Final report

Changes in investment market regimes in the post-Covid era

MS-E2177 Seminar on Case Studies in Operations Research
May 30, 2023

Tuomas Myllymäki (Project Manager)
Roosa Ilvonen
Aake Kesälä
Konsta Parkkali
Lassi Ruoppa

Client: Veritas

Contents

1 Introduction

1.1 Background

Our client Veritas is a Finnish occupational pension insurance company that offers statutory occupational pension insurance (TyEL and YEL) solutions for employees of companies and entrepreneurs in the private sector. Veritas is responsible for the pension security of more than 116,000 people (2022) and pays out more than 600 million euros in pensions annually. Veritas invests the funds it collects as pension contributions for future pensions globally in a diversified way with ∼32% in fixed income investments, ∼39% in equity investments, ∼16% in real estate, and the rest in alternative investments. [1]

The influence of the macroeconomic conditions on investments has gained interest among investors in recent times. Especially the wake of global crises such as the Covid-19 pandemic and the financial crisis of 2008 has sparked fear about the resulting downturn and the effects it may have on the investment scheme. The dynamic nature of financial markets, characterized by regimes of fairly persistent market conditions, underlines the need for investment companies, such as pension funds, and investors to gain a deeper understanding of past crises' impacts on markets and how different asset classes perform in the ever-changing world.

Veritas aims to make secure investments that guarantee solvency and long-term profitability, irrespective of prevailing economic conditions. Veritas focuses on constructing portfolios that generate the best possible returns, ensuring sustainable funding for pensions. They aim to develop a well-yielding robust portfolio that can withstand market uncertainty.

Pension funds, such as Veritas, are interested in understanding the regime switching in asset class returns dependence structures, e.g., correlation structures, as it can significantly impact portfolio performance and make informed decisions about asset allocation, diversification, risk management, and performance. [5, 30]

1.2 Objectives

Our case study aims to gain insight into the evolution of the asset class dependence structure over time. Our main objective is to give scientifically founded suggestions on improving the portfolio construction process based on our analysis and findings of historical asset class dependence regimes. Additionally, we seek to provide our client with a perspective on how to incorporate various macroeconomic conditions into portfolio construction decisions. We aim to develop a methodology for detecting diverse asset class dependence regimes through analysis of historical time series data including multiple asset classes, such as stock indexes, commodities, and macroeconomic indicators such as inflation.

We investigate the inter-dependencies of asset class returns to identify presence of the regime clustering over time. In addition, our study seeks to assess the effectiveness of various clustering techniques and the explanatory capacity of an extensive range of assets and economic variables, with the ultimate goal of developing a robust model for regime identification. Finally, we aim to estimate the asset class risk parameter values in different regimes and explore the conditional tail behavior.

1.3 Report outline

The rest of this report is organized as follows. Chapter 2 presents a review of the relevant literature. Chapter 3 provides an overview of the data used in this study, along with the preprocessing steps that were taken. The methodology used in this study is described in Chapter 4. Chapter 5 presents the results and main findings, which are discussed in more detail in Chapter 6. Finally, Chapter 7 concludes.

2 Literature review

2.1 Stock-bond correlation

Asset class returns are non-stationary and show volatility clustering [28]. In the case of asset returns, volatility can be defined as the rate at which the price of a security increases or decreases for a set of returns. It is measured as the standard deviation of the annualized returns for a given period. Equity volatility is a likely trigger for stocks and bonds moving in opposite directions as when volatility rises, investors tend to move away from the riskier investments, i.e., stocks, to the safer investments, i.e., bonds due to the uncertainty. [6]

The stock-bond correlation is of interest to investors and financial analysts as it provides insight into the behavior of two key asset classes in financial markets. This correlation can have significant implications for e.g., portfolio diversification and risk management. When the correlation is high, it may be challenging for investors to diversify portfolios and thereby spread risk.

Before the 2000s, the stock-bond correlation remained mainly on the positive side and as a result, equities have served as the dominant provider of returns in many portfolios. Bonds have been diluting equity risk and, in addition, have been able to deliver positive returns when equity markets have suffered losses [6]. However, the relationship has taken a turn in the last 20 years, as the correlation has turned negative, as seen in Figure 1.

The literature suggests that stock-bond correlations exhibit frequent fluctuations that can be discerned through analysis of historical data over short time horizons. Nonetheless, such short-term oscillations do not appear to have any sustained implications on the long-term stock-bond relationship, which has remained mostly negative since the early 2000s. [33]

The expected stock and bond returns can be written as

$$P_{stock} = E[\sum_{t=1}^{\infty}(\frac{1+G}{1+Y_t+ERP_t})^{\cdot}D] \tag{1}$$

and

$$P_{bond} = E[\sum_{t=1}^{T}\frac{C_t}{1+Y_t}+\frac{100}{(1+Y_T)^T}] \tag{2}$$

where G is the expected growth rate of dividends D , Y_t is the expectations of future short-term rates and the required bond risk premium, ERP_t is the required equity risk premium embedded in the discount rate, and C_t is a fixed cash flow coming from coupon payments and a par value 100 of a bond. [33]

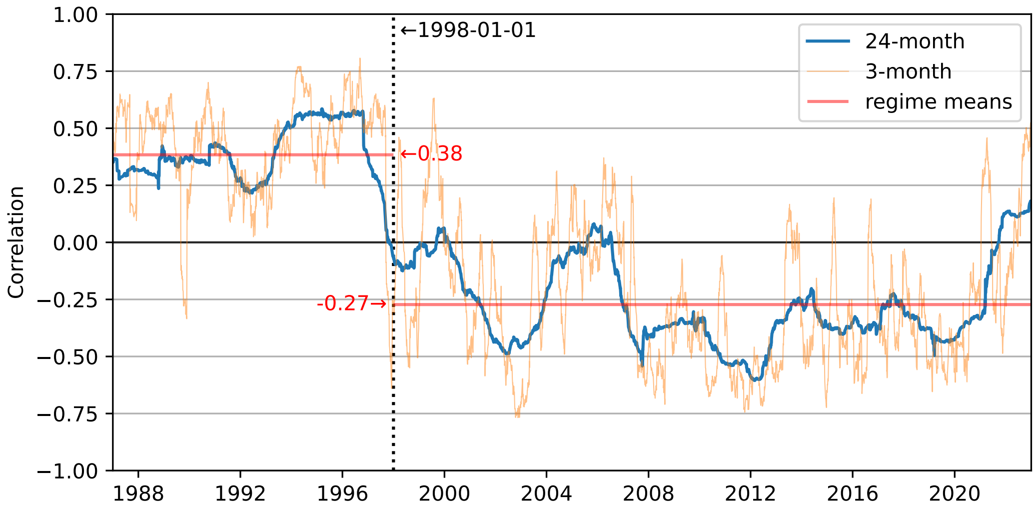

Figure 1: Rolling stock-bond correlation (SPX - G4O2) with 3-month and 24-month rolling windows.

Equations (1) and (2) reveal that both the returns of stocks and bonds exhibit common underlying factors that result in correlated movement. Inflation, for instance, can cause short-term correlation to increase by influencing short-term interest rates, which impacts both asset returns. Apart from volatility and inflation, other macroeconomic factors, such as growth news, may also contribute to changes in the stock-bond correlation. These factors may create discrepancies between the performance of stocks and bonds, thus accounting for cyclical fluctuations in correlation regimes. [33]

Prior research has identified core inflation, inflation uncertainty, equity volatility, treasury yields, the output gap, unemployment, and growth as the most significant variables explaining correlation trends when examining historical data spanning from 1950 to 2021 [33, 16]. For example, the random forest algorithm has been shown to be useful when identifying the explanatory variables from multivariate macroeconomic data [33]. Among these variables, inflation stands out as the most influential driver of the correlation. Positive inflation shocks drive positive correlations, whereas negative shocks are associated with a decline in correlations. However, the impact of correlation regimes on long-term portfolio returns appears to be minimal, as evidenced by previous studies [33]. Moreover, over the long-term, trends in growth, inflation, and interest rates may exert similar influences on the stock-bond relationship, resulting in a positive correlation, despite short-term fluctuations, such as growth resulting in a negative correlation in the near term [16].

2.2 Regime switching

Historical evidence indicates that financial markets exhibit evolving patterns of behavior over time. The changes often persist for a longer period. Although the financial markets can experience abrupt switches in response to, for example, sudden shock situations, the patterns of change are typically long-term and are the result of slower-moving forces as discussed in Section 2.1. Regime-switching models are specifically developed to identify and capture the changes in market behavior based on different macroeconomic variables. [3]

Regime switches in the financial markets often coincide with changes in, e.g., regulation and policies. The macroeconomic environment typically responds to changes by adjusting interest rate behavior, which can later be identified and captured using a regime-switching model.

By combining information about the macroeconomic scheme and how the market has reacted, regime-switching models can provide useful knowledge on the more complicated processes driving security returns. [3] Market regime switches can be detected by analyzing the statistical properties of the market data time series, such as the changes in market volatility. When a significant volatility change occurs, it can indicate a regime switch. [9]

To investigate the dynamics of the stock-bond correlation and regime changes in financial markets, it is necessary to select appropriate variables. One possible method for identifying the most influential factors in historical stock-bond relationship changes is by utilizing the random forest algorithm or the principal component analysis (PCA). Previous studies suggest that decision tree models, i.e., random forest, allow the effective ranking of important macro determinants in explaining the changes in the correlation regimes well. [33]

The combination of unsupervised clustering techniques, specifically k-means and hierarchical clustering, with supervised classification algorithms like the random forest, has been suggested to enhance the accuracy and robustness of detecting regime switches. This hybrid approach has been evaluated using actual financial data and compared to other commonly used techniques, including hidden Markov switching models and autoregressive models. The outcomes reveal that the hybrid method surpasses the other techniques in identifying regime switches in financial markets in terms in terms of accuracy and stability. [2]

The literature also suggests that a mixture of normal distributions econometric model that captures nonlinear dynamics in the joint distribution of stock and bond returns can be an effective approach. The proposed model is based on a Markov switching framework that allows for regime switches in the correlation between the stock and bond returns and it has been tested on empirical data and compared to other commonly used models, e.g., constant correlation models. The results show that the mixture model provides a better fit for the data to capture the changes in the correlations between asset returns over time and thus outperforms the traditional linear models in capturing nonlinear dependencies in stock-bond returns. [13]

One way to model univariate returns series is by using a simple Markov switching process, where the parameters are determined by a state variable unique to the asset in question S_{it} taking values s_{it} = *1, ..., k_i*, where k_i is the number of states for the ith series

$$y_{it} = \mu_{is_{it}} + \sum_{j=1}^{p} a_{ij,s_{it}} y_{it-j} + \sigma_{is_{it}} u_{it}, \quad i = 1, ..., n, \quad u_{it} \sim IIN(0,1) \tag{3}$$

where state transitions are determined by a fixed transition probability matrix

$$P(S_{it} = s_{it} | S_{it-1} = s_{it-1}) = p_{s_{it} s_{it-1}}, \quad s_{it}, s_{it-1} = 1, ..., k_i.$$

In the proposed model, each regime is considered to be a first-order homogeneous, irreducible, and ergodic Markov chain. [13]

2.3 Tail dependency

Tail dependency analysis using copulas, a function that explains the link between a multivariate distribution function and its one-dimensional marginal distribution functions [22], has gained interest among financial researchers and investors. The upper and lower tail dependence coefficients explain the likelihood that two extreme ends of events occur at the same time [29]. Knowing such probabilities acts useful when planning on how to efficiently allocate resources in portfolio construction. For example, being able to estimate if two or more assets are going to crash simultaneously, is likely to have a great value in the diversification process.

The tail dependence can be described as follows. Let X and Y denote two random variables with a joint distribution $F_{X,Y}(x,y)$ and continuous marginal distribution functions F_X and F_Y. Formally, the copula can be defined as:

$$\mathbf{C}(u,v) = P(U \leq u, V \leq v), \;\; U \equiv F_X(X), \; V \equiv F_Y(Y),$$

where u_X, $u_Y \in [0,1]$. The upper dependence distribution coefficient can be defined as:

$$\lambda_U \equiv \lim_{u \to 1^-} P(Y > F_Y^{-1}(u) | X > F_X^{-1}(u)) = \lim_{u \to 1^-} \frac{1 - 2u + C(u,u)}{1-u} \in [0,1],$$

provided the limit exists. Similarly, the lower tail dependence coefficient can be defined as:

$$\lambda_U \equiv \lim_{u \to 1^+} P(Y \geq F_Y^{-1}(u) | X \geq F_X^{-1}(u)) = \lim_{u \to 1^+} \frac{C(u,u)}{u}. \; [27,\ 29,\ 12]$$

Previous research suggests that dynamic lower tail dependence estimators, such as the Patton model [23] and DCC model (Dynamic Conditional Correlation) [11], outperform static estimators, such as a mixture copula model. It has also been recognized that if the sample size is small, the mixture copula model may perform better than some dynamic models. [29]

When studying the tail dependency structure with copulas, it is important to know the potential pitfall that arises from using an unsuitable model in the analysis. In such case, the obtained results from the analysis may not be reliable and should not be utilized in further analysis in e.g., portfolio construction. [29, 12]

3 Data and preprocessing

Our client Veritas provided us with a data set consisting of 88 economic time series for identifying financial regimes. The data set consisted of a wide variety of economic factors from the following categories: 35 macroeconomic indices, such as consumer confidence (e.g., CONCCONF Index), inflation (e.g., CPI YOY Index), and economic surprise (e.g., CESIUSD Index), 12 interest rates indices, such as G4O2 Index and HE00 Index, 23 stock market indices, such as SPX index and NDWUENR Index, 3 currency indices, such as EURUSD

Index, 6 commodities indices, such as corn and copper, 5 private market indices, such as the Preqin private equity index, and 4 equity indices. For the full list of economic indices in the data set, the reader is referred to Appendix A. The given data was of mixed types of observations, including year-on-year changes, monthly changes, and pure values. Furthermore, the time series in the data set spanned different time periods, which added another layer of complexity to the analysis. For this reason, somewhat heavy pre-processing of the given data was necessary.

In addition to each time series in the data set spanning a different time period, the frequency of observations also differed. In total, the data contained 6 different frequencies: yearly, quarterly, monthly, biweekly, weekly, and daily. We chose to use all data in a monthly format, as the majority of the original data set was monthly to begin with. As such, using monthly data provided a good balance between preserving the information value contained in higher frequency data while also having as much usable data as possible. Time series with a frequency higher than monthly were down-sampled to monthly by calculating the monthly mean of the data points. On the other hand, quarterly data was up-sampled to monthly using linear interpolation. All time series with a frequency lower than quarterly were discarded during preprocessing.

3.1 Feature transformations

Another issue with the original data set was that the drastically different scales of the time series made comparison difficult. For example, the NAPMPMI Index was a series with values normalized to be within the interval [0,100], while USURTOT Index, which represents the US unemployment rate, is given as a percentage. Furthermore, data such as stock indices present another issue, as the general trend in the data is upward. As such, before we could use the data for conducting any analysis, we had to perform some feature transformation. Here by feature, we refer to the individual economic time series.

As we were not very familiar with the different features, we decided to consult our client and their experts on the matter. The client provided us with more information on the different time series in the data set and their suggestions for possible feature transformations to perform and test. The suggestions for feature transformations are mostly based on the client's expert opinions and intuition and thus most of them we cannot justify with scientific literature.

For macroeconomic indices, Veritas generally suggested using either a rolling mean or a percentage change to capture the trend depending on the index. On the other hand, for features that represent monetary assets, such as the stock market and currency-related features, Veritas suggested calculating the log return of each observation. The log return is given as

$$\text{log-return} = \log \frac{x_{k+1}}{x_k},$$

where x_k is the observation in the underlying series at time k. Log return is commonly used in economic literature to normalize for example time series that represent stock prices [21]. Finally, the scale of each feature was normalized by computing the z-score of each observation

$$z\text{-score} = \frac{x - \bar{x}}{s},$$

where $\bar{x}$ and s are the sample mean and sample standard deviation of the underlying series respectively [20]. For a full list of feature transformations considered for each time series in the data set, the reader is referred to Appendix A. In addition to the transformations in the Appendix A, we also applied rolling mean with various window sizes to the variables in order to catch longer regimes and to reduce noise. For the detailed lists of all different combinations of variables and rolling window sizes used, see Appendix B.

