RESEARCH ARTICLE

2023, vol. 10, issue 1, 269-283
https://doi.org/10.5281/zenodo.8241412

DYNAMIC INTERDEPENDENCE BETWEEN ASSET CLASSES: A SPECTRAL CO-CLUSTERING AND VAR ANALYSIS

Andrei-Dragos Popescu [1]
Cristi Spulbar [2]

[1] PhD Student, Doctoral School of Economic Sciences, University of Craiova, Craiova, Romania
andrei@scx.holdings
[2] Professor, Doctoral School of Economic Sciences, University of Craiova, Craiova, Romania
cristi.spulbar@feea.ucv.ro

Abstract

This article proposes a new approach for identifying groups of assets that exhibit similar behavior under various market conditions using Spectral Co-Clustering with VAR modeling. Our approach uses VAR models to capture the dynamic interdependence between different asset classes and applies Spectral Co-Clustering to identify groups of assets that exhibit similar patterns of behavior. The method is evaluated on a dataset of asset prices, and its performance is compared to existing methods using various metrics. Results show that our proposed method outperforms other existing methods. The proposed approach can help investors identify groups of asset classes that behave similarly under different market conditions.

Keywords: *Spectral Co-Clustering, Financial Digital Assets, Crypto Assets, Financial Markets, European Markets.*

Introduction

The emergence of digital assets, such as crypto assets, has created new opportunities and challenges for investors, regulators, and researchers. Digital assets are characterized by high volatility, rapid innovation, and a lack of established valuation models, which make them difficult to analyze using traditional financial methods.

The increasing availability of financial data and advances in computational methods have led to the development of new approaches for analyzing complex financial systems. One such approach is the Spectral Co-Clustering methodology combined with a Vector Autoregressive (VAR) approach, which has been used to investigate the dynamic correlations among different financial markets, asset classes, and sectors. In recent years, however, researchers have begun to apply novel approaches, such as the Spectral Co-Clustering methodology to gain insights into the behavior and dynamics of digital assets and their interactions with traditional financial markets.

The Spectral Co-Clustering methodology is a powerful tool for identifying groups of assets that share similar characteristics or exhibit similar behaviors over time. This methodology involves applying a clustering algorithm to a matrix of pairwise correlations between assets, in order to identify groups of assets that have high intra-group correlations and low inter-group correlations. By using the Spectral Co-Clustering methodology, researchers can gain insights into the structure and dynamics of financial systems, and identify important interdependencies and spillover effects between different markets and sectors.

When combined with a VAR approach, the Spectral Co-Clustering methodology can provide a powerful framework for modeling the dynamic relationships between different financial variables. The VAR model is a popular econometric tool for analyzing the interdependencies between multiple time series, and has been widely used in finance and economics to study topics such as asset pricing, risk management, and monetary policy. By using a VAR model in conjunction with the Spectral Co-Clustering methodology, researchers can capture the complex interdependencies and feedback loops that exist between different financial variables, and generate useful insights for investors, policymakers, and other stakeholders.

In this study, we review the literature on the application of the Spectral Co-Clustering methodology combined with a VAR approach in financial markets, combined with financial digital assets. We begin by discussing the theoretical underpinnings of the Spectral Co-Clustering methodology and the VAR model, and then review a selection of studies that have applied this approach in various financial contexts. By examining these studies, we aim to provide a comprehensive overview of the ways in which the Spectral Co-Clustering methodology combined with a VAR approach can be used to analyze financial systems, and identify the key findings and contributions of this research. Through our review, we aim to provide insights into the key findings and contributions of this research, as well as identify the challenges and opportunities for future research in this field.

Bi-clustering is a data analysis technique that involves clustering both rows and columns of a data matrix simultaneously. It has found widespread applications in many fields, including bioinformatics, text mining, and image analysis. In finance, bi clustering can be used to identify groups of stocks that exhibit similar behavior under different market conditions. Vector Autoregression (VAR) models, on the other hand, are widely used in econometrics to model time-series data.

Spectral Co-Clustering is a powerful bi-clustering method that combines spectral graph theory and co-clustering. It can be used to identify clusters of rows and columns that exhibit similar patterns of behavior in a given data matrix. Spectral Co-Clustering has been applied in various fields, including text mining, bioinformatics, and image analysis. In finance, it can be used to identify groups of assets that exhibit similar behavior under different market conditions.

Overall, this research presents a novel approach for identifying groups of assets that exhibit similar behavior under different market conditions.

