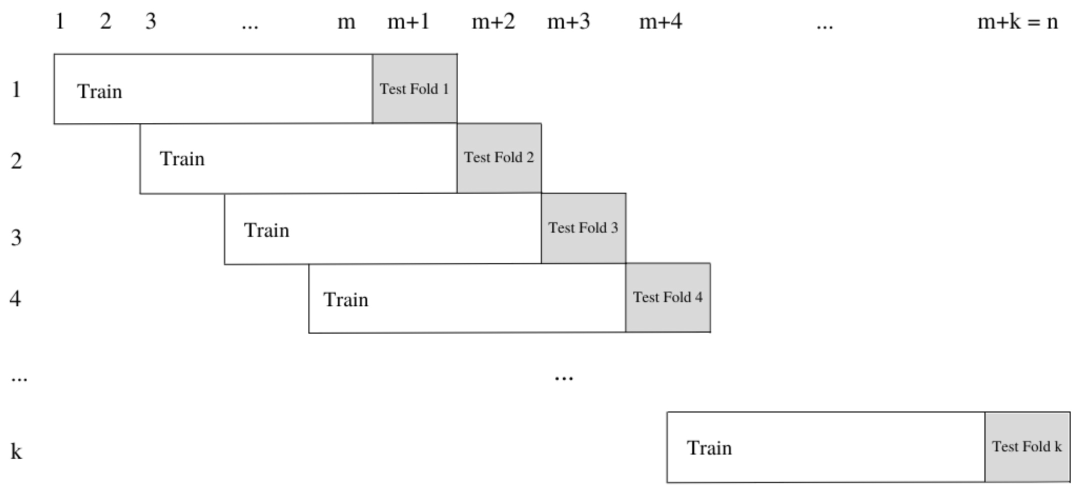

Figure 4: Time based cross-validation schema

erage of the squared difference between the real and predicted values. Its cross-validation estimate is given by:

$$\mathrm{CV}_{\mathrm{MSE}} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y})^2 \tag{16}$$

Although the Root Mean Squared Error is more commonly used in several applications, particularly when the outcome is numerical, because it operates on the same scale as the data. It calculates how distant the observed values are on average from the model results, and its cross validation estimate is defined as:

$$\mathrm{CV}_{\mathrm{RMSE}} = \sqrt{\mathrm{CV}_{\mathrm{MSE}}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y})^2} \tag{17}$$

Another popular measure is the Mean Absolute Error. Bergmeir et al. (2018) demonstrated that MAE may also be applied to the theoretical results of their experiments with cross-validation procedures on time series. It assesses the average size of predicted errors without taking into account their direction, and its cross-validation estimate can be written as:

$$\mathrm{CV}_{\mathrm{MAE}} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}| \tag{18}$$

In the average, MAE weights all individual differences equally. RMSE, on the other hand, assigns a fairly significant weight to large errors since they are squared before being averaged. As a result, the RMSE will always be bigger than or equal to the MAE. The greater the gap between them, the greater the variation in the sample's individual errors. If they are equal, then all mistakes have the same magnitude.

The values of the cross-validation estimates for the three filtering methods under consideration were compared using the approach mentioned above to see which produced better forecasts of the economic series. The cross-validation was performed with eight folds, taking into consideration the number of existing observations, the computational effort, and the results obtained by it.

3.4.2 Standard Deviation

The standard deviation was calculated to aid in comprehending the features of the business cycles. The business cycle volatility can be assessed by the standard deviation of the cyclical component that was achieved through the different filtering techniques. The standard deviation of a random variable corresponds to the positive square root of the variance (Wooldridge, 2012):

$$\mathrm{sd}(X) \equiv +\sqrt{\mathrm{Var}(X)}, \mathrm{Var}(X) \equiv \mathrm{E}\left[(X-\mu)^2\right] \tag{19}$$

Standard deviation is easier to use and more commonly used in comparison to variance because it respects the units of measurement used. Thus, this measure was used to be able to quantify the average distance of the cycle produced by each filtering method from its mean value.

3.5 Power BI

Many believe Microsoft Power BI to be the best self-service business intelligence solution on the market. It has recently surpassed tough technologies such as QlikView and Tableau to take the top rank. As Knight et al. (2018) defined it, Power BI may be one of the most properly titled products ever created by Microsoft, providing analysts and developers with a comprehensive business intelligence and analytics playground in an interestingly lightweight application. The processes of data exploration, data modeling, data visualization, and data sharing can be made elegantly easy by utilizing a single solution, Microsoft Power BI.

Power BI is a SaaS, Software as a Service, available on the Azure cloud, thus the Microsoft product team develops and adds new features to the product using a cloud-first philosophy. There are two methods for sharing the outcomes with other people. The first and most popular way is through the Power BI Service, which customers may use for a modest monthly subscription fee. The other option is through Power BI Report Server, which can be accessed through either a SQL Server license with Software Assurance or a Power BI Premium subscription. Both ways require the use of Power BI Desktop, which is freely accessible. Defining data discovery and data preparation activities, organizing data models, and designing appealing data visualizations for reports are all possible with the Power BI Desktop application. (Knight et al., 2018) It is more than simply a visualization tool, which is one of the reasons it is regarded as an excellent self-service business intelligence tool.

This application has several features that set it apart from others. The main ones are: DAX, the Data Analysis Expressions language, which permits to quickly and efficiently investigate the data model using advanced business logic; a data wrangling tool that allows to manipulate and transform data into a format suitable for data analysis, Power Query; and pre-packaged interactive visuals that allow to show the data in a manner that consumers can understand. Power BI can also handle massive volumes of data (up to hundreds of millions of rows), allow read-only data in the model, promoting security and integrity, and data from a number of sources may be easily integrated.