3.2 Correlation pairs

Our client had originally stated that out of the 82 economic indices in the data, 39 were of interest in terms of correlation. However, the number of unique correlation pairs that could be formed from the 39 different time series is more than 700. As such, to narrow the scope of our analysis, we asked the client to provide a list of at most 10 correlation pairs that were of specific interest. The correlation pairs in question are listed in Table 1. These pairs were used for conducting all of the analyses presented in the following sections of this report. Rolling mean with varying window sizes was also applied to the computed correlations in order to reduce noise and to reveal long term trends.

Table 1: List of correlation pairs of specific interest provided to our team by Veritas.

First index	**Second index**
G4O2 index	NDDUWI Index
G4O2 index	EURUSD index
EG04 index	NDDUWI Index
EG04 index	EURUSD index
H0A0 index return	NDDUWI Index
H0A0 index return	EURUSD index
NDDUWI Index	EURUSD index

4 Methods

Based on the literature review, the client's view, and prior knowledge, PCA and Random Forest are selected as the methods of dimension reduction, and k-means, DBSCAN, and Gaussian mixture are selected as the clustering methods.

4.1 Markov switching models

Markov switching models are a type of statistical model that can be used to analyze time series data when the underlying data-generating process changes over time. They allow for the possibility that the relationship between variables, such as the correlation structure between different asset classes, may change over time.

Markov switching models are based on the principles of the Markov chain, which is a stochastic process used to describe how uncertain and unobserved outcomes occur. In the

Markov chain, the probability of moving from one state to another depends only on the current state and not the past, i.e., the Markov property. In the models, the underlying process that generates the data is assumed to be a Markov process that can take on a finite number of states. Each state is associated with a different statistical distribution, e.g., normal distribution. Then, the data in each state is generated from the corresponding distribution. [17, 14]

The state transitions are controlled by a set of transition probabilities, which determine the probability of moving from one state to another. The probabilities are represented as a transition matrix, where each element represents the probability of transitioning from one state to another. [17]

The parameters of the Markov switching model can be estimated using maximum likelihood estimation or Bayesian estimation. Maximum likelihood estimation involves finding the parameters that maximize the likelihood function, or the probability of observing the data given the model parameters. It utilizes an iterative algorithm known as expectation maximization. Bayesian estimation relies on drawing samples from a joint distribution of the parameters, states, and transition probabilities. [17]

Markov switching models and models that are derived from Markov switching principles have been explored in previous studies on the use of finding regime switches in financial markets. They can be used to analyze the relationship between, for example, stocks and bonds and identify different states of their joint distribution [13, 14, 8], which is why we decided on implementing and testing a Markov switching model described by Equation (3) in this study.

4.2 Dimensionality reduction

In the analysis of financial markets and regime switches, high-dimensional data is commonly used making the analysis complex. To address this challenge, we utilized dimensional reduction techniques for efficient clustering and regime identification. After considering previous research and the client's perspective, the principal component analysis (PCA) and random forest were selected as methods for dimension reduction. The idea with dimension reduction is to transform a high-dimensional dataset into a new dataset with fewer variables while minimizing the loss of information.

4.2.1 Principal component analysis

We decided to implement PCA as a technique for dimensionality reduction before clustering the data due to its success in the context of regime switching in financial markets based on previous research [2]. Another property of PCA is also that it can reveal underlying commonality between the variables in a dataset. For example, the growth of different equity indices are most likely driven by a few common factors, which can be analyzed using PCA.

PCA is a method that identifies linear combinations of the original variables that maximize the variance in the data that is explained by the transformed variables. The resulting transformed variables are called the principal components, with the first one capturing the most variance in the data. The following principal component captures the largest amount of the remaining variance, and so on. By selecting the principal components that explain the data the most and discarding the rest, the original dataset's dimensionality can be reduced effectively. [15]

The PCA transformation is an orthogonal linear transformation that transforms a vector of the original variables into a new coordinate system. Given a dataset of n observations and p variables with finite mean μ and finite covariance matrix Σ, we can represent the data as a data matrix $X = [x_1, x_2, ..., x_p]$, where x_i is a column vector represents the ith variable. The principal component transformation is now done as

$$x \to y = \Gamma^T(x - \mu),$$

where $\Gamma \in R^{p\times p}$ is orthogonal, $\Gamma^T\Sigma\Gamma = \text{diag}(\lambda_1, \ldots, \lambda_p)$ and $\lambda_1 \geq ... \geq \lambda_p$ are the eigenvalues of Σ in order from the largest to smallest. Now, we can choose the first k components as our new dataset with reduced dimensionality as this maximizes the variance of the data in k dimensions. [15]

4.2.2 Random forest

Random Forests are supervised learning algorithms used both in classification and regression problems. As an approach, they construct a large number of decision trees on a randomly selected subset of the original data set. Then, it combines the results of the individual trees to make predictions. The algorithm works well on high-dimensional data sets with noisy features and is robust to missing data and imbalanced data sets.

The random forest can be used for dimensionality reduction. The idea behind this is to rank the importance of each input variable in the algorithm's decision-making process and use only the most important variables for further analysis.

In this study, we decided to implement the random forest for dimensionality reduction and feature selection, as it has been shown to work well in the context of regime switch detection. [33]

4.3 Clustering algorithms

Based on previous studies and the requirements of the project, k-means clustering, Gaussian Mixture models, and DBSCAN were chosen as the primary clustering techniques for regime detection. These methods have have shown to be effective in identifying market regimes [14, 33, 2, 13]. All of these methods were used together with the dimension reduction method random forest. We also performed clustering analysis using k-means on the principal components selected by random forest.

4.3.1 k-means clustering

K-means clustering is a simple unsupervised algorithm that uses input vectors without referring to known or labeled outcomes. K-means cluster analysis aims to partition n-measured observations of interest into k-homogeneous clusters. Clustering aims to find patterns in the unlabeled data set and cluster them together. [15]

The k-means algorithm is based on calculating distances from "centers". Using the k-means algorithm, each observation belongs to a cluster with the nearest mean. The idea is to define k centroids, one for each cluster, and place them as far away from each other as possible. First, the cluster's center point is determined at the beginning at random. Then, each data point is associated with its nearest centroid using the Euclidean distance spacing measurements. [15]

Once this is done, new k centroids are re-calculated as the clusters' centers resulting from the first steps. Afterward, a new binding is done between the data points and the new centroids. This "loop" is then repeated until no more changes are done, i.e., the centroids will not move anymore. [15]

The number of clusters used in the k-means clustering algorithm needs to be specified and it can be chosen by several methods. In this study, we implemented the elbow method to choose the optimal number of clusters. The elbow method is a visual method where the cost function value produced by different values of k, the number of clusters, is used. The value of k at which the improvement of the average distortion of the clusters declines the most is called the elbow and it is where dividing the data into further clusters should be stopped. The elbow method is expressed by the sum of squared error presented as:

$$SSE = \sum_{k=1}^{K} \sum_{i \in S_k} ||y_i - c_k||^2$$

with k as the number of clusters formed, c_i is the ith cluster, and x is the data present in each cluster. [15]

We performed clustering analysis using the k-means algorithm due to its success in previous research. It has been shown to work in the identification of different regimes, such as the financial crisis of 2008. [2]

4.3.2 Gaussian mixture models

We decided to implement the Gaussian mixture model for the regime clustering after reviewing the literature. A mixture of normal distributions in econometric models has been shown to capture the nonlinear dynamics of stock and bond returns well used in regime switch detection. [13]

Gaussian Mixture Models (GMM) are probabilistic models that represent a set of data points as a mixture of several multi-dimensional Gaussian distributions that best model the input dataset, i.e., it assumes that data points are generated by combining multiple Gaussian distributions with different means and variances. Each Gaussian in the mixture has a mean that defines its center, a covariance, that defines its width, and a mixing probability that defines how big or small the Gaussian function will be.

For n multivariate data, the probability density function of a Gaussian Distribution is

$$G(X|\mu, \Sigma) = \frac{1}{\sqrt{2\pi}|\Sigma|} \exp\left(-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu))\right),$$

where μ is the n dimensional mean vector and Σ is the $n \times n$ covariance matrix of the distribution. The probability given in a mixture of k Gaussian, where k is the number of distribution

$$p(x) = \sum_{j=1}^{k} \phi_j G(X|\mu, \Sigma),$$

where ϕ_j is the prior probability of the jth Gaussian

$$\sum_{j=1}^{k} \phi_i = 1 \text{ and } 0 \geq \phi_j \geq 1.$$

GMM has some advantages over other clustering algorithms, such as that it can handle data with complex and irregular shapes, and it can detect clusters of different sizes and densities [19]. However, similar to k-means, one must specify the number of clusters beforehand, which can be difficult in practice. Since, in the case of this study, the number of clusters decided based on the results of the elbow method, expectation maximization can be used to estimate the mixture model's parameters.

4.3.3 Density-Based Spatial Clustering of Applications with Noise

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is an unsupervised clustering algorithm that can be used to group data points based on their density in a given space. It does not require one to specify the number of clusters, and it can find clusters of arbitrary shapes and sizes as well as of varying densities. In addition to being able to detect varying data, it is also rather robust to outliers.

The DBSCAN algorithm clusters data together based on the local density of the data point. It requires only two parameters: *epsilon*, the radius of the circle that is to be created around each data point to check density, and *MinPts*, the minimum number of data points required inside said circle for it to be considered as a core data point. First, the algorithm finds all the neighbor points within epsilon using the Euclidean distance. Then, it classifies that data point into *core*, *border*, and *noise* based on the specified value of *MinPts*. If the neighboring number of points is less than the value of *MinPts*, then it is classified as *border*. If there are no other data points within *epsilon*, then the point is classified as *noise*.

In our implementation, we used a value of 0.5 for *epsilon* and a value of 3 as the minimum number of data points required inside the circle with radius *epsilon*.

One benefit of using DBSCAN in analyzing the interdependencies of different asset classes is its ability in detecting nonlinear relationships between the different variables. [13]

4.3.4 Ordering points to identify the clustering structure algorithm

The Ordering Points to Identify the Clustering Structure (OPTICS) algorithm is a density-based clustering algorithm similar to the DBSCAN algorithm. Like DBSCAN, OPTICS uses the density of a point in a dataset to determine and identify clusters instead of using and relying on a pre-determined number of clusters like the k-means. The OPTICS algorithm can be considered as an extension of the DBSCAN algorithm that addresses some of the limitations of the DBCSAN, such as the sensitivity to the choice of hyperparameters, discussed in Section 4.4. Like DBSCAN, OPTICS utilizes two parameters: *epsilon* and *MinPts*.

However, unlike DBSCAN, OPTICS takes also into consideration the points that are part of a more densely packed cluster: each cluster is assigned a core distance that describes the minimum reachability distance required for a point to be considered as a center of a cluster. [4]

4.4 Optimizing model hyperparameters and features

Markov switching models and all of the chosen clustering algorithms feature at least one or two hyperparameters for which we must choose a value. However, the optimal values for the hyperparameters are generally not immediately obvious based on the data. To optimize the hyperparameter values in each of our models, we employed a simple grid search. In grid search, we first select a finite set of reasonable values for each hyperparameter. In our case, we generally restricted the number of reasonable values between two to five values per hyperparameter. Subsequently, to determine the sets of reasonable hyperparameter values, the model is fitted for each possible parameter combination. The parameter combinations are then compared using some metric that describes how well the model fits the data. The metric may also include a term that penalizes more complex models to avoid overfitting. Finally, the parameter combination that yields the best results according to the metrics is chosen [18]. In some cases when the number of hyperparameters was small, we also relied on visual inspections of the clusters. This was done especially together with the PCA & k-means clustering method as also the number of features used in the clustering was small, and thus visual inspection was straightforward.

Since in the case of some features neither literature nor our client could provide definitive answers on what the most favorable feature transformation would be, we had to consider combinations of different feature transformations in addition to optimizing hyperparameters. In cases where the optimal feature transformation is uncertain, the choice is generally made between using the original value or computing a percentage change, and also possibly applying rolling average with a suitable window size. As such, we constructed three feature sets. In the first set, percentage change was used for all features for which it was considered a viable choice. In the second set, we used either a 12-month rolling mean or the original value of the variable. It should be noted that for many of the variables, the transformation was the same in both sets, as only one feature transformation was considered viable. In the third feature set the goal was to unveil longer trends. Thus a 5-year rolling average was applied to all time series starting from latest on January 1990 in order to have a sufficiently long timespan for the data. This feature set was constructed already during the explanatory data analysis and used as the only feature set with the PCA & k-means clustering method.

Inflation data was considered separately from the rest, as it often ranks among the most important economic descriptors based on both our exploratory data analysis and literature (see e.g., [33]). For inflation data, we tested the following three transformations: change measured in percentage and three and five-year trailing inflation, that is three and five-year rolling mean. All combinations of the inflation feature sets and the two feature sets described above were then considered in testing, that is, the grid search used for hyperparameter optimization was performed for all feature set combinations. Choosing an optimal combination of feature transformations was however only a concern in the case of clustering algorithms since Markov switching models only take one time series as an input. For our present problem, the input in question is the correlation between one of the seven correlation pairs introduced in Section 3.2.

4.4.1 Markov switching models

For the class of Markov switching models that we employed, there were a total of three hyperparameters to consider: the number of regimes k, the order of the autoregressive lag polynomial p, and whether the variance is constant or switches between regimes (boolean). The sets of reasonable values for k and p were selected to be $\mathcal{K} = \{2, 3, 4\}$ and $\mathcal{P} = \{1, 2, 3\}$ respectively. The set for the possible number of regimes follows from the fact that k is rarely higher than 3 in literature. In fact, $k = 2$ has been used extensively in the literature and remains a popular choice even today [24]. Similarly, the order of the autoregressive lag polynomial was constrained between 1 and 3 since values higher than 4 are rarely found in the literature. As an example, Guidolin and Timmermann used the values $p = 0$ and $p = 1$ [13], while $p = 4$ was used when Markov switching models were first introduced [14].

We adopted the Hannan-Quinn information criterion (HQC) for comparing the models fitted with different parameter combinations. The HQC is given as

$$\text{HQC} = -2L_{\max} + 2m \log(\log(n)),$$

where $L_{\max}$ is the maximum of the log-likelihood, m is the number of parameters in the model and n is the number of data points. The HQC is a criterion for model selection that provides an alternative to the Akaike information criterion (AIC) and Bayesian information criterion (BIC). HQC was chosen instead of the more popular AIC and BIC since it has been previously used for hyperparameter optimization of Markov switching models in literature [13]. Similar to AIC and BIC, a lower value of HQC is preferred.

4.4.2 Clustering algorithms

For the k-means model the only parameter that was optimized using grid search was the number of means. We tested seven different values for the number of means k, where $k \in \{2, 3, 4, 5, 6, 7, 8\}$. For Gaussian mixture models, there were two hyperparameters to optimize: the number of clusters k and the covariance type c_{type}. The sets of reasonable hyperparameter values used for grid search are listed in tables in Appendix D Table 31. For the DBSCAN method, Table 32 contains the hyperparameter values used in the gridsearch. Similarly in Table 33 for OPTICS.

All sets of reasonable clustering hyperparameters were chosen based on exploratory and visual data analysis, and the assumptions regarding the lengths and sizes of the regimes.

For comparing the clustering quality achieved with different combinations of hyperparameters, we employed a total of three different metrics for assessing the goodness of fit: silhouette coefficient, Calinski-Harasbasz index and Davies-Boulding index. The aforementioned metrics were chosen since a ground truth, which was not available, is not required for computing any of them. We chose to use multiple metrics to measure the goodness of fit as thoroughly as possible. Each clustering quality metric is discussed in more detail in the following paragraphs.

The first metric we used for model comparison was the silhouette score. For a data point $i \in \mathcal{C}_I$, that is datapoint i in the Ith cluster, the silhouette coefficient $s(i)$ is given as

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}. \tag{4}$$

Equation 4 $a(i)$ represents the average dissimilarity of data point i to all other data points in cluster $i \in \mathcal{C}_I$, defined as follows

$$a(i) = \frac{1}{|\mathcal{C}_I| - 1} \sum_{j \in \mathcal{C}_I, i \neq j} d(i, j),$$

where $|\mathcal{C}_I|$ is the cardinality of the cluster and $d(i, j)$ is some similarity metric. Similarly, in Equation 4, $b(i)$ is the minimum mean dissimilarity between i and any cluster $\mathcal{C}_J \neq \mathcal{C}_I$. This is computed as follows

$$b(i) = \min_{J \neq I} \frac{1}{|\mathcal{C}_J|} \sum_{j \in \mathcal{C}_J} d(i, j).$$

We note that from the definition of $a(i)$ it follows that $s(i)$ is only defined when $|\mathcal{C}_i| > 1$ and $s(i) = 0$ if $|\mathcal{C}_i| = 1$ [25].