Literature review

One notable application of this methodology is in the analysis by Sorensen K. (2014), were the focus was on the market graph, constructed from time series of price return on the American stock market. Two different methods originating from clustering analysis in social networks and image segmentation are applied to obtain graph partitions and the results are evaluated in terms of the structure and quality of the partition. Along with the market graph, power law graphs from three different theoretical graph models are considered. This study highlights topological features common in many power law graphs as well as their differences and limitations. The results showed that the market graph possess a clear clustered structure only for higher correlation thresholds. By studying the internal structure of the graph clusters, she found that they could serve as an alternative to traditional sector classification of the market.

Nagy L. and Ormos M. (2018) introduced a spectral clustering-based method to show that stock prices contain not only firm but also network-level information. They clustered different stock indices and reconstructed the equity index graph from historical daily closing prices. Their research has shown that tail events have a minor effect on the equity index structure and the covariance and Shannon entropy do not provide enough information about the network. However, Gaussian clusters can explain a substantial part of the total variance.

The research performed by Saliner JMG (2019), spectral clustering techniques are used to find clusters in the stock market with the aim of comparing the structure of the clusters with the industry classification of companies. They have used these representation techniques to visualize the increase of systemic risk associated with an economic crisis.

In the study by Yosra B. S. et al. (2022), the authors proposed a model-based co-clustering algorithm for mixed data, functional and binary. Co-clustering aims to identify block patterns in a dataset from a simultaneous clustering of rows and columns. The proposed approach relies on the latent block model, and three algorithms are compared for its inference: stochastic EM within Gibbs sampling, classification EM and variational EM. The proposed model is the first co-clustering algorithm for mixed data that deals with functional and binary features. The model has proven its efficiency on simulated data and on real data extracted from live 4G mobile networks.

In the era of big data, there are increasing interests on clustering variables for the minimization of data redundancy and the maximization of variable relevancy. Existing clustering methods, however, depend on nontrivial assumptions about the data structure. In the research performed by Chen Y. and Yang H (2016), they reformulate the problem of variable clustering from an information theoretic perspective that does not require the assumption of data structure for the identification of nonlinear interdependence among variables. Nonlinear interdependence among variables poses significant challenges on the traditional framework of predictive modeling. The researchers propose and apply the use of mutual information to characterize and measure nonlinear correlation structures among variables, by developing Dirichlet process (DP) models to cluster variables based on the mutual-information measures among variables. They applied orthonormalized variables in each cluster which are integrated with group elastic-net model to improve the performance of predictive modeling.

Both simulation and real-world case studies showed that the proposed methodology not only effectively reveals the nonlinear interdependence structures among variables but also outperforms traditional variable clustering algorithms such as hierarchical clustering.

In another study by Guangwei S. et al. (2018), two co-clustering methods based on smooth plaid model (SPM) and parallel factor decomposition with sparse latent factors (SLF-PARAFAC) are respectively applied to synthetic data set and investors' transaction-level data set from the China Financial Futures Exchange. In their research the comparison between the two methodologies is clearly observed.

Both SLF-PARAFAC and SPM are efficient, robust, and well suited for discovering trading ecosystems in modern financial markets. The analysis recognized temporal pattern differences of various trader types and the results helped to develop a thorough understanding of trading behaviors, and to detect patterns and irregularities.

The analysis performed by Bennett S. et al. (2022) focused on multivariate time series systems, it has been observed that certain groups of variables partially lead the evolution of the system, while other variables follow this evolution with a time delay; the result is a lead–lag structure amongst the time series variables. The researchers demonstrated that the web of pairwise lead–lag relationships between time series can be helpfully construed as a directed network, for which there exist suitable algorithms for the detection of pairs of lead–lag clusters with high pairwise imbalance. Within the framework, the authors considered a number of choices for the pairwise lead–lag metric and directed network clustering model components. The framework is validated on both a synthetic generative model for multivariate lead–lag time series systems and daily real-world US equity prices data. In their opinion the method is able to detect statistically significant lead–lag clusters in the US equity market. We agree that the study shows the nature of these clusters in the context of the empirical finance literature on lead–lag relations, and demonstrate how these can be used for the construction of predictive financial signals.