Programming languages, like R and Python can be integrated and used to bring new functionalities to the ones mentioned about Power BI. Because R is a language designed primarily for data analytics, it is an excellent complement to Power BI. For years, data analysts have used R to conduct tasks such as data wrangling and visualization. Therefore, it is possible that some functionalities may not exist yet in Power BI but have been used in R. Python has grown in popularity as a data analytics programming language over the past decade. One of the qualities that makes this language so appealing is that it is suitable not only for data analytics but also for ordinary programming tasks. Python makes easy communication with APIs, whereas Power Query makes the same process quite difficult. This makes Python as well a great component to Power BI. Wade (2020) cited some features that can be achieved with the leverage of this languages in Power Bi: create costume visualizations in a pretty easy process; use sophisticated techniques, not accessible in Power Query or DAX, to do extensive string manipulations; without the requirement for Power BI Premium, apply data science to data models and interact with

Microsoft Cognitive Services; interact with third-party data APIs to efficiently augment data models.

4 Analyses and Results

The output of the filtering techniques is shown in figure 5 for the case of Portuguese GDP between 1977 and 2022. Although we can detect some noise produced by the median filter that the other filters do not display, the three filters appear to yield a relatively similar trend. Overall, this trend has been rising over the years, with the exception of the years 2008-2014. The effects of the 2008 recession were still being felt in 2010, when the most severe and longest Portuguese recession started, from which the economy only began to recover in 2014, and it is at this time that we can see the GDP trend growing again.

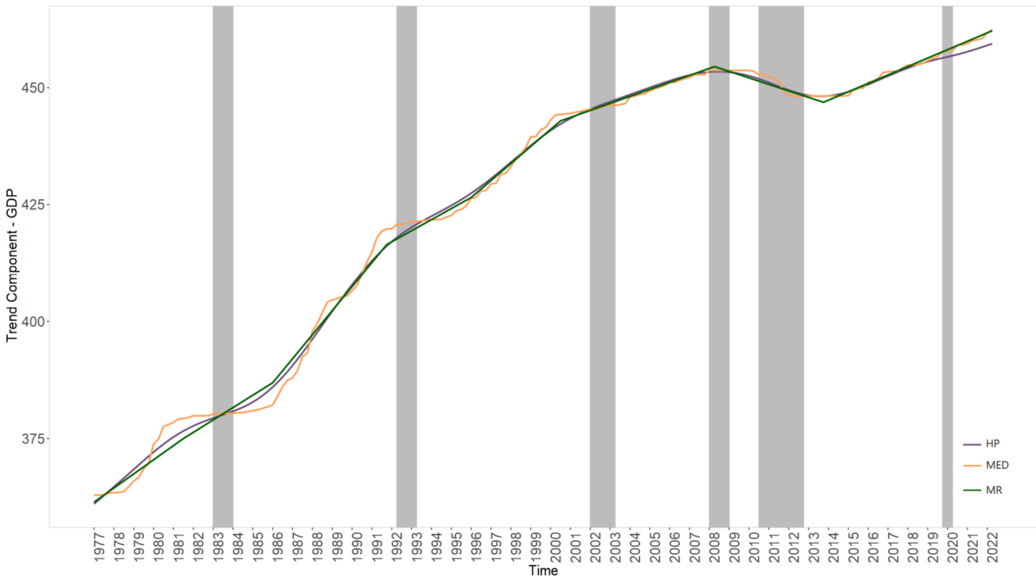

Figure 5: Trend component of GDP

When computing the first differences, the change in the trend over the course of the quarters is obtained, as illustrated on figure 6. It is visible that the trend's growth exhibits quite different behaviors across all three methods. Nonetheless, it is clear that the trend growth in the GDP has been declining. In recent years, this number has been significantly lower than it was at the start of this series.

The MED filter produces more distinct growth; it is more volatile to changes and captures all sharp changes in the trend component's growth. Nevertheless, the MR filter is the one providing the most interesting results.

The plotted results demonstrate the MR filter's outstanding ability to recognize spontaneous periods of differential growth in the time series. The graph depicts some cyclical patterns in the trend's growth. The distinct trend growth levels might be seen as different phases of the economy, with the average trend growth values changing in each period (Assunção and Fernandes, 2017). While all the three filters produce comparable estimates for GDP trend growth that emphasize distinct phases, the MR filter does it in a very natural manner. It is simple to identify eight periods in GDP trend growth and understand where each one begins and ends visually. These intervals appear very naturally, and the recessions found in table 1 are within the weakest periods found here, except for the last one. The other filters, particularly HP, can also show this; however, the start and conclusion of the periods are less evident, and the trend growth produced by the other filters must be averaged to describe each period.

The MR filter clearly creates a more consistent and robust growth than the other filters. This is visible near the end of the sample, where it nicely handles the changes associated with the COVID-19 pandemic, eliminating the oscillation created by the others, which demonstrated to be very affected by end-of-sample data, especially when structural changes are considered.

The growth of the trend component of the remaining series under study is presented in the appendix, section 7.2. In these series, especially in exports, the median filter is excessively noisy, so it is also presented the comparison between the HP filter and the MR filter, where the results mentioned here are confirmed quite well.