We used the Euclidean distance as the similarity metric when computing the silhouette coefficients, that is

$$d(i, j) = \sqrt{(i - j)^2}.$$

By definition $-1 \leq s(i) \leq 1$. The higher the silhouette coefficient is, the further the data points in cluster $\mathcal{C}_I$ are from the other clusters. Correspondingly, $s(i) = 0$ indicates that the data point is essentially of the boundary of the adjacent cluster, while a negative silhouette coefficient implies that the data point may have been assigned to an incorrect cluster [25].

The computed silhouette coefficients can be used in two distinct ways. Firstly, we can plot the silhouette coefficients of each data point and assess the quality of the clustering using the plot. An example of a silhouette plot is presented in Figure 2. Secondly, we can compute a global silhouette coefficient by calculating the mean of all silhouette coefficients. The global coefficient can be used to represent the quality of the clustering across the entire dataset as a single value, which is often convenient [25].

The second metric used for comparing models fitted with different parameter combinations was the Calinski-Harasbasz index (CH index). The CH index is defined as the ratio of the between-cluster dispersion and the inter-cluster dispersion. For a data set E of size n_E that has been clustered into k clusters, the CH index is

$$\mathrm{CH} = \frac{\mathrm{tr}(B_k)}{\mathrm{tr}(W_k)} \times \frac{n_E - k}{k - 1},$$

where $\mathrm{tr}(B_k)$ is the trace of the between-cluster dispersion matrix and similarly $\mathrm{tr}(W_k)$ is the trace of the within-cluster dispersion matrix. The dispersion matrices are given as

$$W_k = \sum_{q=1}^{k} \sum_{x \in \mathcal{C}_q} (x - c_q)(x - c_q)^\top,$$

$$B_k = \sum_{q=1}^{k} n_q (c_q - c_E)(c_q - c_E)^\top,$$

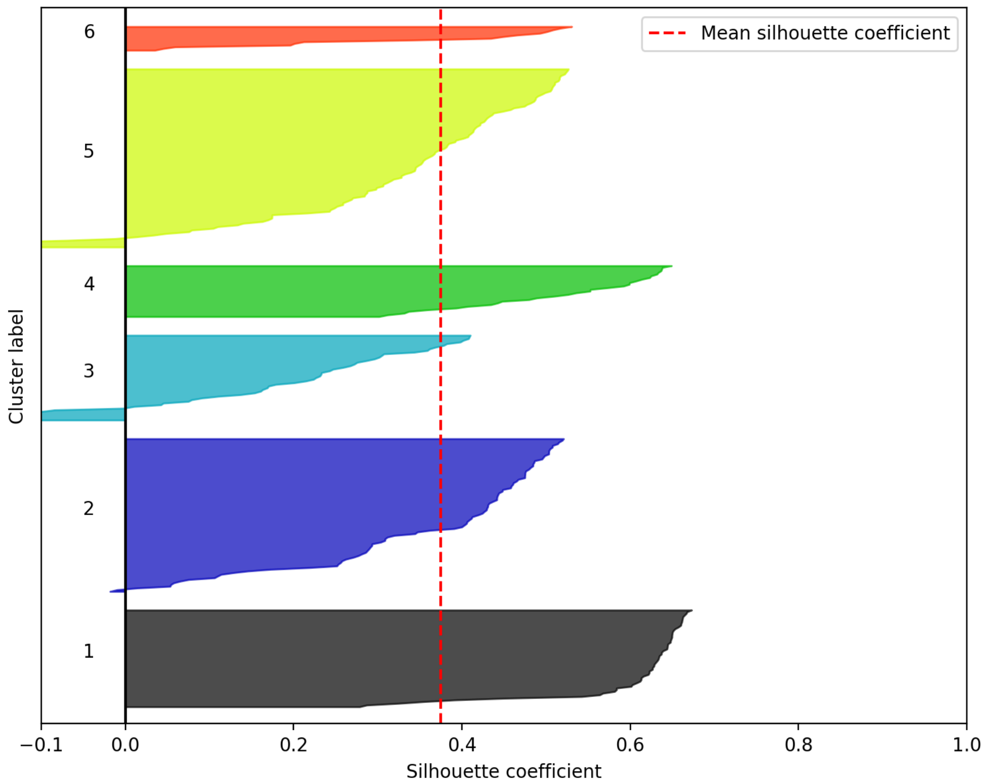

Figure 2: An example of a silhouette coefficient plot. The dashed red line denotes the mean silhouette coefficient across all data points. Note how some data points in clusters 2, 3, and 5 have a negative silhouette coefficient.

where $\mathcal{C}_q$ is the set of data points in cluster q, c_q is the center of the corresponding cluster, c_E is the center of the entire data set and n_q is the number of points in the qth cluster. A higher CH index indicates a better clustering quality [7].

The third and final metric we used for model comparison was the Davies-Boulding index (DB index). The DB index is

$$\mathrm{DB} = \frac{1}{k}\sum_{i=1}^{k}\max_{i\neq j} R_{ij},$$

where k is the number of clusters and R_{ij} is the similarity measure of clusters $\mathcal{C}_i$ and $\mathcal{C}_j$. Thus, the DB index measures the mean similarity between each cluster $\mathcal{C}_i$ and the cluster that is most similar to it. The similarity measure R_{ij} is

$$R_{ij} = \frac{s_i + s_j}{d_{ij}},$$

where s_i is the mean distance between each point in the ith cluster and its center and d_{ij} is the distance between the centers of clusters $\mathcal{C}_i$ and $\mathcal{C}_j$. Contrary to the silhouette coefficient and CH index, a lower value indicates a better clustering quality in the case of the DB index [10].

In addition to the performance metrics presented above, the Bayesian information criterion (BIC) was used when comparing Gaussian mixture models fitter with different parameter combinations. The BIC is formally defined as

$$\mathrm{BIC} = m \log(n) - 2L_{\max},$$

where $L_{\max}$ is the maximum of the log-likelihood, m is the number of parameters in the model and n is the number of data points [32]. The BIC could only be used for Gaussian mixtures since it is the only clustering algorithm among the algorithms we employed for which a likelihood function can be constructed.

5 Results

5.1 Markov switching models

We performed a grid search for each correlation pair using the hyperparameter values described in Section 4.4.1. The function `MarkovAutoregression` provided in the Python library statsmodels [26] was used for fitting each model and computing the Hannan-Quinn information criterion. Based on the HQC values, we identified the best model for each correlation pair. As an illustrative example, we discuss the results for the correlation pair G4O2 index, NDDUWI Index. The HQC values for each model for the aforementioned correlation pair are in Table 2. The best performing parameter combination in this case is $k = 2$, $p = 1$ with switching variance, although all of the HQC values are quite similar. Notably, using switching variance instead of constant variance appears to have the greatest effect on the HQC in terms of model hyperparameters. We conjecture that increasing the set of reasonable hyperparameter values is unlikely to yield better performing models, since the HQC appears to be increasing as k and p increase. Results for other correlation pairs were similar, thus we omit them for the sake of brevity.

Visually the regimes identified by the best Markov-switching models generally appear to be quite random, with regimes switching abruptly and for seemingly no reason. Furthermore, the regimes may even switch every month at times. None of the aforementioned are desirable properties when attempting to identify longstanding financial regimes that would preferably line up with historical events in the would economy to a certain degree. An example of the regimes given by the best model for the correlation pair G4O2 index, NDDUWI Index is presented in Figure 3. The regimes identified in other correlation pairs were visually similar and did not line up between different correlation pairs. The visual results thus suggest that Markov switching models are not suitable for identifying financial regimes in correlations.

Another problem we observed with Markov switching models was that many of the fitted parameters were not statistically significant even for the best models we found with grid search. This result further suggests that Markov switching models are a poor choice for the task at hand. As such, since the preliminary results for simple models were extremely poor, we decided to discard Markov switching models as a viable option for regime identification in correlations.

Table 2: Hannan-Quinn information criterion values for all models fitted to the correlation pair G4O2 index, NDDUWI Index. In the table, the best performing model has been highlighted.

k	p	**Variance**	**HQC**
2	1	Constant	-274.34
	2		-241.35
	3		-236.24
3	1	Constant	-274.96
	2		-269.81
	3		-264.63
4	1	Constant	-254.95
	2		-244.49
	3		-248.05
2	1	Switching	**-312.18**
	2		-275.25
	3		-269.33
3	1	Switching	-306.62
	2		-283.90
	3		-272.82
4	1	Switching	-293.73
	2		-269.79
	3		-278.23

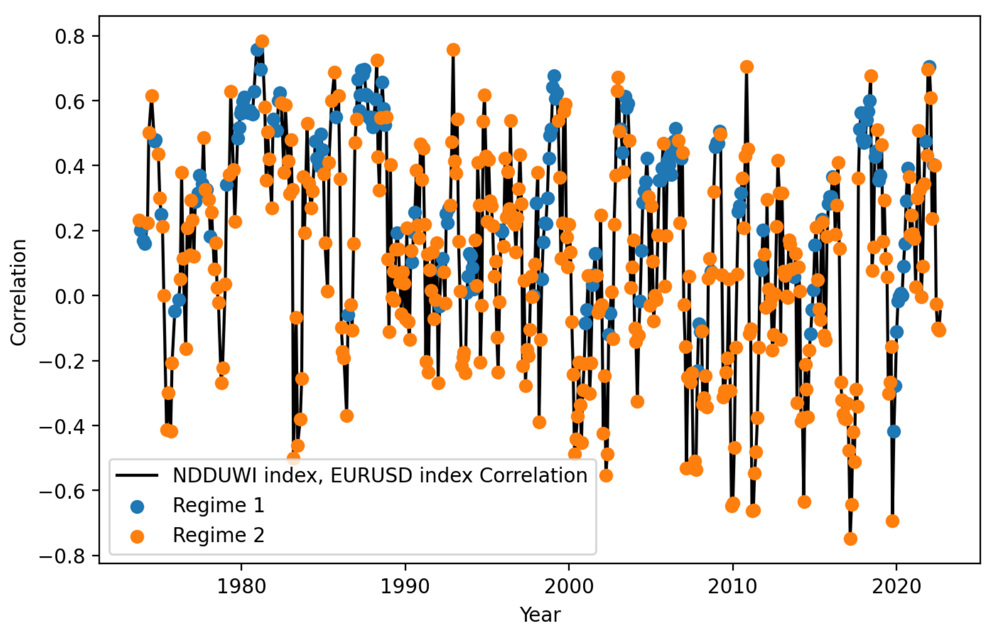

Figure 3: The best Markov switching model found with grid search for the correlation pair NDDUWI index, EURUSD index. The hyperparameters of the model are $k = 2$, $p = 1$ and switching variance.

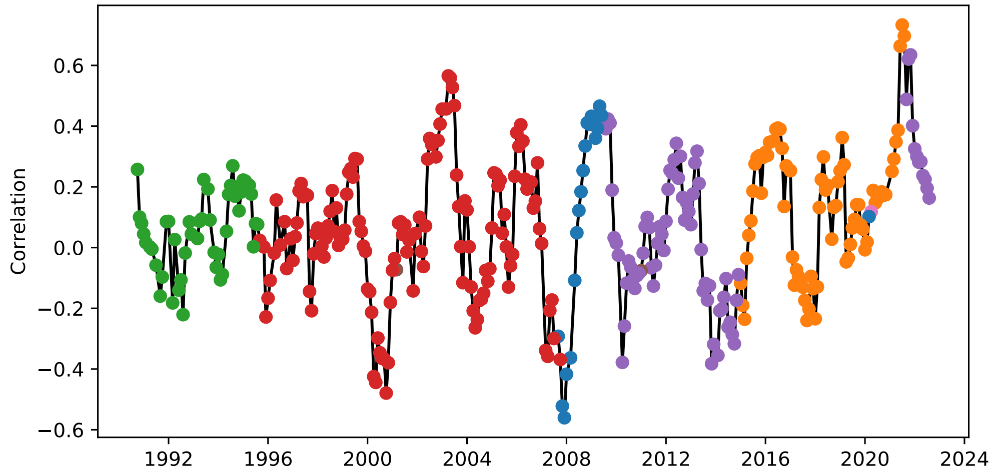

Figure 4: Rolling correlation of EG04 index and NDDUWI index clustered with k-means using $k = 7$.

5.2 Random forests combined with clustering algorithms

5.2.1 k-means clustering with more intricate feature transformations

The performance of the k-means model was tested with 2–8 means for the different correlation pairs and features. Based on performance metrics and visual inspection the models with 7 means performed the best with most correlation pairs. With the 7 means, the model was able to identify multi-year time intervals that, based on visual inspection, followed macro economic events such as the 2008 financial crisis.

The features used for fitting the model are listed in the appendix. Based on the results of the random forest regression the most important features always included inflation metrics regardless of the correlation pair. The inflation features include, for example, trailing inflation and rolling standard deviation of the yearly change of the inflation level. The rolling standard deviation of the inflation level was the only "more intricate" data transformation that made it through the random forest dimension reduction.

The clustering for the rolling correlation of EG04 index and NDDUWI index is shown in Figure 4. The clusters form clear time intervals whose start and end can be associated with some macro economic event. It is important to note that the models do not necessarily re-enter earlier clusters if some exceptions are not taken into account.

5.2.2 Gaussian mixture models

The hyperparameter combination that yielded the best performing GMM model for all correlation pairs was $k = 6$, $c_{\text{type}} = \text{full}$. The model in question performance of the model was quite good both in terms of the clustering quality metrics and visual cluster quality. The fact that the best performing model was the same for each correlation pair is an interesting result in itself, considering that correlation acts quite differently between different pairs.

However, the result is also quite convinient, since it saves a lot of time when it comes to further analysis and model comparisons.

While models with $k \in \{4, 5, 6\}$ and any type of covariance performed similarly in terms of most performance metrics aside from the DB index, the model with $k = 6$ clusters and $c_{\text{type}} = full$ consistently performed the best in terms of the Bayesian information criterion. Furthermore, the BIC was generally considerably lower for all models with $c_{\text{type}} = full$ compared to any other covariance type.

Visually the regimes identified by the model appeared reasonable for the best model. Clusters form regimes that span multi-year time periods when overlayed with the underlying correlation and conversely to Markov switching models the regime changes appear reasonable. An example of the clusters for the correlation pair G4O2 index, NDDUWI Index is shown in Figure 5.

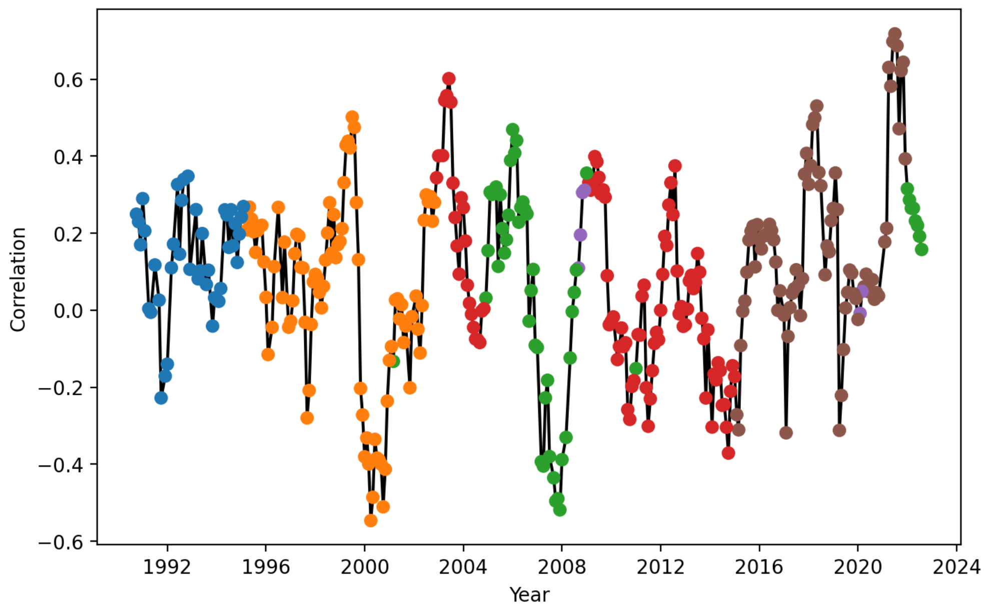

Figure 5: Clustering formed by the best GMM model for the correlation pair G4O2 index, NDDUWI Index.

The features used for fitting the best model and the most important features for each correlation pair according to random forest regression are listed in the appendix. Notably, inflation metrics appeared to have a significant effect on the model, and using a 5 year trailing inflation stabilized the model greatly in comparison to either 1 or 3 year trailing inflation. Furthermore, using the percentage change transformation for the appropriate features appeared to give visually slightly better results.