While spectral techniques have been successfully applied for clustering undirected graphs, the performance of spectral clustering algorithms for directed graphs (digraphs) is not in general satisfactory: these algorithms usually require symmetrising the matrix representing a digraph, and typical objective functions for undirected graph clustering do not capture cluster-structures in which the information given by the direction of the edges is crucial. To overcome this, Cucuringu M. et al. (2020), proposed a spectral clustering algorithm based on a complex-valued matrix representation of digraphs. The researchers analyzed the theoretical performance on a Stochastic Block Model for digraphs in which the cluster-structure is given not only by variations in edge densities, but also by the direction of the edges. The significance of the study is highlighted on a data set pertaining to internal migration in the United States: while previous spectral clustering algorithms for digraphs can only reveal that people are more likely to move between counties that are geographically close, this approach is able to cluster together counties with a similar socio-economical profile even when they are geographically distant, and illustrates how people tend to move from rural to more urbanized areas.

The Spectral Co-Clustering methodology combined with a VAR approach has been applied in various financial markets, including digital assets and traditional finance markets. This approach allows for the clustering of both assets and time periods, which can provide valuable insights into the relationships and dynamics between financial assets.

Overall, the Spectral Co-Clustering methodology combined with a VAR approach has proven to be a valuable tool for analyzing the relationships and dynamics between financial assets in both digital and traditional finance markets. By clustering both assets and time periods, this approach can provide valuable insights into the underlying factors driving the movements of financial digital assets and markets.

These studies demonstrate the broad applicability of this methodology, by revealing underlying clusters of assets and modeling their interrelationships over time, this approach can provide valuable insights for investors, policymakers, and researchers alike.

Further, this can provide evidence of the applicability and versatility of the Spectral Co-Clustering methodology combined with a VAR approach in financial markets, and demonstrate the overview that can be gained by examining the dynamic relationships between different asset classes and sectors.

Methodology

In this article, we propose to use Spectral Co-Clustering with VAR models to identify groups of assets that exhibit similar behavior under different market conditions. Specifically, we will use VAR models to capture the dynamic interdependence between different markets, and then apply Spectral Co-Clustering to identify groups of assets that exhibit similar patterns of behavior. We will apply our proposed method to a dataset of asset prices and evaluate its performance using various metrics.

The datasets used for this research are from January 2016 till October 2022 for the following financial digital assets, markets and indexes: Bitcoin (BTC), Ethereum (ETH), Dow Jones Industrial Average (DJIA), Standard and

Poor's 500 Index (SP500), Nasdaq Composite (IXIC), Euronext 100 (N100), Euro Stoxx 50 (STOXX50), Nikkei Index (N225), Shanghai Stock Exchange Composite Index (SSE) and Hang Seng Index (HSI).

Spectral Co-Clustering

Spectral co-clustering is a data clustering technique that simultaneously clusters both the rows and columns of a matrix. This technique is based on the use of the eigenvalues and eigenvectors of a matrix to extract its low-dimensional structure.

Let A be an $\mathrm{n} \times \mathrm{m}$ non-negative matrix, where each row represents an object, and each column represents a feature. The goal of spectral co-clustering is to group together objects and features that are like each other.

We define a bipartite graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ where $\mathrm{V} = \mathrm{V}_1 \cup \mathrm{V}_2$ is the set of vertices and E is the set of edges. V_1 corresponds to the *n* rows of A and V_2 corresponds to the *m* columns of A. There is an edge between vertices *i* in V_1 and *j* in V_2 if and only if $\mathrm{A}(\mathrm{i}, \mathrm{j})$ is non-zero.

We can represent the bipartite graph G as an adjacency matrix B, where $\mathrm{B}(\mathrm{i}, \mathrm{j}) = 1$ if there is an edge between vertex i in V_1 and vertex j in V_2, and 0 otherwise. The diagonal matrix D_1 is defined as $\mathrm{D}_1(\mathrm{i}, \mathrm{i}) = \sum_{\mathrm{j}} \mathrm{B}(\mathrm{i}, \mathrm{j})$ and the diagonal matrix D_2 is defined as $\mathrm{D}_2(\mathrm{j}, \mathrm{j}) = \sum_{\mathrm{i}} \mathrm{B}(\mathrm{i}, \mathrm{j})$. The matrix $\mathrm{L} = \left[\mathrm{B} - \mathrm{D}_1^{1/2} \mathrm{A} \mathrm{D}_2^{1/2}\right]$ is the normalized Laplacian of the bipartite graph.

The co-clustering is performed by computing the k smallest eigenvectors of L,where k is the desired number of clusters. The rows of the matrix of k eigenvectors represent the cluster memberships of the objects, and the columns represent the cluster memberships of the features.