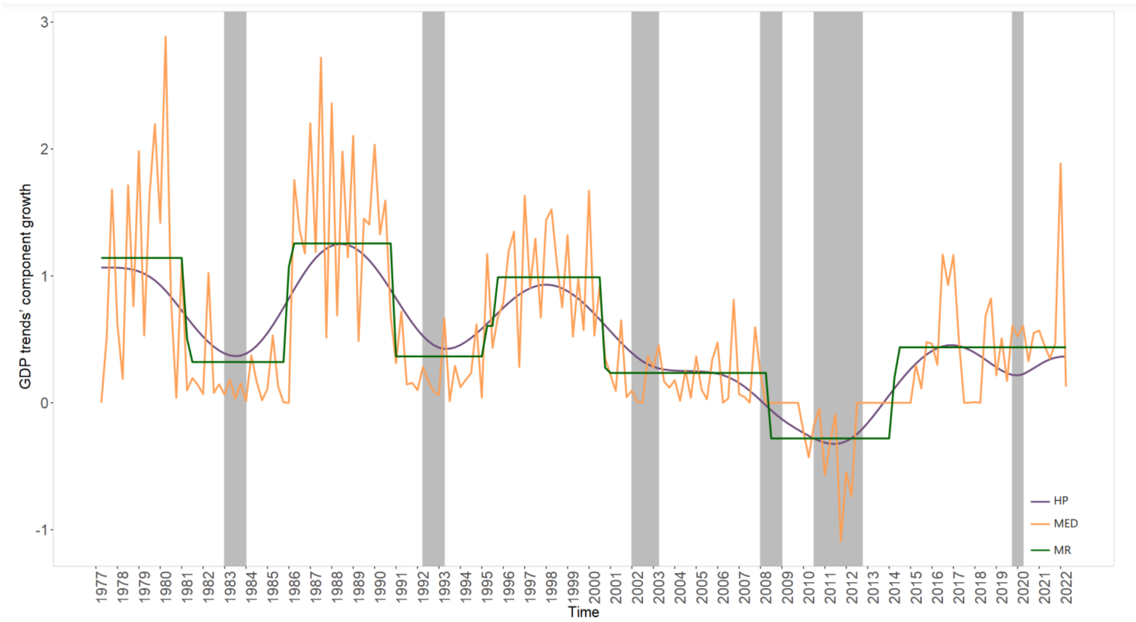

Figure 6: GDP trends' component growth

The cross-validation results also match what was visually analyzed. As shown in table 2, the MR filter is the one that presents the smallest root mean squared error for all four series, meaning that this filter is the one that produces a smaller error variance; the average distance between observed and predicted values is smaller for this technique. Regarding the mean absolute error, the lowest values also coincided with the MR filter, except for the investment (by a small difference of 0.1 in comparison with the HP filter MAE). In general, the MR filter was the source of minor errors in the majority of cases.

	HP		MED		MR	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
GDP	3.48	5.30	5.46	10.04	**3.30**	**5.05**
Exports	8.05	12.61	12.67	26.44	**7.37**	**11.95**
Investment	**2.82**	3.95	3.86	4.67	2.83	**3.91**
Private Consumption	3.73	5.79	5.30	9.79	**3.60**	**5.63**

Table 2: Cross-validation results

At first sight, looking at figure 7, the cyclical component produced by the three filters appears to be fairly similar, but at the beginning of the sample, the differences between them are more evident. Despite this, it is clear from the straighter pattern that the median filter generates a less volatile cyclical component. This is a result of the non-linear filter's capacity to capture sharp changes in trend growth, as previously observed. Other than pure cyclical behaviors, this filter is less affected by noise and shifts. We can also see that the reported decreases in the cyclical pattern of the median filter match the previously dated recessions fairly well. Although the other two filters correspond to these recessions, they nonetheless generate significant declines that the median filter does not produce and do not correspond to any previously dated recession.

It is also worthy of note the massive reduction in GDP's economic cycle caused by the COVID-19 pandemic in 2020, driving it to reach the lowest value recorded throughout the whole series. Regarding the other series under study, all of them also meet the behavior mentioned here; the median filter stands out for a more constant pattern. This can be seen in section 7.3. The export series generates a cyclical component that is much more inconstant than the others.

The results of the visual analysis are in agreement with the results obtained by the