5.2.3 DBSCAN and OPTICS

We also tested DBSCAN and OPTICS models. Since OPTICS is a modified version of DBSCAN, we can consider the two models together. The best perfoming hyperparameters changed depending on the correlation pair considered and for one correlation pair OPTICS

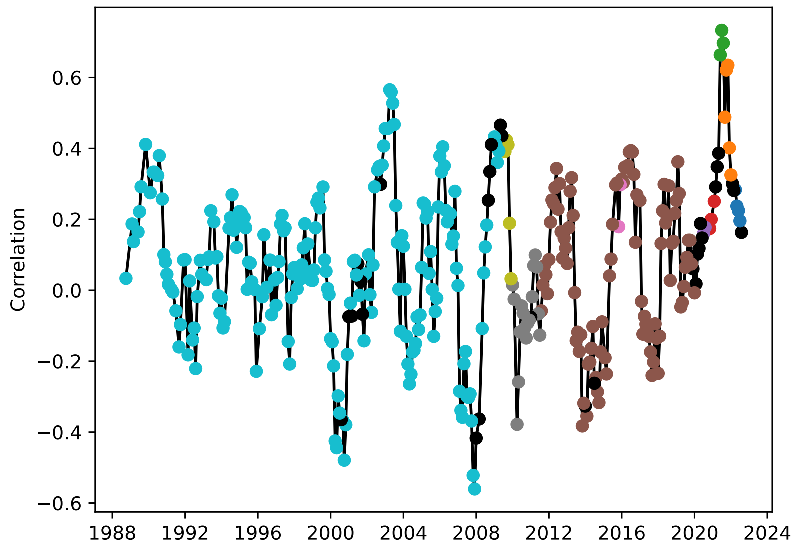

Figure 6: Rolling correlation of EG04 index and NDDUWI index clustered with DBSCAN. Note that black circles represent noise points. $\epsilon = 1.4$ and minimum sample count of 2

performed better than DBSCAN. However, for 6 out of 7 correlation pairs, DBSCAN outperformed OPTICS when considering performance metrics described before. Table 33 and Table 32 contain the hyperparameters used in the grid search for OPTICS and DBSCAN.

While the best hyperparameters changed for correlation pairs, for the majority of them the best hyperparameters were $\epsilon = 1.2$ or $\epsilon = 1.4$ and the minimum number of datapoints being 2, using DBSCAN approach. This suggests that the grid search did find some optimal value for some majority subset of the correlation pairs. Nonetheless there was at least one outsider being the correlation pair between the NDDUWI index and EUR/USD rate. For this pair, the optimal model was DBSCAN with $\epsilon = 0.6$ with minimum sample size of 2. Also for H0A0-NDDUWI index pair, the better model was OPTICS.

Overall, DBSCAN and OPTICS methods do not perform well. The clusters change quite rapidly and there is often a significant amount of datapoints labelled as "noise". We can however get some reasonable clusters as depicted in Figure 6. We will omit this model from further analysis.

5.3 Principal component analysis combined with k-means clustering

We decided to take a slightly different approach in finding the regimes with the PCA & k-means clustering method. The idea was to search for longer regimes by using 5-year rolling means of the chosen variables shown in Appendix B. We also used a longer, 4-year rolling correlation to obtain less noisy and more steady correlation data. Yet another difference was that we included the correlation in the final clustering as a variable similar to the chosen principal components. The idea was that with this approach it could be more likely to obtain clusters where the correlation would stay the same inside each cluster and differ from cluster to cluster.

5.3.1 The principal components

First, we conducted the clustering using principal components that were selected with the more standard approach, i.e., choosing the first k principal components that explain for example 90% of the total variance of the original data. However, this approach was not very successful as the number of principal components used in the clustering ended up being quite large, and only an handful of those components actually explained the correlation of the examined variables.

Therefore, we combined the two dimension reduction methods by first conducting PCA and then using random forest to the principal components similarly as in 5.2 to choose the components used in the clustering. In Table 3, all of the principal components that we used in clustering the seven different correlation pairs. The first principal component (PC1) has a negative correlation with most of the equity indexes, such as the S&P 500. It also correlates positively with interest rate indexes and the spot price of gold. Therefore, PC1 could be interpreted as a measure of how much risk the market is willing to take in the sense that with a large value for PC1 we could expect to see movement in the market to generally less volatile assets such as bonds and gold. The second principal component (PC2) correlates positively with the absolute level of inflation (CPI YOY Index) and the Federal Funds Target Rate (FDTR Index). It also correlates positively with the change in an index tracking consumer confidence (CONCCONF Index) and negatively with the change in an index tracing unemployment levels (USURTOT Index), which indicate that PC2 obtains high values in a thriving economy as in such times inflation and interest rates are often also high. The third (PC3) and fifth (PC5) principal components are quite similar. They correlate negatively with unemployment rates and positively with consumer confidence indices both in the EU and in the USA. The fourth principal component (PC4) correlates positively with indices following the prices of goods (EUPPEMUY Index) and new orders received by the industry (TMNOCHNG Index). It also correlates positively with the prices of e.g., copper and wheat. The sixth principal component (PC6) correlates positively with value equity indices and the development of the gross domestic product (GDP) of the USA, and negatively with momentum equity indices.

In Table 3, the amount of the variance of the original data explained by each component, the features that explain the components the most, and correlation between the components and those features. As expected, the components which explain the data the most were also chosen consistently by random forest as the components that explain the correlation the most. This is reasonable as the principal components that explain the variance the least often contain just noise.

The principal components chosen for each correlation pair are in Table 4. These were chosen using the criterion that each of the components must explain at least 10% of the correlation. For the correlation pair NDDUWI Index and H0A0 Index, there is only one component, PC1, which explained over 10% of the correlation. With other pairs there were two components. PC1 is chosen in five out of the seven correlation pairs, and in four of

Table 3: List of principal components used in the clustering, the variables which explain the components the most, and the correlations between them.

	Explained variance	**Feature**	**Correlation**
PC1	27%	H0A0 index spread	0.20
		XAU BGNL Curncy	0.20
		M1WOQU Index	-0.20
		M1WO000V Index	-0.20
		BFCIUS Index	-0.20
		SPX Index return	-0.19
PC2	21%	CPI YOY Index	0.25
		FDTR Index	0.19
		CONCCONF Index change	0.24
		IP CHNG Index change	0.20
		NAPMPMI Index	-0.23
		USURTOT Index change	-0.24
PC3	13%	CONCCONF Index	0.31
		EUCCEMU Index	0.31
		USURTOT Index	-0.29
		VIX index change	-0.26
PC4	8.8 %	EUPPEMUY Index	0.23
		TMNOCHNG Index	0.36
		HG1 COMB Comdty	0.30
		W1 COMB Comdty	0.28
PC5	7.9 %	CONCCONF Index	0.23
		EUCCEMU Index	0.31
		CONSSENT Index change	0.26
		W USURTOT Index	-0.30
PC6	6.9%	Equity indices Value	0.21
		Equity indices Momentum	-0.26
		M1WO000G Index	-0.21
		GDP CQOQ Index change	0.29

them it is the component that explains the variance the most. However, the first principal component does not seem to explain the correlation of EG04 Index with the other variables of interest. PC3 was chosen with three pairs, all of them having either EURUSD Index or EG04 Index as the other variable of interest. In the correlation pair of these variables PC3 explained most of the data at 67% importance score. The principal components 2, 5 and 6 were chosen only for one correlation pair each.

Table 4: The principal components chosen by random forest for each correlation pair.

Correlation pair	**PCs chosen by RF**	**Importance score**
NDDUWI Index G4O2 Index	PC1 PC2	50% 23%
NDDUWI Index EURUSD Index	PC1 PC4	62% 15%
NDDUWI Index EG04 Index	PC6 PC3	50% 19%
EURUSD Index EG04 Index	PC3 PC4	67% 13%
NDDUWI Index H0A0 Index return	PC1	85%
EURUSD Index H0A0 Index return	PC1 PC3	78% 14%
EURUSD Index G402 Index	PC5 PC1	50% 35%

5.3.2 Clustering results

The number of clusters (and thus k) was chosen mainly by visually inspecting how the clusters looked like when the correlation was plotted as a time series or when plotting the correlation with respect to one of the principal components. The different metrics defined is Section 4.4.2 were also used in the assessment, but these metrics were often not in line with the observations made in visual inspections as the metrics almost always preferred smaller k to a larger k.

The results were varying in terms of how interpretable and continuous the obtained clusters were. In Figure 7 we present the clustering results of a chosen subset of 4 correlation pairs, in which the clusters were the most reasonable and interpretable. With three of the four correlation pairs (NDDUWI Index - G4O2 Index, NDDUWI Index - EURUSD Index, EURUSD Index - G402 Index) the time periods of the clusters are almost exactly the same. As an example, we provide the distribution of PC1 values in the three clusters plotted agains the first correlation pair in Figure 8. The first cluster starts in 1990 and ends roughly by the end of 2001. This cluster is best described with a negative value of PC1, perhaps an indication of a time period where stocks performed better than less volatile assets. The time period of the first cluster is also described with a high value for PC2 indicating high absolute interest rates and inflation levels. In this time period, the stock-bond (NDDUWI Index - G4O2 Index) correlation was positive (although decreasing) together with EURUSD Index

- G402 Index correlation. The period ends roughly in the time of the dot-com bubble, after which the interest rates started to fall together with equity indices.

The second cluster spans two time periods, first from 2002 to the beginning of 2008 and then from 2014 onward. In this cluster, PC2 obtains mostly negative values indicating a time of low interest rates. The first principal components attains values near zero and the fifth mainly positive values. In this cluster the stock-bond correlation is already mostly negative although quite volatile. The correlation of EURUSD Index with both NDDUWI Index and G4O2 Index is generally positive.

The third cluster is located in the period between 2008 and 2014. This cluster is characterized by positive values for PC1 and PC2 and negative values for PC5. The positive values for PC1 are explained by the poor performance of stocks and increase in e.g., gold prices. The positive PC2 values would indicate high interest rates or inflation, however, interest rates at this time period were near zero. As we are using 5-year rolling averages in computing the principal components, the values of the components in short time periods like this do not necessarily describe very well the time period itself, but are heavily affected also by the previous years. In this case, the years leading to the financial crisis of 2007-2008 can have an strong effect. More detailed statistics of these clusters are combined in Table 5. The fourth correlation pair (EG04 Index - EURUSD Index) in Figure 7 does not have similar clusters. These clusters are more separated in time, and quite hard to interpret.

The third correlation pair (NDDUWI Index - EG04 Index) did not form clear clusters with the principal components 6 and 3 chosen by random forest. Perhaps one reason for this poor performance was that the sixth principal component only explained 6.9% of the variance of the original data. Moreover, clustering the correlation pairs containing H0A0 Index did not yield desirable end results, as the clusters were quite scattered in time.

5.4 Suggested optimal model

In this section, we compare the performance of the best models for identifying financial regimes, which we presented in Sections 5.3 and 5.2. Based on the comparison, we then suggest which models should be used for regime identification by the client and use the best models to analyze the identified regimes in more detail. Because the model construction and selection process for the model in Section 5.3 differs significantly from the process for the models in Section 5.2, the two model classes are not necessarily comparable. As such, we do not perform a comparison for the PCA model at all in this section. Conversely, the models that combine random forests with clustering are strictly comparable, thus we conduct a model comparison and use the best model for conducting the regime analysis.

We begin the model comparison by analyzing the numerical performance metrics we used for comparing clustering quality in the model selection process. The silhouette coefficient, CH index and DB index of each model for each correlation pair are presented in Table 6. As we can see from the table, k-means generally performs the best in terms of at least two of the three performance metrics for most correlation pairs. As an exception, the OPTICS algorithm performs best in terms of both silhouette coefficient and DB index for the correlation pair H0A0 index return, NDDUWI Index. When the clustering quality of k-means is the best, it is often noticeably better than the other two models, especially in terms of the CH index. On the contrary, the CH indices of DBSCAN and OPTICS are considerably

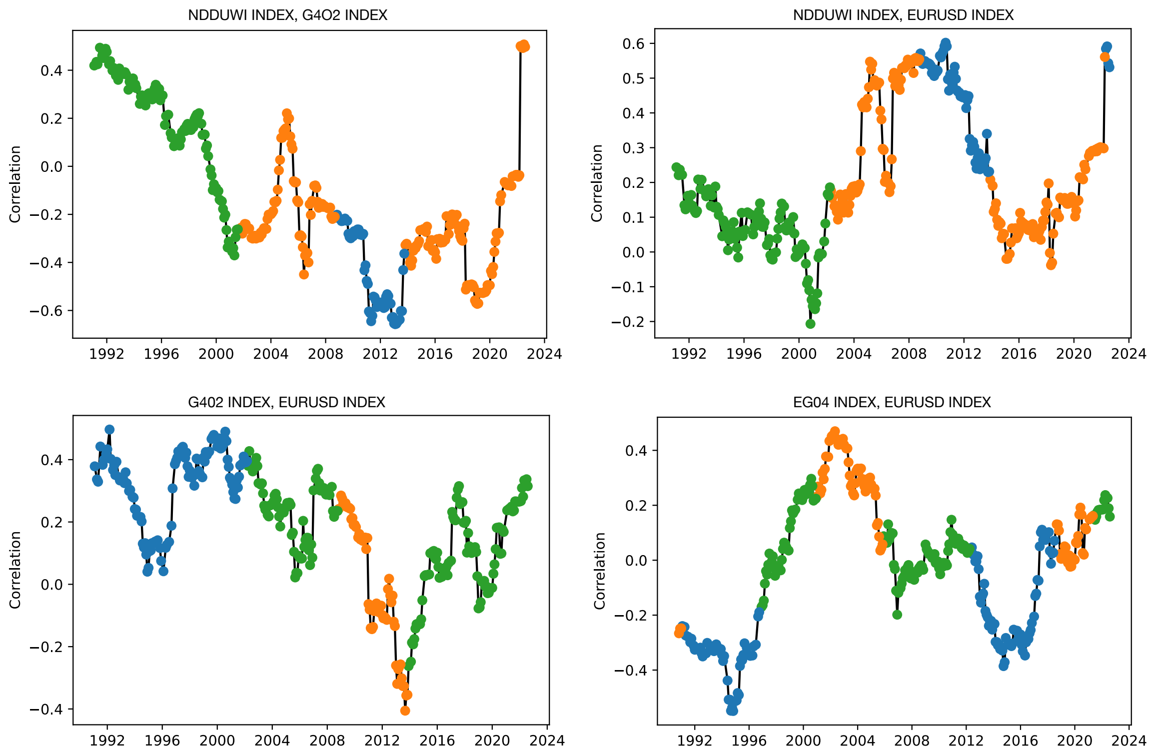

Figure 7: Examples of clustering results using PCA and k-means

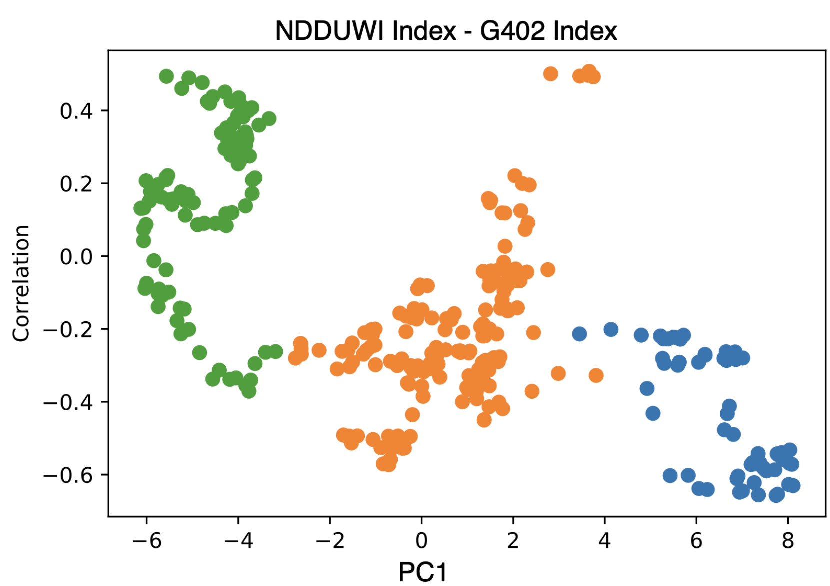

Figure 8: PC1 plotted against the correlation between NDDUWI Index and G4O2 Index. The colors represent different clusters.

Table 5: Statistics of the clusters obtained with PCA and k-means. Here, with CP1, CP2 and CP7 we refer to the correlation pairs NDDUWI Index, G4O2 Index; NDDUWI Index, EURUSD Index; and EURUSD Index, G402 Index; respectively. LTD refers to lower tail dependence coefficient.