The spectral co-clustering algorithm can be summarized in the following steps:

- Construct the bipartite graph G from the matrix A;
- Compute the normalized Laplacian matrix L;
- Compute the k smallest eigenvectors of L;
- Cluster the rows and columns of the matrix of k eigenvectors into k clusters.

The spectral co-clustering algorithm can be extended to handle weighted and directed graphs, as well as to incorporate additional constraints on the clusters. This technique has been successfully applied to a variety of applications, including text mining, image segmentation, and gene expression analysis.

Vector Autoregressive Models

Vector Autoregressive (VAR) models are a type of time series model used to analyze the relationship between multiple variables over time. In a VAR model, each variable is modeled as a linear function of its own past values and the past values of all other variables in the system.

Suppose we have p variables, and we want to model the time series for each variable over *T* time periods. We can write the VAR(p) model as follows:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{J,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_J \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \phi_{1,1,i} & \phi_{2,1,i} & \cdots & \phi_{J,1,i} \\ \phi_{1,2,i} & \phi_{2,2,i} & \cdots & \phi_{J,2,i} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{1,J,i} & \phi_{2,J,i} & \cdots & \phi_{J,J,i} \end{bmatrix} \begin{bmatrix} y_{1,t-i} \\ y_{2,t-i} \\ \vdots \\ y_{J,t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{J,t} \end{bmatrix}$$

Equation 1 – General Expression for VAR model

In this equation, $\mathrm{y}_{\mathrm{j,t}}$ represents the value of the j-th variable at time t, c_{j} represents the intercept for the j-th variable, $\phi_{\mathrm{i,j,k}}$ is the coefficient of the j-th variable in the i-th equation and at the k-th lag, and $\varepsilon_{\mathrm{j,t}}$ is the error term for the j-th variable at time t. The coefficient matrix ϕ is a $\mathrm{J} \times \mathrm{J} \times \mathrm{p}$
array, where each $(\mathrm{i}, \mathrm{j}, \mathrm{k})$ element represents the coefficient of the j-th variable in the i-th equation and at the k-th lag.

To estimate the coefficients of the VAR model, we typically use Maximum Likelihood Estimation (MLE) or Bayesian methods. In MLE, we maximize the likelihood function given by the joint distribution of the p-dimensional time series, which is assumed to be multivariate normal.

Bayesian methods involve specifying prior distributions for the coefficients and then updating them using the data to obtain the posterior distribution. This allows us to incorporate prior knowledge about the parameters into the model.

Once we have estimated the VAR model, we can use it for a variety of purposes, such as forecasting or impulse response analysis. Impulse response analysis involves studying the response of the system to a shock in one of the variables.

Bi-Cluster Vector Autoregressive Model

Spectral co-clustering and VAR models are effective methods to analyze complex multivariate time series data. By identifying subgroups of variables through spectral co-clustering, we can analyze relationships between variables within each subgroup using VAR models. This approach can reveal relationships that might be hidden when examining the data as a whole. Furthermore, spectral co-clustering reduces the complexity of the analysis, which facilitates result interpretation. In conclusion, combining spectral co-clustering and VAR models can yield valuable insights for the analysis of multivariate time series data.

Co-clustering is a powerful method for classifying rows and columns in a matrix, allowing for the clustering of observations for each asset and identification of their behavior. By applying co-clustering to multivariate time series data, subgroups of variables that share similar patterns over time can be identified, which may reveal hidden relationships that are not apparent when analyzing the data as a whole. This approach can help to reduce the complexity of the analysis by breaking down the data into more manageable subgroups, allowing for a more focused examination of relationships within the data.

To combine spectral co-clustering and VAR models, a procedure can be followed as outlined below. First, apply spectral co-clustering to the multivariate time series data to identify subgroups of variables that exhibit similar patterns. This is done by constructing a matrix of the data, where each row represents a time step and each column represents a variable, and then performing spectral co-clustering on the matrix. Next, fit a VAR model to each subgroup of variables identified by the spectral co-clustering algorithm by estimating the coefficients of a linear model. To identify relationships between variables within each subgroup, examine the coefficients of the VAR models and search for patterns that suggest causal connections between variables. This analysis can help to uncover concealed relationships within the data that may be difficult to identify when analyzing the data as a whole, for this case study, we will identify 2 subgroups in our data.

The co-clustering algorithm is applied to the data set using the Spectral Co-clustering class, with 2 clusters, specified to identify assets with similar behavior.

VAR modeling is applied to each cluster of assets to investigate the causal relationships between them.