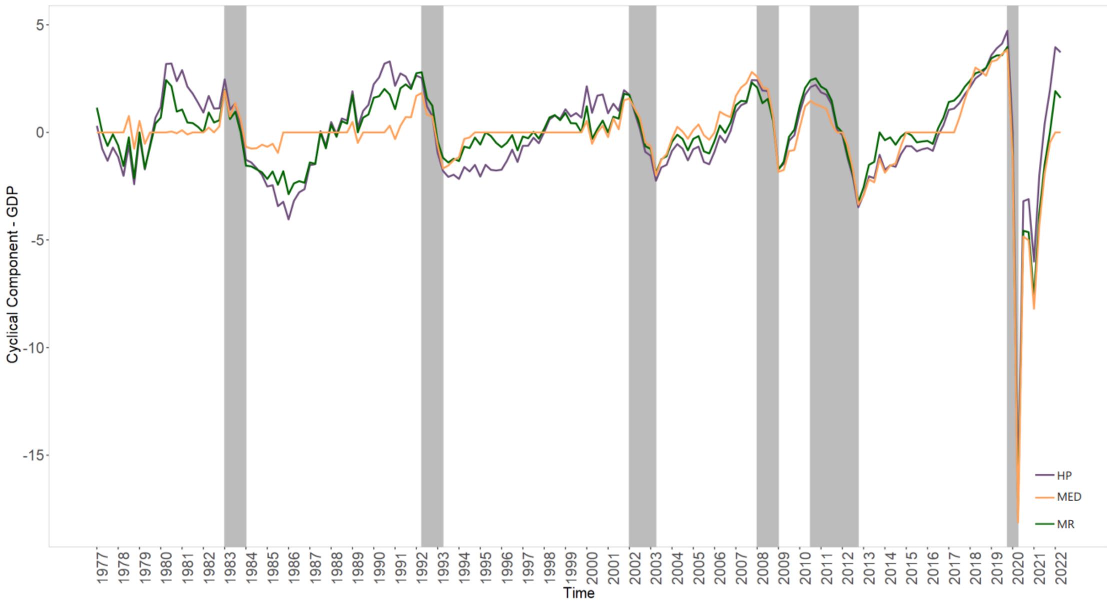

Figure 7: Cyclical component of GDP

standard deviation, table 3 demonstrates that the Median filter is the filtering technique with a smaller average distance between the cyclical component and its mean value. Therefore, we can classify it as the filter that produces less volatile cyclical components. Despite being a non-linear filter, the cyclical component of the MR filter is more akin to that of the HP filter.

	HP	MED	MR
GDP	2.24	**1.95**	2.09
Exports	5.82	**4.97**	6.84
Investment	5.41	**4.61**	5.86
Private Consumption	2.44	**1.84**	2.71

Table 3: Standard Deviation of the cycle

5 Discussion

According to the results, the MR filter generates a better trend estimate and a cyclical component that is quite comparable to the HP filter. It was also demonstrated that the MED filter produces a worse trend component but a better cyclical component than the linear filter. This is consistent with Assunção and Fernandes (2022) previous study, which discovered that nonlinear filters may capture discrete changes in the trend growth of economic series better than linear filters. They concluded that, despite its noisy nature, the Median filter provides noise reduction and is useful in signaling recessions. They also discovered that the MR filter yields reliable estimates of trend growth, which allow for meaningful analyses of diverse economic time periods (Assunção and Fernandes, 2022).

Overall, the findings are consistent with the notion that nonlinear filters outperform linear filters in terms of accurately forecasting the trend to produce a useful definition of the business cycle. However, it is not the purpose to reject linear filters since their advantages are still worth recognizing, and it is critical that their weaknesses be examined so that they may be improved or fixed.

There are some limitations to this study that should be taken into account when interpreting the results. One of the most significant limitations of this work is the limited number of filters used to generate conclusions. Consider testing a variety of other filters, both linear (such as the bandpass filter and the moving average filter) and non-linear (such as the jump process technique). This would have been a better approach. Hence, it should be reiterated that the conclusions mentioned are based only on what has been tested and exposed during the work.

Another limitation that prevents us from having a more efficient, totally clear, and entirely safe idea of the operability of the statistical filters studied here as sources of faithful business cycles is the lack of a reference dating method for the Portuguese cycle accepted as official for the Portuguese economic reality. Preferably, Portugal would have a similar economic monitoring system as the United States. Although a similar approach has been applied, adapted, and used before, it is neither official nor purely created for the Portuguese case. The NBER is responsible for conducting a full and rigorous chronology of the turning points in the North American economy. This chronology is based on the measurement of a variety of cyclical indicators, namely GDP, and is consistent with the concept of the business cycle as a fluctuation between most economic indices.

6 Conclusion

Through this study, it was hoped to examine the potential of linear and nonlinear filters applied to aggregated series of the Portuguese economy in order to determine which ones allow for the best decomposition results, correctly extracting the trend in order to obtain a better definition of the economic cycle.

The graphic and numerical results supported the initial ideas presented through existing literature that nonlinear filters can really perform better than linear filters. The HP linear filter is unquestionably the most popular, and it does produce good results, owing to its ability to predict smooth trend growths and informative cycle components. However, it mostly fails because its characteristics, common to linear filters, make the trends it produces too predictable and regular. The data is too smoothed, which may not represent the reality of the trends' growth. Thus, the cycles it generates are also disrupted because they receive high frequencies that have been wrongly removed from the other components, making them noisy. By establishing a robust, clear, and solid trend nature against structural and less significant changes in the economy, the MR filter has firmly demonstrated its ability to produce consistent trend growth estimates. This trend allows for the examination of distinct periods of economic growth that may be of interest. Regarding the MED filter, the noise of the predicted cycles is actually lower compared to the other cycles. The cyclical component produced by this filter is less volatile than the others and also a good estimate, matching the previously dated recessions. However, its worse performance regarding the trend component should be noted; the trend growth produced by this filter is extremely noisy.

Therefore, it can be recognized that the MR filter has better trend estimates and a cyclical component very similar to that of the HP filter, despite its piecewise pattern, and that the MED filter has a worse trend component but a better cyclical component. Under the methodologies used, the data, and the assumptions, it is safe to say that in fact, non-linear filters can generate better performance than linear ones.