	Cluster 1	**Cluster 2**	**Cluster 3**
Timespan	1990-2001	2002-2008 2014-	2008-2014
CP1 mean	0.15	-0.22	-0.45
CP1 std.	0.23	0.21	0.16
CP1 LTD	0.19	0.31	0.12
CP2 mean	0.06	0.22	0.45
CP2 std.	0.09	0.17	0.12
CP2 LTD	0.34	0.34	0.22
CP3 mean	0.33	0.16	-0.01
CP3 std.	0.12	0.14	0.20
CP3 LTD	0.23	0.33	0.40
PC1 mean	-4.6	0.54	6.6
PC1 std.	0.83	1.4	1.1
PC2 mean	2.4	-2.1	2.1
PC2 std.	3.4	2.6	0.73
PC4 mean	-0.17	-0.14	0.068
PC4 std.	1.8	2.7	1.3
PC5 mean	-0.35	0.54	-0.92
PC5 std.	1.7	2.2	1.9

lower than those of k-means and GMM for every correlation pair.

While numerical performance metrics generally suggest that k-means clustering is the best performing algorithm when it comes to identifying financial regimes, it is still not the clear winner for every correlation pair. Following the numerical performance comparison, we examined the visual quality of the regimes identified by each model. Firstly, we overlayed the clusters with the underlying correlation. If a clustering is visually reasonable, it should form continuous time spans, or regimes, that last multiple years and where all points within the span belong to the same cluster. Visually both k-means and GMM appear to form reasonable clusters that correspond to some type of continuous regimes. There are however some rare cases where points appear to be erroneously clustered and the regime switches for a month or two before switching back. Such points are most likely a result from the rather noisy data set that was used for fitting the models, and are not a substantial issue as they only appear for some correlation pairs. For DBSCAN and OPTICS, the number of identified regimes, is often unreasonably high, with over 20 identified regimes. Furthermore, when the regimes are overlayed with the underlying correlation, we observe that the regimes switch sporadically and often very fast, lasting only a few months at times.

In addition to analyzing the visual regime quality, we examined the overall visual cluster quality using the entire feature set. That is, we inspected how well separated the clusters

Table 6: A comparison of performance metrics for the best model for each clustering algorithm and correlation pair. The best performance metric for each correlation pair is highlighted. In the table, SC denotes the mean silhouette coefficient.

Correlation pair	**Algorithm**	**SC**	**CH index**	**DB index**
G4O2 index, NDDUWI Index	k-means	**0.36**	**127.00**	1.06
	GMM	0.22	65.30	1.63
	DBSCAN	0.27	68.31	**0.86**
G4O2 index, EURUSD index	k-means	**0.45**	**189.80**	**0.84**
	GMM	0.33	108.68	1.25
	DBSCAN	0.15	38.20	0.98
EG04 index, NDDUWI Index	k-means	**0.40**	**115.87**	0.88
	GMM	0.19	52.45	1.76
	DBSCAN	0.15	38.94	**0.84**
EG04 index, EURUSD index	k-means	**0.48**	**313.27**	**0.72**
	GMM	0.37	179.74	0.92
	DBSCAN	0.10	30.59	1.12
H0A0 index return, NDDUWI Index	k-means	0.32	**109.41**	1.24
	GMM	0.25	82.11	1.18
	OPTICS	**0.41**	66.15	**0.66**
H0A0 index return, EURUSD index	k-means	**0.41**	181.14	**0.92**
	GMM	0.37	**201.67**	1.02
	DBSCAN	0.19	44.66	0.94
NDDUWI Index, EURUSD index	k-means	**0.45**	**257.68**	0.81
	GMM	0.37	145.75	1.05
	DBSCAN	0.27	79.18	**0.78**

appeared when overlayed with the entire feature set. To visualize the high dimensional feature set, we used t-distributed stochastic neighbor embedding (t-SNE) for reducing the dimensionality to two. t-SNE is a statistical method often used for visualizing high dimensional data, that gives each datapoint a new location in a two or three-dimensional space [31]. Once again, k-means and GMM appeared to give the best clustering quality, although clusters blending together and noise points were still present. Similar to the previous visual analysis, DBSCAN and OPTICS were visually slightly worse and some clusters appeared totally random when visualized with t-SNE. An example of a t-SNE visualization for the correlation pair H0A0 index return, EURUSD index is presented in Figure 9.

For a model to be considered appropriate for identifying financial regimes, in addition to performing well numerically and visually, it must be stable when it comes to changes in the data. Ideally, the regimes the model identifies would remain the same even if some of the original data was removed. To assess the robustness of our best clustering models, we removed 20% of the original data randomly and performed the clustering again using the same hyperparameters. We then compared the regimes identified with incomplete data to the regimes identified with the full data set. A clustering model was considered stable, if

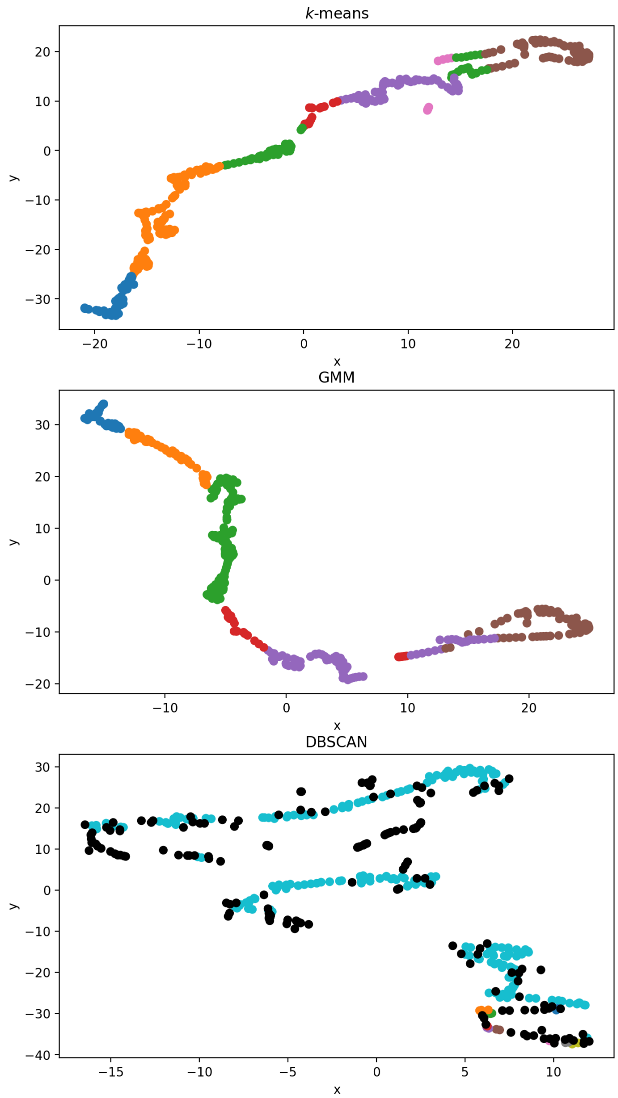

Figure 9: t-SNE visualization of the clusters for the correlation pair H0A0 index return, EURUSD index. Note that the black dots in the DBSCAN figure were classified as noise.

the regimes remained mostly unchanged with incomplete data. If the beginning or end of a regime switched by a couple of months or a year at most the model was still considered stable. The stability of clustering was only assessed visually by plotting the original clustering and the clustering achieved when 20% of the data set was removed side by side.

We observed that k-means was essentially always the most stable model. The regimes identified with incomplete data were almost exactly the same for most correlation pairs, changing only by around a year at most. As an exception, for the correlation pair EG04 index, NDDUWI Index some instability was present, but even then k-means was clearly the most robust of the tested models. On the other hand GMM produced the second most stable results. However, the stability of the GMM models was still somewhat poor in comparison to k-means. For two correlation pairs at least half of the clusters changed noticeably, while for two other correlation pairs a majority of the clusters changed at least slightly. Finally, the stability of DBSCAN and OTPICS was generally quite abysmal. When 20% of the data was removed, both the number of regimes, and their starting and ending points changed drastically.

An example of visual stability comparison for the correlation pair G4O2 index, EURUSD index is presented in Figure 10. As we can see from the figure, regimes identified with k-means remain essentially unchanged between the full and incomplete data sets, while regimes identified with GMM change to a certain degree. For example, the regime beginning from around 1995 is extended to 2008 from 2002 when we remove 20% of the original data set. Lastly, the regimes identified with DBSCAN change to something completely different.

In conclusion, the model presented in Section 5.2.1, which combines random forests with k-means clustering and more intricate feature transformations, generally performs the best in numerical performance metrics, visual regime quality and regime stability. As such, we suggest using the model in question for identifying financial regimes. Furthermore, we also suggest the model presented in Section 5.3, which performed well based on our testing, but could not be compared to the other models.

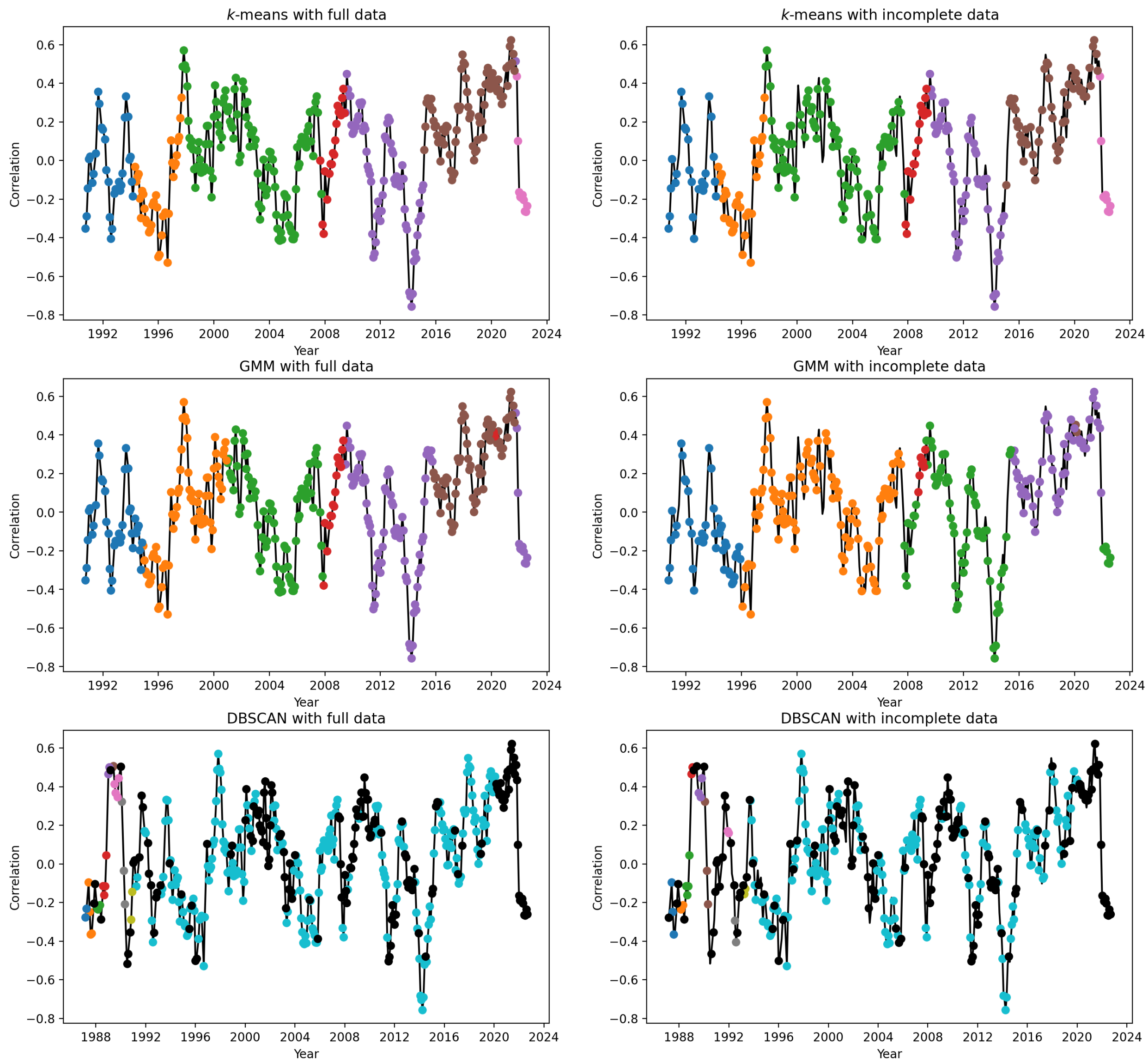

Figure 10: Stability comparison of the three clustering models for the correlation pair G4O2 index, EURUSD index when 20% of the original data set was removed at random. Note that the black dots in the DBSCAN figures were classified as noise.

5.5 Analysis of identified regimes

Using the best model presented in section 5.2.1, we note that the regimes change depending on which particular correlation pair we examine. Therefore in this section we examine three of the pairs (EG04 - NDDUWI, G4O2 - NDDUWI, and NDDUWI - EURUSD) individually and the rest of the pairs (EG04 - EURUSD, G402 - EURUSD, H0A0 - EURUSD, and H0A0 - NDDUWI) together. When examining the regimes, we try to understand the main differences using the transformed values of the different economic indicators.

5.5.1 EG04 vs NDDUWI indices

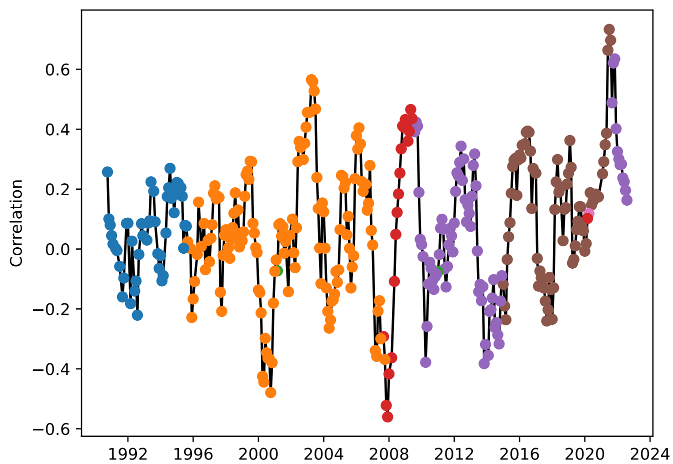

Figure 11: NDDUWI vs EG04

From Figure 11 we can see different regimes. The regimes have different sizes in terms of time and we can also see that some of the regimes appear in multiple different times. For example, the red regime corresponds to the 2008 financial crisis and is very short-lived. On the other hand, the brown regime corresponds to the relative upwards trajectory after 2015. However, the different regimes do not necessarily correspond to different correlation values, they correspond more to overall economic environments.

Examining the regimes more closely, we can begin from the blue regime. This regime lasts from the start of our dataset, 1.10.1990 until 1.9.1995. During this regime, the correlation coefficient has a mean of 0.06 with standard deviation of 0.12. The minimum and maximum values for correlation are -0.22 and 0.27 respectively. This regime has a high inflation CPI index of 2.00 on average. The average CPI index over the whole dataset is -0.08.

As inflation lessens, we get to the second regime, this regime is the longest lasting one and lasts almost continuously from 1.9.1995 to 1.11.2007. The regime also holds before and after the dot-com bubble collapse of late 90s. During this regime, the correlation coefficient obtains the minimum and maximum values of -0.48 and 0.57 respectively, with mean of 0.04 and with standard deviation of 0.22. The CPI index for this regime is -0.08 which is the mean value for the whole dataset, however it does fluctuate quite a lot within the regime, obtaining extreme values of -0.57 and 1.29.

The next regime is the red regime corresponding to the 2008 financial crisis. This regime lasts only from 1.11.2007 to 1.6.2009. Correlation during this regime increases from the minimum of -0.56 to 0.47 with mean of 0.10 and standard deviation of 0.35. The defining features of this regime are the high PPI YoY index of 0.73 on average. The EUPPEMUY index is also quite high having a mean value of 1.17.

Now, consider the next regime, colored purple on the graph. This regime is the majority regime between 1.7.2009 and 1.12.2014. The regime also re-emerges in the coronavirus pandemic lasting from 1.9.2021 until the end of the dataset at 1.8.2022. During this regime the correlation coefficient had a mean of 0.06. The extreme values are -0.38 and 0.63 with standard deviation of 0.24. The defining features of this regime are the high PPI YoY index standard deviation, having the mean value of 1.72, and a low CPI index of -0.67 on average.