A Granger causality test is performed to identify the causal direction between each pair of assets in a cluster.

The p-value of the test is then stored in a Granger causality matrix, which is used to plot a heatmap that visualizes the causal relationships between the assets in a cluster.

In the context of Granger causality testing, a lower p-value indicates that the lagged values of one variable significantly improve the prediction of the other variable, and therefore suggests a causal relationship between the two variables. A p-value less than 0.05 is commonly used as a threshold for statistical significance, meaning that if the p-value is less than 0.05, we can reject the null hypothesis of no causality and conclude that there is likely a causal relationship between the two variables. In summary, the lower the p-value, the more significant is the relationship between the two variables.

Results

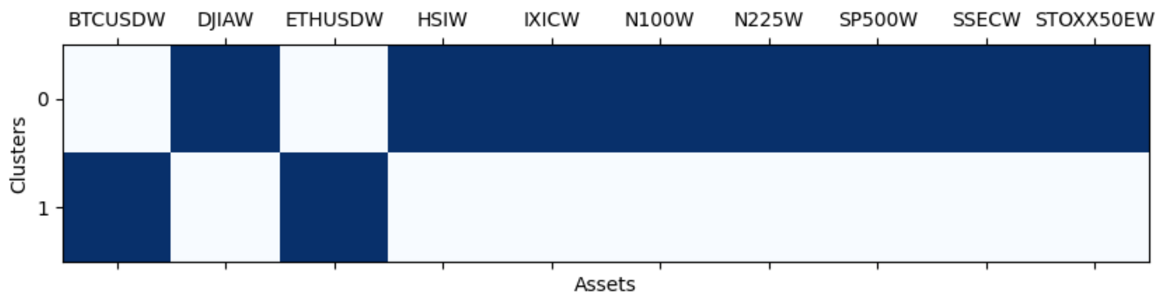

Figure 1 – Bi-Cluster of Markets

Alt text: There are two rows of squares defined by the colors Blue and White, which is determined by each type of market. The BTC and ETH representative markets are separated by the clustering methodology.

The previous figure displays how the Co-Clustering method can differentiate between traditional and digital assets, grouping Bitcoin and Ethereum in one cluster and the remaining assets in the second cluster. The ability of the Co-Clustering algorithm to identify and separate the two types of assets provides a valuable insight into the underlying similarities and differences between these types of assets. The grouping of Bitcoin and Ethereum in a separate cluster suggests that these two assets may share similar market characteristics, such as high volatility or unique trading patterns. Furthermore, by separating traditional assets from digital assets, the Co-Clustering method provides a useful tool for asset managers and investors to make informed decisions about their investment portfolios, given the differences in risk and return profiles between these two types of assets. Overall, the ability of the Co-Clustering method to identify and group similar assets provides a valuable tool for analyzing complex financial data and can yield insights that may not be apparent through other analysis methods.

In our study, we grouped together markets that had repeated variables. By doing so, we created subgroups consisting of three or more variables and applied the model to each of these subgroups. This allowed us to account for any potential correlations among the variables within each subgroup and obtain more accurate estimates for the parameters in our model.