All the computation and presentation of results were supported by Power BI, a computer tool of growing importance in the business intelligence world. It is worth noting that this platform greatly facilitated data processing, visualization, and overall support for many Python and R scripts over the course of this work. The most significant advantage seen was the ability to create several visualization dashboards that allow for very easy extraction of insights from all the generated graphs and tables. This type of visualization allows you to hide the entire computation underneath and only see the various results obtained, whether through a statistical filter or an economic series. Both of the programming languages incorporated into Power BI delivered good results; the functions developed were implemented in a fairly similar manner, simply matching the specifications of each language. According to the literature and as demonstrated, both R and Python are an excellent complement to Power BI, require little computing effort, and are easy to integrate into the program. Despite all of this tool's benefits, it would be a significant improvement to integrate complex routines, such as those that have to be implemented using programming languages outside of the application, in a more straightforward manner. The native language of Power BI, DAX, is limited and does not allow for the implementation of advanced functions and procedures, much like VBA, which we use in another Microsoft platform, Excel. However, it is worth noting that this platform has grown in popularity in the business world and has proven to be adaptable to changes and improvements. As such, as a final note about this tool, here is a suggestion to integrate more sophisticated routines into Power BI so that it can continue to be used as a tool for data visualization, modeling, and exploration, but also as a predictive tool.

Because business cycles are an essential component of the economy and because knowing their origins and effects may help firms, consumers, and governments make wise decisions, it is crucial to keep researching them in the future. For instance, knowing when and how severe future economic occurrences like recessions or expansions will be may help create policies to mitigate their effects, and knowing the economic growth determinants can help businesses decide what investments to undertake. Therefore, it is necessary to establish a more precise definition of business cycles and trends in economic series, especially by using methods like those in this study. Furthermore, it is critical to explore opportunities to improve or optimize the algorithms that underlie robust filtering techniques in order to improve their accuracy and efficiency, which is why some authors have been proposing improvements to the original statistical filters. It is also critical to maintain and promote the use of new statistical modeling approaches, machine learning, or even new business intelligence platforms, such as the one employed here, in order to get increasingly informative findings.

7 Appendix

7.1 Cross-Correlation of GDP with the other series

The cyclical components of a few other series tend to move in a certain way that is synchronized with GDP. These traits are referred to as comovements. We can measure these comovements through the correlation of the cyclical deviations of each series with the cyclical deviations of real GDP. When the correlation is near to one, it can be said that a series is procyclical and its cyclical components move in the same direction as those of the GDP. When the correlation value is nearly one but has the inverse sign, the series is countercyclical and the cycle component of the series move in the opposite way of the cycle component of the GDP. The series is uncorrelated with the GDP cycle when the correlation value is near to zero; it does not change in a predictable manner concurrent with the cycle. (Kydland and Prescott, 1990)

	GDP		
	HP	MED	MR
GDP	1	1	1
Exp	0.5841637	0.8199236	0.5956607
Inv	0.6107651	0.3588964	0.5271634
PrivC	0.8881123	0.9372244	0.8362876

Table 4: Cross-correlation results

From these results it is concluded that all series exhibit a procyclical movement relative to GDP, and this is true for the cyclical component generated by any of the filters used. Only the median filter produces a cyclical behavior of investment that is countercyclical.