The final regime is the one colored brown in the graph. This regime starts in the beginning of 2015 and lasts until 1.8.2021. During this regime, the correlation coefficient varies between -0.24 and 0.73, with mean value of 0.14 and standard deviation of 0.21. During this regime, the CPI index is quite low (-0.41 on average) and also the EUPPEMUY index is quite low, having a mean value of -1.16.

Looking at the regimes as a whole, we can conclude that the regimes are clearly different in terms of their features, most important of which are the inflation and bond yield ones. However, the correlation coefficient does not change between the regimes, at least not in a straightforward way.

5.5.2 G4O2 vs NDDUWI indices

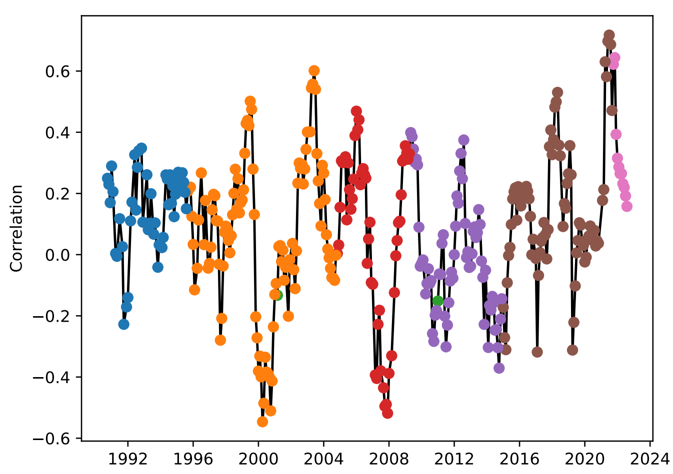

Figure 12: G4O2 and NDDUWI correlation clustering

Figure 12 shows the clustering using the best k-means model for the G4O2 and NDDUWI indices correlation. We perform a similar analysis as for the previous correlation pair.

The first cluster is the blue one which corresponds to the time before 1.8.1995. Correlation coefficient has a mean value of 0.15 with extreme values of -0.23 and 0.35. Again the most distinguished feature is the CPI inflation measure, with a mean value of 2.24.

As inflation lessens, we get the next regime. This regime is again the longest continously lasting, starting from 1.9.1995 and lasting until 1.10.2004. During this regime, the correlation coefficient ranges from -0.55 to 0.60 with mean of 0.07. This regime has a lower inflation than the first regime, at on average value of 0.16. Also during this regime, the German 10 year government bond yield (GDBR10 Index) is quite high, having average value of 1.61.

Swithing into the third regime, colored red, which is active during the 2008 financial crisis. However, it begins already at 1.11.2004 and lasts until 1.4.2009. During this regime, the correlation coefficient varies between -0.52 and 0.47 with mean of 0.07. This regime has a quite large EUPPEMUY index of 0.93 and low inflation of -0.45.

The regime after the 2008 financial crisis is active from 1.5.2009 until 1.12.2014. During this regime, the correlation varies between -0.37 and 0.40 with mean of -0.02. This regime is identified by low inflation of 0.75 and a high PPI index standard deviation of 0.88.

The second to last regime lasts from the beginning of 2015 until 1.9.2021. The correlation varies from -0.32 to 0.72 with mean of 0.15. This regime has again low inflation measure of -0.66 on average and a low EUPPEMUY index of -1.11.

The final regime starts from 1.10.2021 and lasts until the end of our dataset at 1.8.2022. The correlation has a mean value of 0.33 and varies between 0.16 and 0.64. This regime corresponds to a larger inflation than the previous one. The inflation measure has a value of 0.06. Also the PPI measure has a large value of 1.27. The inflationary pressures correspond to the end of the pandemic and economies reopening.

Holistically, the regimes do correspond to specific macro-economic conditions, however they do not provide that much valuable insight in the behaviour of the correlation coefficient.

5.5.3 NDDUWI vs EUR/USD

Next, we perform the similar analysis to the NDDUWI equity index and EUR/USD currency rate. We have again plotted the correlation and regimes in figure Figure 13. We will perform an analysis of the regimes as previously.

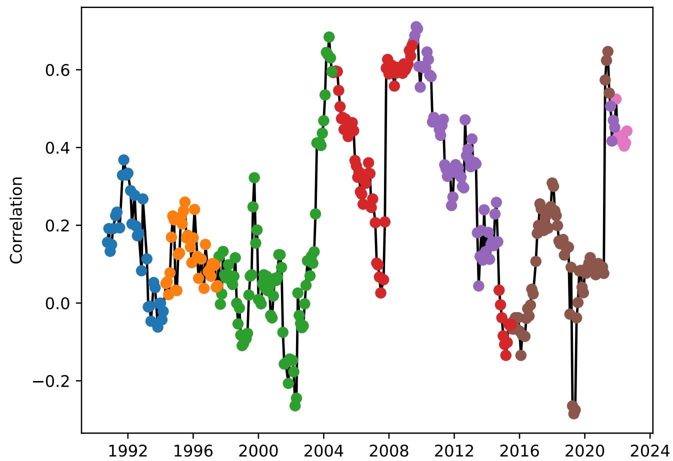

Figure 13: NDDUWI - EURUSD

We can clearly see that the regimes are somewhat different than in the previous 2 cases. Furthermore, we can see that there are more times when the regimes reappear in different points in time. We can also notice that there are regimes with differening lengths, as in previous cases.

The first regime is again from the start of our dataset until 1.4.1994. The correlation coefficient varies between -0.06 and 0.37 with mean of 0.15. This regime is again characterized by having a high CPI core value of 2.46.

The second regime (orange)is rather short as well and it only lasts from 1.4.1994 until 1.7.1997. The correlation constant stays within the same values as during the previous regime, with mean value of 0.12, however, it does not vary as much. The core inflation measure (CPI XYOY) has decreased to 1.36 on average.

The third regime start on 1.7.1997 and lasts until 1.7.2004. This regime surrounds the dot-com bubble crash at the end of the 90s. The correlation coefficient changes drastically during this regime, varying between -0.26 and 0.68, with mean of 0.09. This regime can be best identified by the even lower standardized CPI value of 0.00.

Moving onto the fourth regime, this regime corresponds to the 2008 financial crash. The regime starts at 1.7.2004 and lasts continously until 1.6.2009. However, it does reappear at the end of 2014 and lasts a couple of months into 2015. The correlation coefficient varies violently in this regime, obtaining a minimum of -0.13 and a maximum of 0.66. This is partly due to the non-continuous timing of this regime. As this regime corresponds to financial turmoils, it is characterized mainly by a low CPI index.

Next, the purple regime appears also multiple times in our history, first starting in 1.7.2009. After the first time, it lasts continuously until 1.9.2014, and reappears in the second half of 2021. The correlation coefficient has a mean value of 0.38, varying between 0.04 and 0.71. This regime is mainly identified by having a large PPI standard deviation, namely 1.76.

The next regime is the brown one, which is largely continuous and lasts from 1.7.2015 until 1.7.2021. The correlation is 0.11 on average with minimum and maximum of -0.28 and

0.65. This regime corresponds to a low PPI value of -1.24 on average and a low EUPPEMUY index of -1.27 on average.

The final regime is mainly at the very end of our dataset. From the start of 2022 to the end of our dataset at 1.8.2022. During this final regime, the correlation coefficient remains mainly the same around the value of 0.43. This regime has a higher EUPPEMUY index than the previous one (3.08 on average), indicating a higher inflation. This corresponds to the real world economic reality in 2022.

As with the previous pairs, we can conclude that the clustering did find some overall macroeconomic conditions. However, that these do not correspond to correlation structure changes, since the correlation can change quite radically within a single regime.

5.6 Other correlation pairs

In this section we briefly examine the rest of the correlation pairs that are of interest to our client: EG04 - EURUSD, G402 - EURUSD, H0A0 - EURUSD, and H0A0 - NDDUWI. The clustering results are presented in Figure 14. The results of the four correlation pairs are very similar based on visual evaluation. For all pairs the formed clusters are similar in length and in over all placement within the time span of the data set. The common theme with the correlation pairs is that there is no distinct trends present in the data. The rolling correlation is continuously moving between negative and positive, never staying in either side for extended periods of time. Since there are no visible trends in the correlation data, we are not able to detect distinct correlation themes from the regimes. The correlation behaves very similarly in all the regimes. Thus, we have clustered the data into regimes; however, the found regimes cannot be called correlation regimes, due to the lack of distinct correlation themes or trends within the clusters.

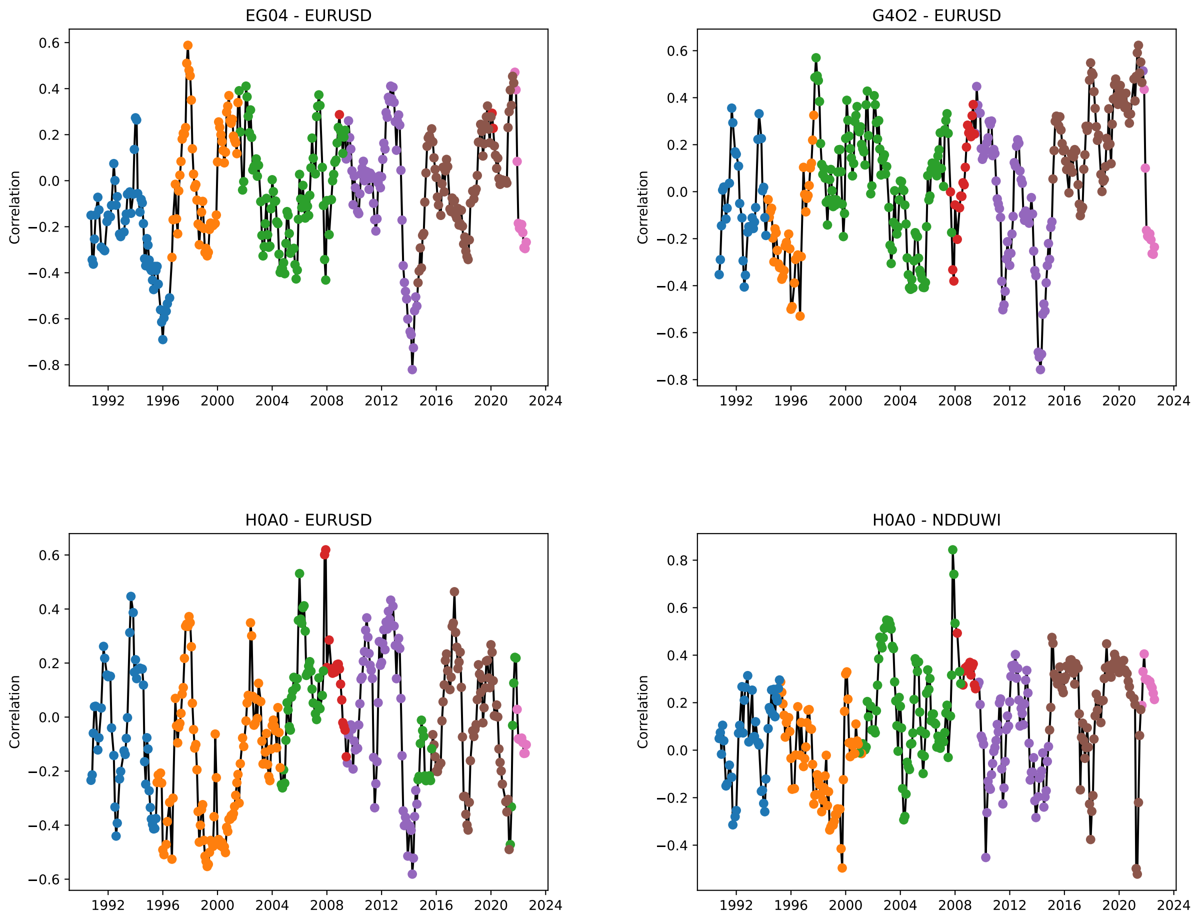

Figure 14: Rolling correlation of the remaining correlation pairs of interest separated in to clusters using the optimal model.

Re-entering previous regimes is clearly not common with the used model. With some correlation pairs (e.g., H0A0 - EURUSD) there is some re-entering, however, not in a large scale and not for longer periods of time. It seems that our model is either not able to detect reoccurring themes in the financial data or there are no reoccurring themes in the data in the first place. The regime changes often happen around some macroeconomic events with all correlation pairs. In addition, most of the clustering results suggest a short "transition regime" between longer regimes near the 2008 financial crisis. These observations support the conclusion that the clusters have some connection to the real life financial events, however, the correlation trends are not different between the clusters.

To conclude, our model is not able to detect regimes between which the correlation structure of the two financial variables would distinctively change. Moreover, there does not seem to be any visually detectable over two-year long trends present in the rolling correlations of the interesting correlation pairs.

6 Discussion

The data set used in this project was only about 30 years in length. Obtaining a good balance between the length of the data set and the number of possible features was very important and we opted for the maximum amount of data, which meant shortening of the longest time series. If we had extended the time span of the data set, we would have lost a large portion of the possible time series. In retrospect, one possibility would have been to select a few of the most promising and longest time series based on previous literature and try to find regimes based on those features and further transformations of them.

With our data set, using random forests as a dimension reduction tool rises a significant drawback. If the used data set contains a lot of similar features that, in our case, well explain the rolling correlation, only those features will be left after the dimension reduction. This reduces the dimensionality in a way that loses a lot of the more intricate information, which could have helped used in detecting the possible regimes. This problem is, however, eliminated in the PCA + random forests approach, in which similar features are "grouped together". On the other hand, the results of this approach can not ve explained intuitively. There were also a few other issues with PCA. Some of our raw data contained a few outlier values that were extremely large compared to the others. Thus, even after normalizing the data, these outlier values affected the results a lot. Furthermore, as we had many similar variables describing the similar indices, e.g., stock indices, PCA most likely weighed these indices too much compared to indices for which we had only one variable.

The proposed methods rarely re-entered previously detected regimes. This either means that there are no reoccurring themes in the financial market or that our method is not capable of detecting the similarity of the clusters based on the features used in the analysis.

The length of the detected clusters is also another point of discussion. If the aim is to find long lasting themes or trends from the financial data, is it feasible to detect clusters or regimes that are only a year in length? In this report, the number of clusters and indirectly the length of the regimes was based on the several performance metrics measuring the quality of the clusters. This approach suggested that the shorter time span of the clusters that, based on visual evaluation, do not seem to propose regimes with distinct correlation structures. However, the model is able to detect many of the macroeconomic events from the past three decades. These crises are separated into short regimes that could be interpreted as "transition regimes". It is, however, impossible to say if the current regime is a "transition regime" or a regime which is going to last for a longer period of time.

Most of the correlation pairs that our client was interested in had no clear structure or trend based on visual inspection. This was a major draw back since the rolling correlation of SPX index (stock) and G4O2 index (bond) (see Fig. 1) had two visually separable correlation regimes. If the correlation pairs of interest did not have any clear changes between positive and negative correlation time periods, the correlation structure was very similar in all of the detected regimes—constantly moving between positive and negative side.

7 Conclusion

This exploratory project set out to analyse the dependence structures between different asset classes, and to see whether it is possible to identify regimes in the data using advanced clustering techniques and a broad set of different economic variables. A further objective was to develop a robust clustering method for detecting the clusters.

We studied different clustering methods including k-means, Gaussian mixture models, DBSCAN and OPTICS, and examined their performance in clustering the data. We used two different approaches in reducing the dimensionality of out dataset and discovering the relevant variables, namely principal component analysis and random forest. The aim of the dimension reduction was to find the features that best explained the rolling correlation of a specific correlation pair. The best model was selected based on several clustering performance metrics such as silhouette coefficient and Calinski-Harasbasz index aided by visual evaluation of the results. These metrics revealed the poor performance of DBSCAN and OPTICS models and thus these models were discarded.

According to the numerical performance metrics the best model for the clustering was k-means with 7 means. Using this model we were able to divide the rolling correlation of the correlation pairs into reasonable regimes. The regime changes followed macroeconomic events such as the 2008 financial crisis, however, the correlation pairs that were of interest to our client did not show any visually separable trends and the rolling correlation was constantly moving between negative and positive. Due to the lack of visible trends lasting for several years, it is evident that a model with many clusters performs well. The obtained clusters do not have distinct correlation structures which would both differ significantly from the other clusters and behave predictably (e.g., be constantly negative) within the cluster. Thus, we were able to divide the data into regimes, however, the detected regimes are not correlation regimes.

Visual evaluation did not indicate long lasting correlation trends within the correlation pairs of interest. Moreover, neither our model nor visual inspection revealed reoccurring correlation themes from the data, since the rolling correlations behaved similarly throughout the time span of the data.