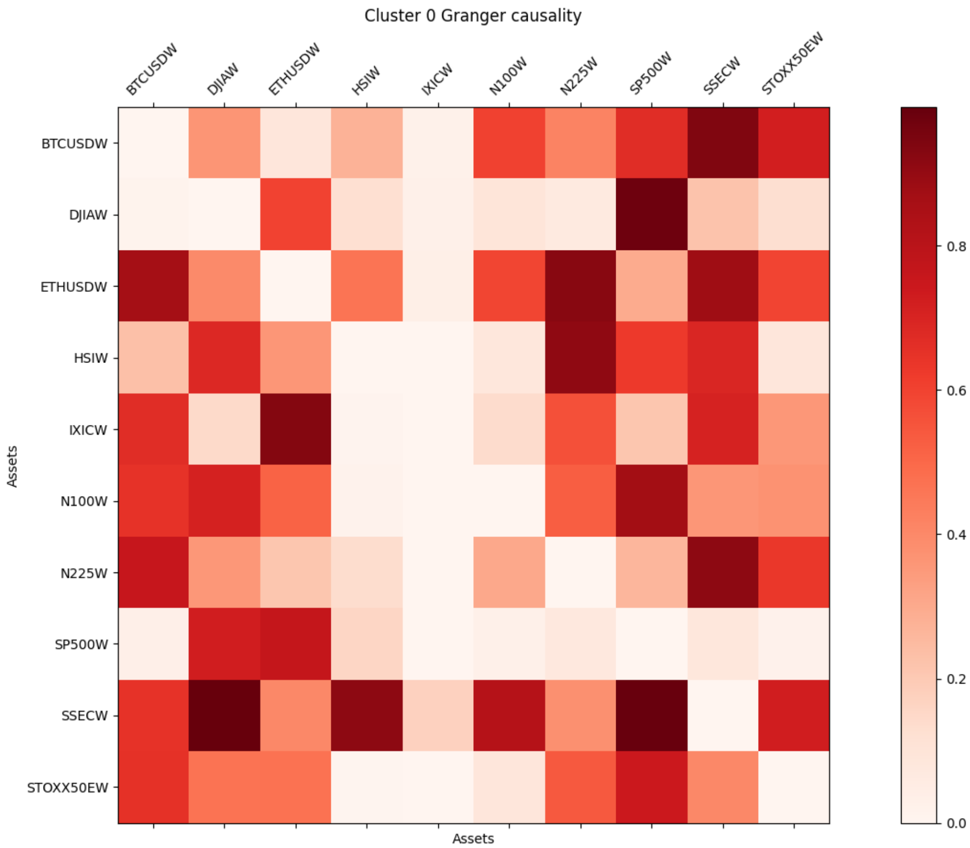

Figure 2 – Heat Map of Granger Causality Test for Cluster 0 (p-values)

Alt text: Heat Map of the Granger Causality Test for Cluster 0 where all the markets analyzed are clustered with colored variations of red squares.

There are several pairs of variables that have a significant relationship between them, as indicated by the p-values less than 0.05.

For example, BTCUSDW has a significant relationship with IXICW (p-value=0.0256) and N100W (p-value=0.022). ETHUSDW has a significant relationship with DJIAW (p-value=0.0127) and N100W (p-value=0.001). SP500W has a significant relationship with DJIAW (p-value=0.0342) and SSECW (p-value=0.0205).

On the other hand, there are some pairs of variables that do not have a significant relationship between them, as indicated by the p-values greater than 0.05. For example, HSIW and SSECW have a p-value of 0.0866, which is not significant. Similarly, IXICW and N225W have a p-value of 0.4678, which is also not significant.

In summary, the table shows that there are significant relationships between some pairs of variables, while others do not have a significant relationship. These findings can be used to inform further analysis of the relationships between these markets and help to identify potential dependencies or opportunities for diversification in an investment portfolio.

Variable 1	Variable 2	P-Value
BTCUSDW	IXICW	0.0256
DJIAW	IXICW	0.0301
DJIAW	SP500W	0.0127
DJIAW	SSECW	0.0364
ETHUSDW	IXICW	0.0136
N100W	IXICW	0.0174
N100W	SSECW	0.0328
N225W	SSECW	0.026
SSECW	STOXX50EW	0.026