7.2 Growth components

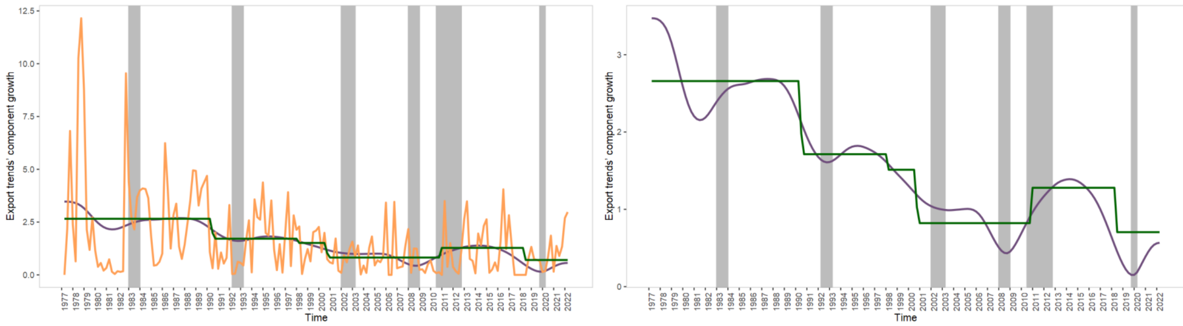

Figure 8: Exports trends' component growth

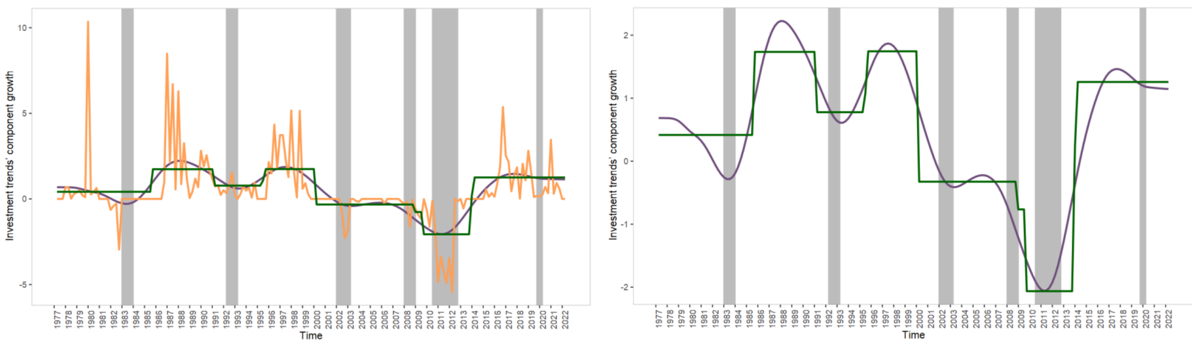

Figure 9: Investment trends' component growth

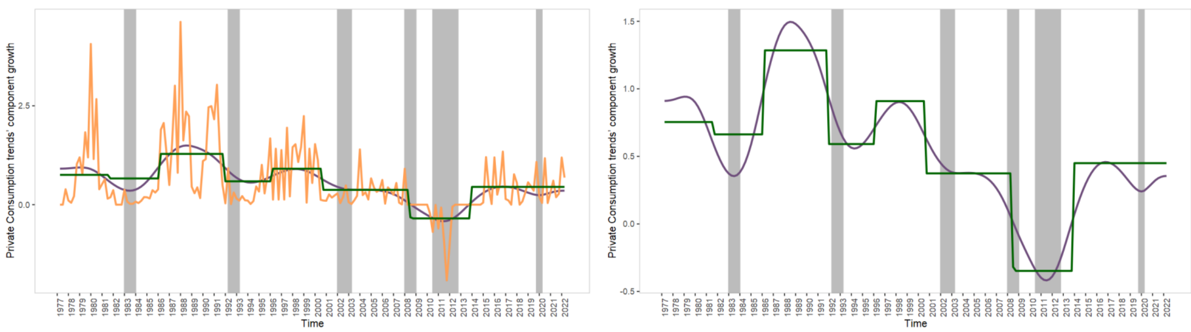

Figure 10: Private Comsumption trends' component growth

7.3 Cyclical components

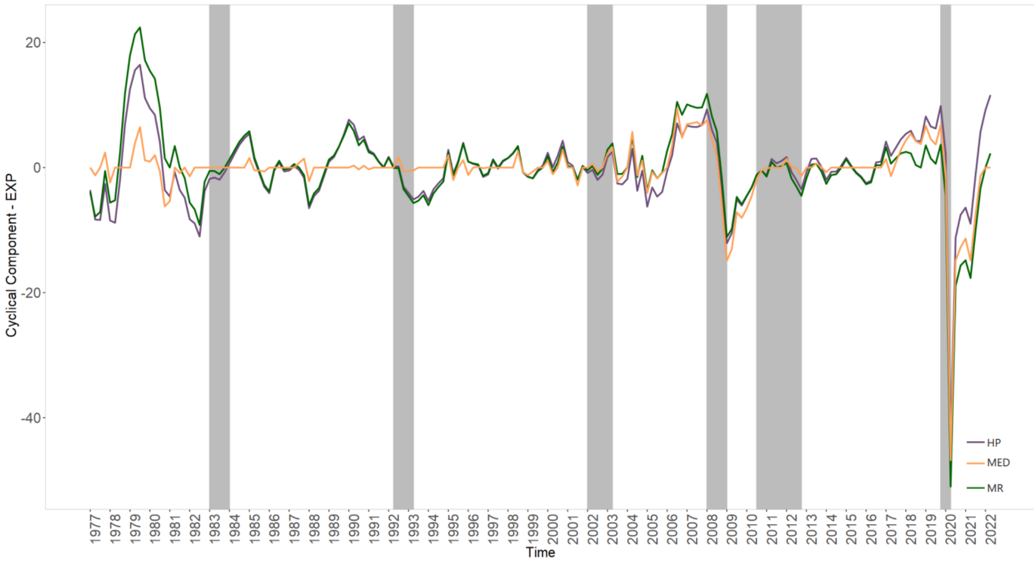

Figure 11: Cyclical component of Exports

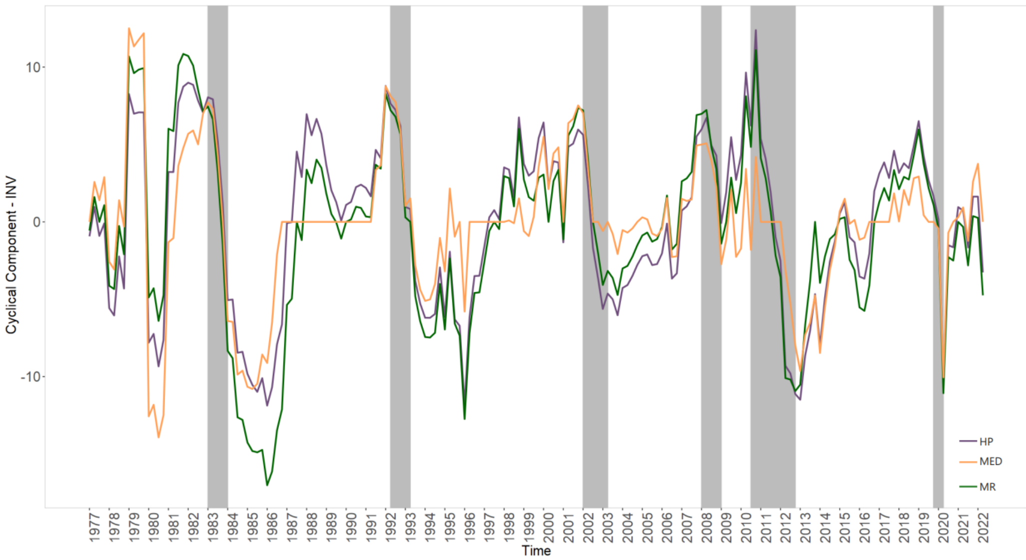

Figure 12: Cyclical component of Investment

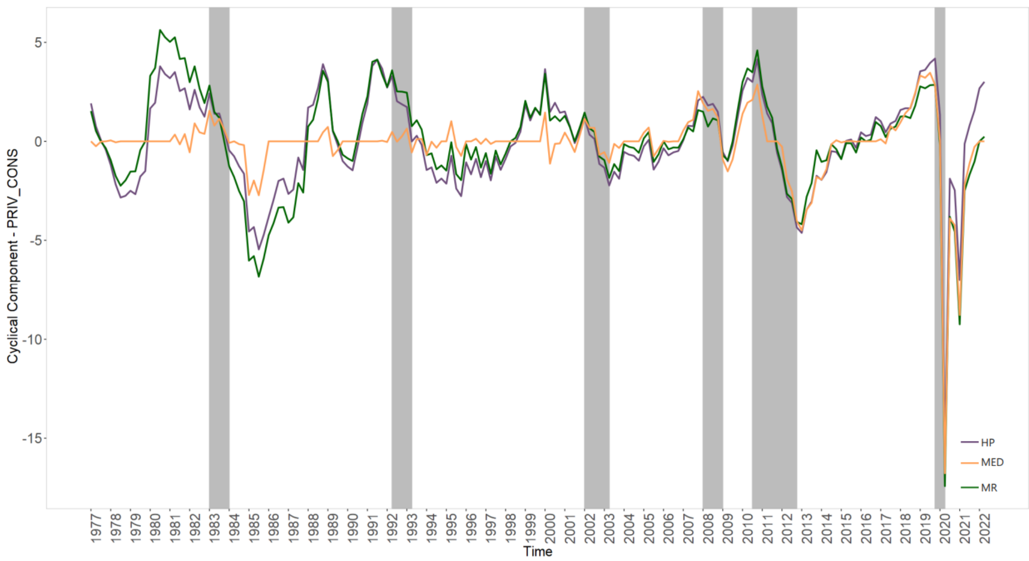

Figure 13: Cyclical component of Private Consumption

7.4 Code

The functions developed to compute the three filters, under the strategies described on section 3.3, are presented in this section.

7.4.1 R code

```
#HP filter
hpfilt <- function(y, lambda = 1600){
  n <- length(y)
  V <- rep(0, n)
  V[1] <- 1
  V[2] <- -2
  V[3] <- 1
  T <- toeplitz(V)
  T[lower.tri(T)] <- 0
  m <- n-2
  D <- T[1:m,]
  D2 <- crossprod(D)
  A <- solve( diag(n) + lambda * D2 )
  x <- A %*% y
  out <- ts(data = x, start = 1, end = n)
}
```

```

#MED filter
medfilt <- function(y, k = 11){
  n <- length(y)
  m <- n + 2 * k
  a <- rep(0, m)
  x <- y
  for (i in 1:m) {
    if (i <= k){
      a[i] <- x[1]
    }
    else if(i > (n+k)){
      a[i] <- x[n]
    }
    else
      a[i] <- x[i-k]
  }
  for (i in 1:n) {
    x[i] <- median(a[i:(i+2*k)])
  }
  out <- ts(data = x, start = 1, end = n)
}

#MR filter
meketon <- function (y, x, eps = 1e-04, beta = 0.97){
  f <- lm.fit(x,y)
  n <- length(y)
  w <- rep(0, n)
  d <- rep(1, n)
  its <- 0
  while(sum(abs(f$resid)) - crossprod(y, w) > eps)
  {
    its <- its + 1
    s <- f$resid * d
    alpha <- max(pmax(s/(1 - w), -s/(1 + w)))
    w <- w + (beta/alpha) * s
    d <- pmin(1 - w, 1 + w)^2
    f <- lm.wfit(x,y,d)
  }
```

```r
  list(coef = f$coef, steps = its)
}
mrfilt <- function (y, theta = 40){
  n <- length(y)
  V <- rep(0, n)
  V[1] <- 1
  V[2] <- -2
  V[3] <- 1
  T <- toeplitz(V)
  T[lower.tri(T)] <- 0