A potential further development could have been to investigate the autocorrelations or other types of dependency structures between the variables of interest. Furthermore, it is possible that using different data transformations, e.g., varying rolling mean window sizes for different economic variables could have yielded more information for the clustering process.

References

[1] Veritas Eläkevakuutus. https://veritas.fi/en/, 2023. Accessed: 5.3.2023.

[2] Peter Akioyamen, Yi Zhou Tang, and Hussien Hussien. A Hybrid Learning Approach to Detecting Regime Switches in Financial Markets. In *Proceedings of the First ACM International Conference on AI in Finance*, 2020.

[3] Andrew Ang and Allan Timmermann. Regime Changes and Financial Markets. *The Annual Review of Financial economics*, 4(1):1–27, 2012.

[4] Mihael Ankerst, Markus M Breunig, Hans-Peter Kriegel, and Jörg Sander. OPTICS: Ordering Points to Identify the Clustering Structure. *ACM Sigmod Record*, 28(2):49–60, 1999.

[5] Gary P. Brinson, L. Randolph Hood, and Gilbert L. Beebower. Determinants of Portfolio Performance. *Financial Analysts Journal*, 42(4):39–44, 1986.

[6] Alfie Brixton, Jordan Brooks, Pete Hecht, Antti Ilmanen, Thomas Maloney, and Nicholas McQuinn. A Changing Stock – Bond Correlation: Drivers and Implications. *The Journal of Portfolio Management*, pages 1–17, 2023.

[7] T. Caliński and J. Harabasz. A Dendrite Method for Cluster Analysis. *Communications in Statistics*, 3(1):1–27, 1974.

[8] Yoosoon Chang, Yongok Choi, and Joon Y Park. A New Approach to Model Regime Switching. *Journal of Econometrics*, 196(1):127–143, 2016.

[9] Jun Chen and Edward P. K. Tsang. Classification of Normal and Abnormal Regimes in Financial Markets. *Algorithms*, 11(12):1–14, 2018.

[10] David L. Davies and Donald W. Bouldin. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:224–227, 1979.

[11] Robert Engle. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economics Statistics*, 20(3):339–350, 2002.

[12] Gabriel Frahm, Markus Junker, and Rafael Schmidt. Estimating the Tail Dependence Coefficient: Properties and Pitfalls. *Insurance: Mathematics and Economics*, 37:80–100, 2006.

[13] Massimo Guidolin and Allan Timmermann. An Eeconometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns. *Journal of Applied Econometrics*, 21(1):1–22, 2006.

[14] James Hamilton. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 57(2):357–84, 1989.

[15] Robert. V. Hogg, Joseph. W. McKean, and Allen. T. Craig. *Introduction to Mathematical Statistics*. 2005.

[16] Nicholas Johnson, Vasant Naik, Niels Pedersen, and Steve Sapra. The Stock-Bond Correlation. *PIMCO Quantitative Research*, pages 1–12, 2013.

[17] Chung-Ming Kuan. Lecture on the Markov Switching Model. *Institute of Economics Academia Sinica*, 2002.

[18] Petro Liashchynskyi and Pavlo Liashchynskyi. Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS. *arXiv preprint arXiv:1912.06059*, 2019.

[19] Geoffrey J. McLachlan and David Peel. Finite Mixture Models. *John & Sons*, 2000.

[20] William M Mendenhall and Terry L Sincich. *Statistics for Engineering and the Sciences*. CRC Press, 2016.

[21] Panna Miskolczi. Note on simple and logarithmic return. *Applied Studies in Agribusiness And Commerce*, 11, 2017.

[22] Roger B Nelsen. *A Introduction to Copulas*. Springer Science & Business Media, 2006.

[23] Andrew Patton. Modelling Asymmetric Exhange Rate Dependence. *International Economic Review*, 47(2):527–556, 2006.

[24] Jeremy Piger. Econometrics: Models of Regime Changes. In *Complex systems in finance and econometrics*. Springer, 2009.

[25] Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[26] Skipper Seabold and Josef Perktold. statsmodels: Econometric and Statistical Modeling with Python. In *9th Python in Science Conference*, 2010.

[27] Abe Sklar. Fonctions de Répartition á n Dimensions et Leurs Marges. *Institut Statique de l'Université de Paris*, 8:229–231, 1959.

[28] Ralph Sueppel. Classifying market regimes. https://research.macrosynergy.com/classifying-market-regimes/, 2023. Accessed: 11.4.2023.

[29] Hendrik Supper, Felix Irresberger, and Gregor Weiß. A Comparison of Tail Dependence Estimators. *European Journal of Operational Research*, pages 728–742, 2019.

[30] Jun. Tu. Is Regime Switching in Stock Returns Important in Portfolio Decisions. *Management Science*, 56(7):1198–1215, 2010.

[31] Laurens Van der Maaten and Geoffrey Hinton. Visualizing Data Using t-SNE. *Journal of Machine Learning Research*, 9(11), 2008.

[32] Ernst Wit, Edwin van den Heuvel, and Jan-Willem Romeijn. 'All Models are Wrong...': an Introduction to Model Uncertainty. *Statistica Neerlandica*, 66(3):217–236, 2012.

[33] Boyu (Daniel) Wu, Beatrice Yeo, Kevin J. DiCurcio, and Qian Wang. The Stock/Bond Çorrelation: Increasing Amid Inflation, but Not a Regime Change. *Vanguard Research*, 2021.

A Data and feature transformations

Table 7: All time series in the data set provided by our client along with all feature transformations considered for them.

Category	Feature	Transformation
ISM & confidence	NAPMPMI Index	None, percentage change
	NAPMNMI Index	None, percentage change
	RSSAEMUM Index	None
	RSTAMOM Index	None
	EUIPEMU Index	None, percentage change
	IP CHNG Index	None, percentage change
	TMNOCHNG Index	None, percentage change
	GRIORTMM Index	None, percentage change
Employment	NFP TCH Index	None
	USURTOT Index	None, percentage change
	UMRTEMU Index	None, percentage change
Economic surpise	CESIUSD Index	None, percentage change
	CESIEUR Index	None, percentage change
	CSIIUSD Index	None, percentage change
	CSIIEUR Index	None, percentage change
Consumer confidence	CONCCONF Index	None, percentage change
	EUCCEMU Index	None, percentage change
	CONSSENT Index	None, percentage change
Financial conditions	BFCIUS Index	None
	BFCIEU Index	None
Inflation	CPI YOY Index	None, percentage change
	CPI XYOY Index	None, percentage change
	ECCPEMUY Index	None, percentage change
	CPEXEMUY Index	None, percentage change
	PPI YOY Index	None, percentage change
	EUPPEMUY Index	None, percentage change
GDP	EUGNEMUQ Index	Percentage change
	GDP CQOQ Index	None, percentage change
Central banks	FDTR Index	None, percentage change
	EUORDEPO Index	None, percentage change
	FARBAST Index	Percentage change
	EBBSTOTA Index	Percentage change
	BJACTOTL Index	Percentage change
Other economic indices	BDIY Index	Percentage change
	VIX index	None, percentage change
Gevernment bonds	G4O2 index	None
	USGG10YR Index	None
	USYC2Y10 Index	None

	USGGBE10 Index	None
	EG04 index	None
	GDBR10 Index	None
	DEYC2Y10 Index	None
	DEGGBE10 Index	None
High yield credit	H0A0 index return	None
	H0A0 index spread	None
	HE00 index return	None
	HE00 index spread	None
USA stocks	SPX index return	Log-return
	SPX index EBITDA	Log-return
	SPX index revenue	Log-return
	SPX index PE ratio	Log-return
MSCI World	NDDUWI Index	Log-return
	M1WOMVOL Index	Log-return
	M1WDHDVD Index	Log-return
	M1WOQU Index	Log-return
	M1WOMOM Index	Log-return
	M1WO000V Index	Log-return
	M1WO000G Index	Log-return
	M1WOSC Index	Log-return
Sectors	NDWUENR Index	Log-return
	NDWUMAT Index	Log-return
	NDWUIND Index	Log-return
	NDWUCDIS Index	Log-return
	NDWUCSTA Index	Log-return
	NDWUHC Index	Log-return
	NDWUFNCL Index	Log-return
	NDWUIT Index	Log-return
	NDWUTEL Index	Log-return
	NDWUUTIL Index	Log-return
	NDUWREIT Index	Log-return
Currency	EURUSD index	Log-return
	EURGBP index	Log-return
	EURJPY index	Log-return
Commodities	CO1 Comdty	Log-return
	XAU BGNL Curncy	Log-return
	HG1 COMB Comdty	Log-return
	C 1 COMB Comdty	Log-return
	W 1 COMB Comdty	Log-return
	S 1 COMB Comdty	Log-return
Private markets & other	Preqin private equity	Log-return
	Preqin Real Estate	Log-return
	HFRXAR Index	Log-return

	HFRXMD Index	Log-return
	HFRXM Index	Log-return
Equity indices	Equity indices Value	None
	Equity indices Momentum	None
	Equity indices Carry	None
	Equity indices Defensive	None

B Feature sets used for model fitting

B.1 PCA combined with k-means

Table 8: Features used with PCA & k-means. All transformations include 5 year rolling mean.

Category	**Feature**	**Transformation (excluding 5 year rolling mean)**
ISM & confidence	NAPMPMI Index	None, percentage change
	NAPMNMI Index	None, percentage change
	IP CHNG Index	None, percentage change
	TMNOCHNG Index	None, percentage change
	GRIORTMM Index	None, percentage change
Employment	NFP TCH Index	None
	USURTOT Index	None, percentage change
Consumer confidence	CONCCONF Index	None, percentage change
	EUCCEMU Index	None, percentage change
	CONSSENT Index	None, percentage change
Financial conditions	BFCIUS Index	None
Inflation	CPI YOY Index	Percentage change
	CPI XYOY Index	Percentage change
	PPI YOY Index	Percentage change
	EUPPEMUY Index	Percentage change
GDP	GDP CQOQ Index	None, percentage change
Central banks	FDTR Index	None, percentage change
Other economic indices	BDIY Index	Percentage change
	VIX index	None, percentage change
Gevernment bonds	G4O2 index	None
	USGG10YR Index	None
	USYC2Y10 Index	None
	EG04 index	None
	GDBR10 Index	None
High yield credit	H0A0 index return	None
	H0A0 index spread	None
USA stocks	SPX index return	Log-return
MSCI World	NDDUWI Index	Log-return

	M1WOMVOL Index	Log-return
	M1WOQU Index	Log-return
	M1WOMOM Index	Log-return
	M1WO000V Index	Log-return
	M1WO000G Index	Log-return
Currency	EURUSD index	Log-return
	EURGBP index	Log-return
	EURJPY index	Log-return
Commodities	CO1 Comdty	Log-return
	XAU BGNL Curncy	Log-return
	HG1 COMB Comdty	Log-return
	C 1 COMB Comdty	Log-return
	W 1 COMB Comdty	Log-return
	S 1 COMB Comdty	Log-return
Equity indices	Equity indices Value	None
	Equity indices Momentum	None
	Equity indices Carry	None
	Equity indices Defensive	None

B.2 k-means clustering with more intricate feature transformations

Table 9: Features used for fitting the k-means model.

Feature	**Transformation**
NAPMPMI Index	Percentage change
NFP TCH Index	None
USURTOT Index	Percentage change
CONCCONF Index change	Percentage change
EUCCEMU Index	Percentage change
CONSSENT Index	Percentage change
BFCIUS Index	None
FDTR Index	Percentage change
VIX index	Percentage change
Equity indices Value	None
Equity indices Momentum	None
Equity indices Carry	None
Equity indices Defensive	None
G4O2 index	None
USGG10YR Index	None
USYC2Y10 Index	None
EG04 index	None
GDBR10 Index	None
DEYC2Y10 Index	None

H0A0 index return	None
BDIY Index change	Percentage change
IP CHNG Index	Percentage change
TMNOCHNG Index	Percentage change
GRIORTMM Index	Percentage change
GDP CQOQ Index	Percentage change
EURUSD index	Log-return
EURGBP index	Log-return
EURJPY index	Log-return
CO1 Comdty	Log-return
XAU BGNL Curncy	Log-return
HG1 COMB Comdty	Log-return
C 1 COMB Comdty	Log-return
W 1 COMB Comdty	Log-return
S 1 COMB Comdty	Log-return
SPX index return	Log-return
NDDUWI Index	Log-return
M1WOMVOL Index	Log-return
M1WOQU Index	Log-return
M1WOMOM Index	Log-return
M1WO000V Index	Log-return
M1WO000G Index	Log-return
CPI YOY Index	5yr rolling mean
CPI XYOY Index	5yr rolling mean
PPI YOY Index	5yr rolling mean
EUPPEMUY Index	5yr rolling mean
CPI YOY Index std	5yr rolling std of yearly change
CPI XYOY Index std	5yr rolling std of yearly change
PPI YOY Index std	5yr rolling std of yearly change
EUPPEMUY Index std	5yr rolling std of yearly change

B.3 Gaussian mixture models

Table 10: Features used for fitting the best GMM model.

Feature	**Transformation**
NAPMPMI Index	Percentage change
NFP TCH Index	None
USURTOT Index	Percentage change
CONCCONF Index change	Percentage change
EUCCEMU Index	Percentage change
CONSSENT Index	Percentage change
BFCIUS Index	None
FDTR Index	Percentage change

VIX index	Percentage change
Equity indices Value	None
Equity indices Momentum	None
Equity indices Carry	None
Equity indices Defensive	None
G4O2 index	None
USGG10YR Index	None
USYC2Y10 Index	None
EG04 index	None
GDBR10 Index	None
DEYC2Y10 Index	None
H0A0 index return	None
BDIY Index change	Percentage change
IP CHNG Index	Percentage change
TMNOCHNG Index	Percentage change
GRIORTMM Index	Percentage change
GDP CQOQ Index	Percentage change
EURUSD index	Log-return
EURGBP index	Log-return
EURJPY index	Log-return
CO1 Comdty	Log-return
XAU BGNL Curncy	Log-return
HG1 COMB Comdty	Log-return
C 1 COMB Comdty	Log-return
W 1 COMB Comdty	Log-return
S 1 COMB Comdty	Log-return
SPX index return	Log-return
NDDUWI Index	Log-return
M1WOMVOL Index	Log-return
M1WOQU Index	Log-return
M1WOMOM Index	Log-return
M1WO000V Index	Log-return
M1WO000G Index	Log-return
CPI YOY Index	60 month rolling mean
CPI XYOY Index	60 month rolling mean
PPI YOY Index	60 month rolling mean
EUPPEMUY Index	60 month rolling mean

B.4 DBSCAN and OPTICS

The same as GMM, table 10.

C Best features identified by random forest

C.1 k-means clustering with more intricate feature transformations

Table 11: Most important features for the correlation pair G4O2 index, EURUSD index.

Feature	**Importance score**
CPI XYOY Index std	0.408131
CPI XYOY Index	0.081414
EUPPEMUY Index std	0.071610
GDBR10 Index	0.065605
EG04 index	0.051604
EUPPEMUY Index	0.046920
H0A0 index return	0.046106
PPI YOY Index	0.031853
DEYC2Y10 Index	0.023735
CPI YOY Index std	0.021704
USGG10YR Index	0.021280
USYC2Y10 Index	0.011106
BFCIUS Index	0.010557

Table 12: Most important features for the correlation pair G4O2 index, NDDUWI index.

Feature	**Importance score**
H0A0 index return	0.214081
CPI XYOY Index	0.154618
USYC2Y10 Index	0.098224
GDBR10 Index	0.082596
DEYC2Y10 Index	0.052076
EUPPEMUY Index	0.047558
PPI YOY Index	0.040350
PPI YOY Index std	0.038344
CPI XYOY Index std	0.036527
EUPPEMUY Index std	0.032020
CPI YOY Index std	0.031738
USGG10YR Index	0.017683
CPI YOY Index	0.016453

Table 13: Most important features for the correlation pair EG04 index, NDDUWI index.

Feature	**Importance score**
CPI XYOY Index	0.221469
USGG10YR Index	0.145865

PPI YOY Index std	0.080774
USYC2Y10 Index	0.075164
H0A0 index return	0.059842
EUPPEMUY Index	0.057421
DEYC2Y10 Index	0.047240
BFCIUS Index	0.026645
CPI XYOY Index std	0.025295
GDBR10 Index	0.025129
GDP CQOQ Index change	0.023403
NFP TCH Index	0.021937
PPI YOY Index	0.021604
CPI YOY Index	0.017128
EUPPEMUY Index std	0.015465
G4O2 index	0.011598
CPI YOY Index std	0.011000
FDTR Index change	0.008942

Table 14: Most important features for the correlation pair EG04 index, EURUSD index.