Table 1 – Significant pairs on granger causality test

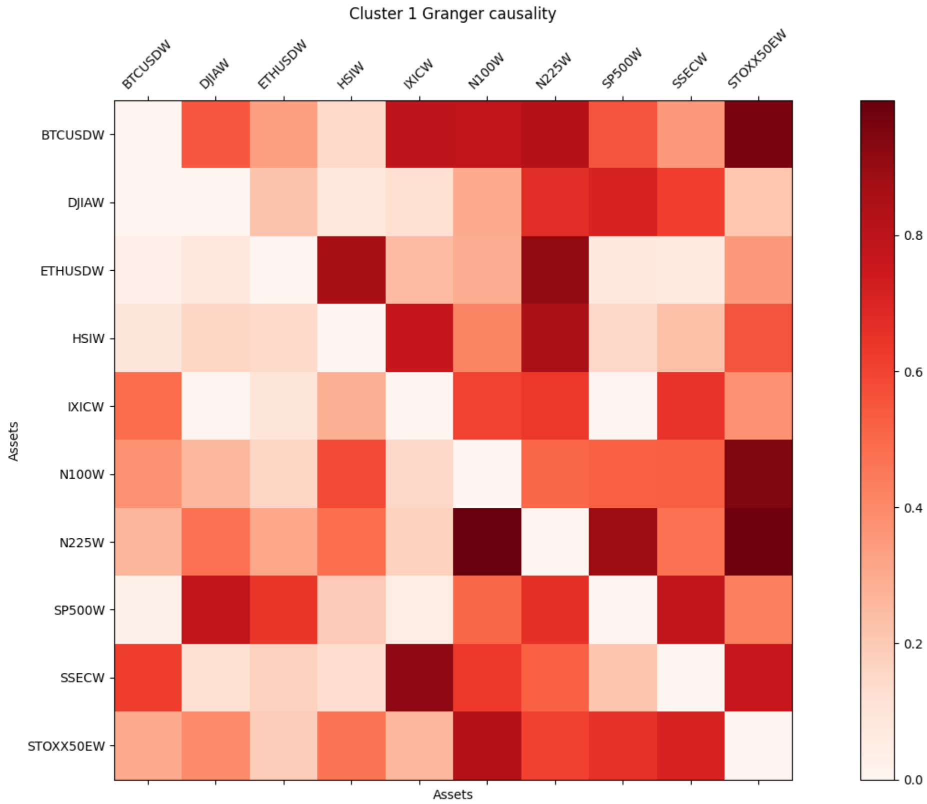

Figure 3 - Heat Map of Granger Causality Test for Cluster 1 (p-values)
Alt text: Heat Map of the Granger Causality Test for Cluster 1 where all the markets analyzed are clustered with colored variations of red squares.

We can see the pairwise Granger causality tests between different financial markets. The diagonal values are zero as they represent the causality of a variable with itself.

For example, BTCUSDW has a significant causal relationship with ETHUSDW (p=0.0037), and DJIAW has a significant causal relationship with SP500W (p=0.023). Additionally, some variables have relatively lower p-values, indicating a stronger causal relationship between them. For example, IXICW has a relatively strong causal relationship with DJIAW (p=0.0055), and SSECW has a relatively strong causal relationship with HSIW (p=0.068).

Overall, this matrix provides valuable insight into the interdependence between financial markets and highlights potential causal relationships between variables. However, it is important to note that Granger causality tests have limitations, and the results should be interpreted with caution.

Variable 1	Variable 2	P-Value
DJAW	N100W	0.0037
DJAW	SP500W	0.0235
ETHUSDW	IXICW	0.0055
ETHUSDW	N100W	0.0807
IXICW	SP500W	0.006
N225W	SP500W	0.0487
SSECW	SP500W	0.0214

Table 2 - Significant pairs on granger causality test

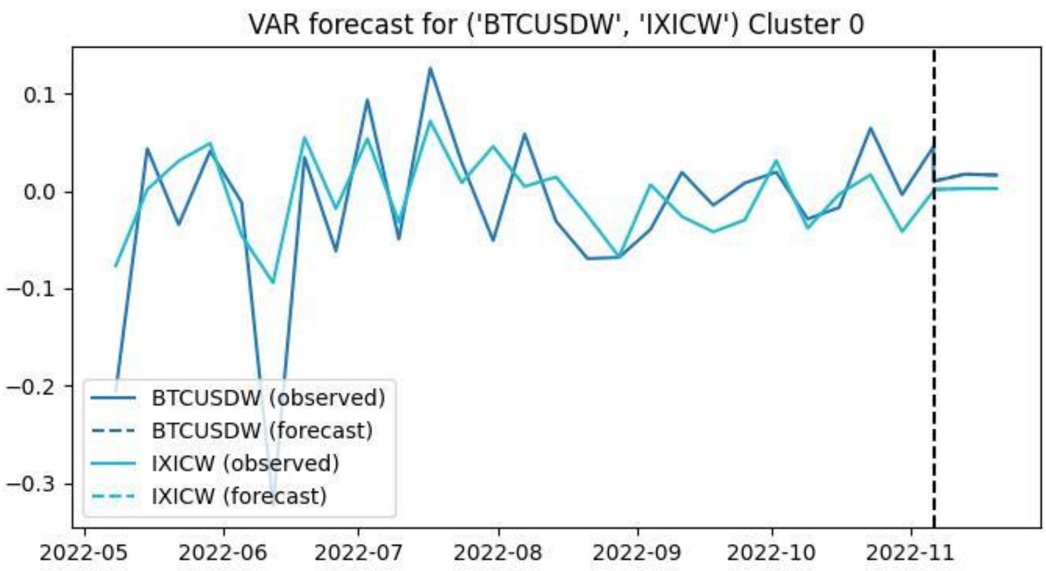

Figure 4 – Forecast for BTCUSD-IXIC

Alt text: The forecast for Bitcoin and Nasdaq Composite markets within a line graph, which is presented after a vertical dotted line; The forecast is stable.