  m <- n-2
  D <- T[1:m,]
  X <- rbind(diag(n), theta * D)
  y_tilde <- c(y, rep(0,m))
  aux <- meketon(y_tilde,X)
  out <- ts(data = aux$coef, start = 1, end = n)
}
```

7.4.2 Python code

```python
#HP filter
def hpfilt(y, vlambda=1600):
    C = [0]*n
    C[0] = 1
    R = [0]*n
    R[0] = 1
    R[1] = -2
    R[2] = 1
    T = toeplitz(c=C, r=R)
    D = T[:-2]
    DT = D.transpose()
    D2 = np.dot(DT, D)
    D3 = np.diag(np.full(n,1))
    D4 = vlambda*D2
    D5 = D3+D4
    A = np.linalg.inv(D5)
    x = A.dot(y)
    return x

#MED filter
```

```
def medfilter(y, k = 11):
    k = 11
    n = len(y)
    m = n + 2 * k
    a = [0]*m
    x=y
    for i in range(1, m+1):
        if (i <= k):
            a[i-1] = x[0]
        elif (i > (n+k)):
            a[i-1] = x[n-1]
        else:
            a[i-1] = x[i-k-1]
    for j in range(1,n+1):
        x[j-1] = statistics.median(a[j-1:(j+2*k)])
    return x

#MR filter
def meketon(y, x, eps = 1e-04, beta = 0.97):
    model = sm.OLS(y, x, missing='drop')
    model_result = model.fit()
    res= sum(abs(model_result.resid))
    n=len(y)
    w = np.array([0]*n)
    d = [1]*n
    its = 0
    cp = y.transpose()
    while (sum(abs(model_result.resid)) - np.dot(cp, w) > eps):
        its = its + 1
        s = model_result.resid * d
        alpha = max(np.maximum(s/(1-w), -s/(1+w)))
        w = w + (beta/alpha) * s
        d = np.square(np.minimum(1 - w, 1 + w))
        mod_wls = sm.WLS(y, x, weights=d)
        model_result = mod_wls.fit()
        coefs= [model_result.params]
        lista=[[model_result.params], [its]]
    return coefs
def mrfilter(y, theta = 40):
```

```
    n = len(y)
    C = [0]*n
    C[0] = 1
    R = [0]*n
    R[0] = 1
    R[1] = -2
    R[2] = 1
    T = toeplitz(c=C, r=R)
    D = T[:-2]
    D1 =  np.diag(np.full(n,1))
    D2 = theta * D
    X = np.concatenate((D1, D2), axis=0)
    m = n-2
    y_tilde = np.concatenate((y, [0]*m), axis=None)
    aux = meketon(y_tilde,X)
    return aux
```