Feature	**Importance score**
GDBR10 Index	0.369993
CPI XYOY Index	0.264815
FDTR Index change	0.082451
USGG10YR Index	0.054080
G4O2 index	0.024022
PPI YOY Index std	0.023253
USYC2Y10 Index	0.022204
CPI XYOY Index std	0.018383
CPI YOY Index std	0.014450
EUPPEMUY Index std	0.013045
H0A0 index return	0.010694

Table 15: Most important features for the correlation pair H0A0 index return, EURUSD index.

Feature	**Importance score**
EUPPEMUY Index	0.291638
PPI YOY Index std	0.140279
CPI XYOY Index	0.110428
CPI YOY Index std	0.102809
PPI YOY Index	0.096644
USYC2Y10 Index	0.041026
EUPPEMUY Index std	0.026582
CPI XYOY Index std	0.026450

CPI YOY Index	0.017924
DEYC2Y10 Index	0.015798
BFCIUS Index	0.014518
G4O2 index	0.012798

Table 16: Most important features for the correlation pair NDDUWI index, EURUSD index.

Feature	**Importance score**
PPI YOY Index	0.260338
EUPPEMUY Index	0.176080
PPI YOY Index std	0.103182
CPI XYOY Index std	0.102252
USGG10YR Index	0.065401
EUPPEMUY Index std	0.052375
DEYC2Y10 Index	0.048407
CPI XYOY Index	0.033174
H0A0 index return	0.031436

C.2 Gaussian mixture models

Table 17: Most important features for the correlation pair G4O2 index, EURUSD index.

Feature	**Importance score**
H0A0 index return	0.237639
CPI XYOY Index	0.178608
USYC2Y10 Index	0.099725
GDBR10 Index	0.092059
EUPPEMUY Index	0.055880
PPI YOY Index	0.053039
DEYC2Y10 Index	0.044535
USGG10YR Index	0.026005
CPI YOY Index	0.023097
EG04 index	0.020865
FDTR Index change	0.010117
TMNOCHNG Index change	0.010023
BDIY Index change	0.009922
GDP CQOQ Index change	0.009490
CO1 Comdty	0.007619
BFCIUS Index	0.007141
C 1 COMB Comdty	0.006899
W 1 COMB Comdty	0.006844

Table 18: Most important features for the correlation pair EG04 index, NDDUWI Index.

Feature	**Importance score**
CPI XYOY Index	0.263072
USGG10YR Index	0.147565
USYC2Y10 Index	0.095675
EUPPEMUY Index	0.063316
H0A0 index return	0.060351
DEYC2Y10 Index	0.050051
BFCIUS Index	0.034601
GDBR10 Index	0.032396
GDP CQOQ Index change	0.026374
CPI YOY Index	0.026307
NFP TCH Index	0.024585
PPI YOY Index	0.023428
G4O2 index	0.012736
FDTR Index change	0.008803
EUCCEMU Index change	0.008331
Equity indices Momentum	0.007865
XAU BGNL Curncy	0.007780
NAPMPMI Index change	0.006385

Table 19: Most important features for the correlation pair EG04 index, NDDUWI Index.

Feature	**Importance score**
CPI XYOY Index	0.242506
USGG10YR Index	0.154054
USYC2Y10 Index	0.096922
H0A0 index return	0.064416
EUPPEMUY Index	0.063192
DEYC2Y10 Index	0.047490
BFCIUS Index	0.033865
GDBR10 Index	0.032630
GDP CQOQ Index change	0.028865
NFP TCH Index	0.027851
CPI YOY Index	0.027545
PPI YOY Index	0.023854
G4O2 index	0.012675
FDTR Index change	0.008982
XAU BGNL Curncy	0.007924
Equity indices Momentum	0.007909
EUCCEMU Index change	0.007738
NAPMPMI Index change	0.006504

Table 20: Most important features for the correlation pair EG04 index, EURUSD index.

Feature	**Importance score**
GDBR10 Index	0.362737
CPI XYOY Index	0.273178
FDTR Index change	0.081523
USGG10YR Index	0.075885
G4O2 index	0.033243
USYC2Y10 Index	0.022438
EUPPEMUY Index	0.014608
H0A0 index return	0.014470
BFCIUS Index	0.011180

Table 21: Most important features for the correlation pair H0A0 index return, NDDUWI Index.

Feature	**Importance score**
BFCIUS Index	0.171909
G4O2 index	0.158575
PPI YOY Index	0.107991
USGG10YR Index	0.103483
CPI XYOY Index	0.067218
EG04 index	0.066249
DEYC2Y10 Index	0.057598
USYC2Y10 Index	0.038673
GDBR10 Index	0.032796
EUPPEMUY Index	0.028452
CPI YOY Index	0.020393
EUCCEMU Index change	0.012693
GDP CQOQ Index change	0.011055
NFP TCH Index	0.010628
EURGBP index	0.007626

Table 22: Most important features for the correlation pair H0A0 index return, EURUSD index.

Feature	**Importance score**
EUPPEMUY Index	0.426411
PPI YOY Index	0.178444
CPI XYOY Index	0.074011
USYC2Y10 Index	0.047798
BFCIUS Index	0.034420
G4O2 index	0.031548
GDBR10 Index	0.029012

CPI YOY Index	0.028843
USGG10YR Index	0.021841
DEYC2Y10 Index	0.013478
EG04 index	0.011308

Table 23: Most important features for the correlation pair NDDUWI Index, EURUSD index.

Feature	**Importance score**
PPI YOY Index	0.274501
EUPPEMUY Index	0.170978
USGG10YR Index	0.144303
CPI XYOY Index	0.072142
DEYC2Y10 Index	0.067487
H0A0 index return	0.053648
BFCIUS Index	0.051815
USYC2Y10 Index	0.040368

C.3 DBSCAN and OPTICS

Table 24: Most important features for the correlation pair NDDUWI Index, EG04 index.

Feature	**Importance score**
CPI XYOY Index	0.194225
USGG10YR Index	0.175906
USYC2Y10 Index	0.072104
IP CHNG Index	0.064016
EUPPEMUY Index	0.050277
CPI YOY Index	0.044179
H0A0 index return	0.043495
NAPMPMI Index	0.035556
FDTR Index	0.030921
H0A0 index spread	0.030621
EUCCEMU Index	0.028082
PPI YOY Index	0.024366
GDP CQOQ Index	0.019944
TMNOCHNG Index	0.017566
FDTR Index change	0.017282
USURTOT Index	0.016744
G4O2 index	0.016063
GDP CQOQ Index change	0.015031

Table 25: Most important features for the correlation pair NDDUWI Index, EURUSD index.

Feature	Importance score
EUPPEMUY Index	0.555408
NAPMPMI Index	0.114058
FDTR Index	0.07998
TMNOCHNG Index	0.039945
USYC2Y10 Index	0.0317
CPI XYOY Index	0.030538
USGG10YR Index	0.020265
H0A0 index return	0.019256

Table 26: Most important features for the correlation pair G4O2 Index, EURUSD index.

Feature	Importance score
CONCCONF Index	0.287641
IP CHNG Index	0.201804
CPI XYOY Index	0.087657
PPI YOY Index	0.052443
USURTOT Index	0.04711
EUCCEMU Index	0.038005
TMNOCHNG Index	0.029049
FDTR Index	0.026283
USGG10YR Index	0.021123
USYC2Y10 Index	0.019795
EUPPEMUY Index	0.019188
GDP CQOQ Index	0.016478
EG04 index	0.016456
H0A0 index return	0.015933
CONSSENT Index	0.010184
NAPMPMI Index	0.009526

Table 27: Most important features for the correlation pair H0A0 index, NDDUWI index.

Feature	Importance score
USGG10YR Index	0.165317
NAPMPMI Index	0.105495
PPI YOY Index	0.083296
EUPPEMUY Index	0.079951
GDP CQOQ Index	0.073585
G4O2 index	0.07171
IP CHNG Index	0.045116
TMNOCHNG Index	0.043558
H0A0 index spread	0.042821
USYC2Y10 Index	0.033551
CPI XYOY Index	0.030578

USURTOT Index	0.02622
EUCCEMU Index	0.022007
EG04 index	0.018978
FDTR Index	0.017242
CPI YOY Index	0.013025
GDP CQOQ Index change	0.009087
C 1 COMB Comdty	0.008895
NFP TCH Index	0.007444

Table 28: Most important features for the correlation pair EG04 index, EURUSD index.

Feature	**Importance score**
CPI XYOY Index	0.192911
USGG10YR Index	0.145946
H0A0 index spread	0.10805
CONCCONF Index	0.077952
PPI YOY Index	0.064264
TMNOCHNG Index	0.051626
EUPPEMUY Index	0.04832
IP CHNG Index	0.043023
CPI YOY Index	0.037961
USYC2Y10 Index	0.033999
EUCCEMU Index	0.028932
FDTR Index change	0.019292
FDTR Index	0.011455
GDP CQOQ Index	0.011203
CONSSENT Index	0.010637
USURTOT Index	0.010482

Table 29: Most important features for the correlation pair H0A0 index, EURUSD index.

Feature	**Importance score**
PPI YOY Index	0.195011
EUPPEMUY Index	0.193489
CPI XYOY Index	0.14064
EUCCEMU Index	0.096153
GDP CQOQ Index	0.043237
IP CHNG Index	0.039414
CPI YOY Index	0.034175
USYC2Y10 Index	0.021454
FDTR Index	0.018491
USURTOT Index	0.017875
PPI YOY Index change	0.016072
EG04 index	0.014831

G4O2 index	0.013704
USGG10YR Index	0.011996
NAPMPMI Index	0.010229
H0A0 index spread	0.010117
GDP CQOQ Index change	0.009299
CONCCONF Index	0.007616

Table 30: Most important features for the correlation pair G4O2 index, NDDUWI index.

Feature	**Importance score**
CPI XYOY Index	0.260466
H0A0 index return	0.11505
EUPPEMUY Index	0.088352
USYC2Y10 Index	0.088081
CPI YOY Index	0.047175
GDP CQOQ Index	0.04148
IP CHNG Index	0.033869
FDTR Index	0.031069
EUCCEMU Index	0.028422
PPI YOY Index	0.028201
CONCCONF Index	0.026964
TMNOCHNG Index	0.020732
H0A0 index spread	0.018684
USGG10YR Index	0.017702
EG04 index	0.016822
USURTOT Index	0.012871
FDTR Index change	0.010939
CONSSENT Index	0.01004

D Hyperparameters

Tables for the different hyperparameters used in 4.4.2.

Table 31: Set of hyperparameters used in grid search for Gaussian mixture clustering.

Parameter	Description	Set of values
k	Number of clusters	$\{2, 3, 4, 5, 6\}$
c_{type}	Covariance type	{full, tied, diagonal, spherical}

Table 32: Set of hyperparameters used in grid search for DBSCAN.

Parameter	Description	Set of values
ϵ	Circle radius	$\{0.2, 0.4 \ldots 5.8\}$
minimum samples	Minimum number of data points	$\{2, 4, 8 \ldots 16\}$.

Table 33: Set of hyperparameters used in grid search for OPTICS.

Parameter	Description	Set of values
ξ	Parameter for xi-clustering	$\{0.05, 0.15 \ldots 1.0\}$
p	Minkowski metric exponent	$\{1.0, 1.5 \ldots 3.0\}$
minimum samples	Minimum number of data points	$\{2, 4, 8 \ldots 16\}$
Metric	Measure of distance used	{Minkowski, cosine}
Cluster method	Method for clustering the points	{Xi, DBSCAN}

E Self assessment

E.1 Following the initial project plan

The scope of our work has remained exploratory throughout the project, with implementation loosely following the initial project plan. When first discussing the research questions within our team and with the client, we knew that a more exploratory "trial and error" scope would be necessary. As we had three wide questions and only a limited amount of previous research discussing our topic, narrowing the exact topic was revealed as a challenge. Overall, we think that we stuck with our initial project timeline and plan relatively well. Small adjustments were done to the phasing of the work, as shown in Figure 15.

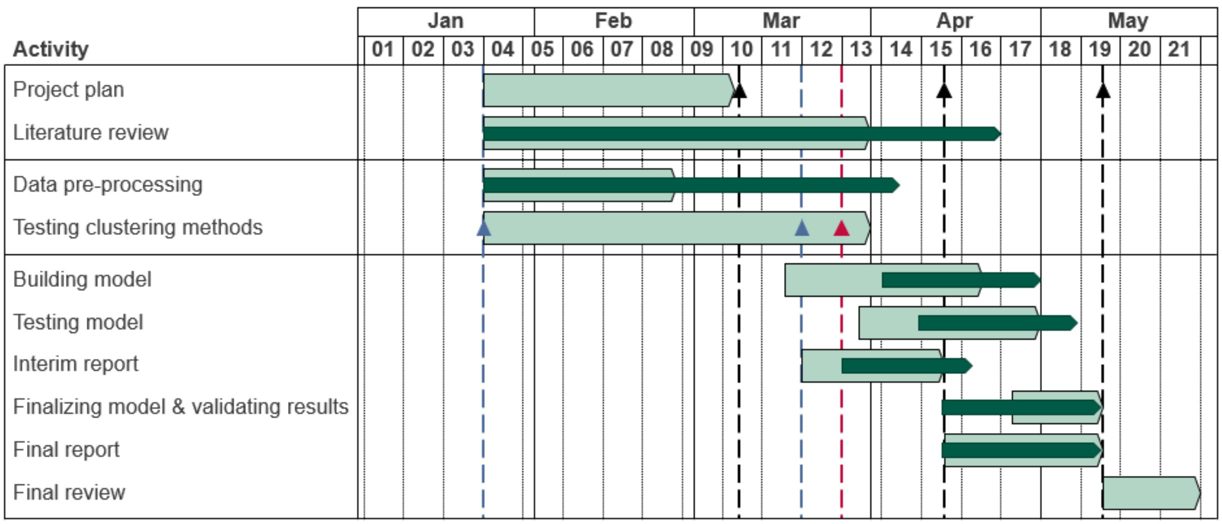

Figure 15: The initial project timeline with modifications to better picture the actual timeline of our work. The light green arrows represent the initial timeline and the darker arrows represent how the team worked. The red line represents an additional workshop held with Veritas in March.

No major departures were done to the original project plan, and to summarize, the initial project plan was followed well. None of the risks identified in the first phase of the work were realized. We have been in contact with the client regularly throughout the project and they have helped guide us with any questions regarding the topic that the team has had.

The workload of this project has remained reasonable. However, distributing the workload evenly within the team could have been more successful, which is explained more in detail in Section E.3.

E.2 Succession of the Project

Overall the project was a success. Although the final results of the project might not have been as satisfactory as expected, we were able to carry out the data analysis with a proper methodology and high standards. During the project we learned a lot about working as a team and managing a project, writing scientific text, presenting, and also about the topic of out project, which was quite new to all of us.

In the beginning of the project, we did not focus enough on the obtained data. The first priority should have been in analyzing the data properly, in the sense that the variables would have been more familiar and could have made the right data transformations right away. We focused maybe a bit too much on technical details such as choosing clustering methods and researching about tail-dependencies. A lot of the hours spend on the project were spend on managing the data, and too little time was left for the actual analysis of the results.

E.3 Improvement

E.3.1 Project team

As mentioned before, the distribution of the workload within the project team could have been better. It was clear that part of the team was contributing more to the work than others. However, it seems that everyone was working on tasks that they felt suited the best to their personal capabilities. This being said, not everyone was allowed to work on all tasks, which is most likely good in terms of finalizing the project, but not so good in terms of everyone getting a good learning experience on the topic.

In addition, the communication within the team could have been better and more assertive. We were, up to the very end, disagreeing on some fundamental aspects of the project such as how the clustering should be conducted and what the regimes actually mean, and never quite got a common view of the objectives as a whole. We could have also worked better as a team as at times, the work felt more individual than group work.

On a more positive note, the project team has supported each other well. Everyone has been open to helping others if necessary. Writing the final report also felt like a joint project, where all members contributed well in both writing and reviewing the report.

E.3.2 Client team

Our team and the client team could have been in contact more often. We met only a handful of times during the work, which affected negatively the professional learning experience of the project. However, as our team members were working in addition to school, having fewer meetings with the client helped with the workload of this course and was necessary from the point of view of the energy levels of the team members.

E.3.3 Teacher(s)

In our opinion, the contact with the teaching staff has been satisfactory. We have received help from the professor if necessary, and the teaching assistant has been very reactive and responded well and fast to any of our questions.