References

Altissimo, F., Marchetti, D. J., and Oneto, G. P. (2000). The italian business cycle: Coincident and leading indicators and some stylized facts. *Giornale degli Economisti e Annali di Economia*, 60(2):147–220.

Assunção, J. B. and Fernandes, P. A. (2017). A Robust Estimation of the Portuguese Real Business Cycles. *NECEP Research Papers*, 1:1–14.

Assunção, J. B. and Fernandes, P. A. (2022). Robust filtering with quantile regression. *NECEP Research Papers*, 2:1–15.

Baxter, M. and King, R. G. (1995). Measuring business cycles: Approximate band-pass filters for economic time series. *NBER Working Paper Series*, (5022):2–16.

Bergmeir, C., Hyndman, R. J., and Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics Data Analysis*, 120:70–83.

Burns, A. F. and Mitchell, W. C. (1946). *Measuring Business Cycles*. National Bureau of Economic Research, New York, 2 edition.

Cochrane, J. (1997). *Time Series for Macroeconomics and Finance*. University of Chicago Press, Chicago.

Cogley, T. and Nason, J. M. (1995). Effects of the hodrick-prescott filter on trend and difference stationary time series implications for business cycle research. *Journal of Economic Dynamics and Control*, 19(1):253–278.

Enders, W. (2014). *Applied Econometric Time Series*. John Wiley and Sons, New York, NY, 4 edition.

Guay, A. and St-Amant, P. (1996). Do mechanical filters provide a good approximation of business cycles? *Bank of Canada Technical Reports*, (78):1–35.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, New Jersey, 1 edition.

Hamilton, J. D. (2017). Why you should never use the hodrick-prescott filter. *The Review of Economics and Statistics*, 100(5):831–843.

Harding, D. and Pagan, A. (2002). Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics*, 49(2):365–381.

Harvey, A. C. and Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–247.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021). *An Introduction to Statistical Learning with Applications in R*. Springer, New York, NY, 2nd edition.

Kim, S.-J., Koh, K., Boyd, S., and Gorinevsky, D. (2009). $\ell 1$ trend filtering. *SIAM Review*, 51(2):339–360.

Knight, D., Knight, B., Pearson, M., Quintana, M., and Powell, B. (2018). *Microsoft Power BI Complete Reference: Bring your data to life with the powerful features of Microsoft Power BI*. Packt Publishing Ltd., Birmingham, 1 edition.

Koenker, R. (2008). Quantile regression computation:from outside, inside and the proximal. `http://www.econ.uiuc.edu/~roger/research/rq/QRComputation.pdf`.

Koopmans, T. (1947). Measurement without theory. *The Review of Economics and Statistics*, 29(3):161–172.

Kydland, F. E. and Prescott, E. C. (1990). Business cycles: real facts and a monetary myth. *Quarterly Review*, 14(2):3–18.

Lucas, R. E. (1977). Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy*, 5:7–29.

Meketon, M. S., Vanderbei, R. J., and Freedman, B. A. (1986). A modification of karmarkar's linear programming algorithm. *Algorithmica*, 1:395–407.

Mohr, M. (2005). A trend-cycle(-season) filter. *European Central Bank Working Paper Series*, (499):4–35.

Mosheiov, G. and Raveh, A. (1997). On trend estimation of time-series: A simpler linear programming approach. *The Journal of the Operational Research Society*, 48(1):90–96.

Phillips, P. C. B. and Jin, S. (2021). Business Cycles, Trend Elimination, And The Hp Filter. *International Economic Review*, 62(2):469–520.

Ravn, M. O. and Uhlig, H. (2002). On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. *The Review of Economics and Statistics*, 84(2):371–376.

Rua, A. (2017). Datação dos ciclos económicos em portugal. *Banco de Portugal Artigos Científicos*, 1:1–17.

St-Amant, P. and van Norden, S. (1997). Measurement of the output gap: A discussion of recent research at the bank of canada. *Bank of Canada Technical Reports*, (79):1–75.

Wade, R. (2020). *Advanced Analytics in Power BI with R and Python: Ingesting, Transforming, Visualizing*. Apress, Indianapolis.

Wen, Y. and Zeng, B. (1999). A simple nonlinear filter for economic time series analysis. *Economics Letters*, 64(2):151–160.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach*. South-Western College Pub, Ohio, 5 edition.

Zhao, S. and Wei, G. (2003). Jump process for the trend estimation of time series. *Computational Statistics Data Analysis*, 42(1):219–241.