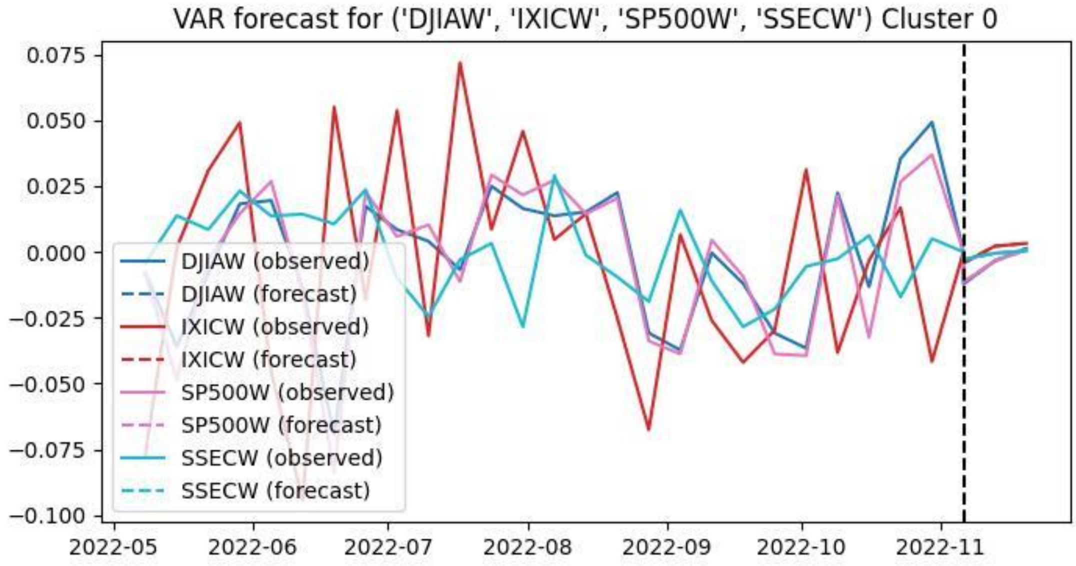

Figure 5 – Forecast for DJIA-IXIC-SP500-SSEC

Alt text: The forecast for Dow Jones Industrial Average, Nasdaq Composite, Standard and Poor's 500 and Shanghai Stock Exchange Composite markets within a line graph, which is presented after a vertical dotted line; The forecast is showing an upward trend.

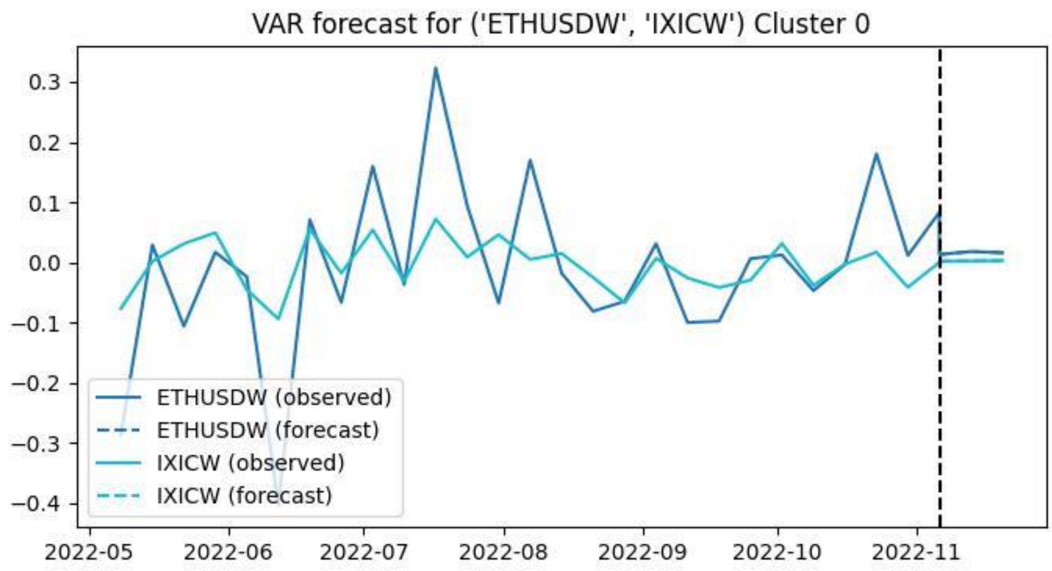

Figure 6 – Forecast for ETHUSD-IXIC

Alt text: The forecast for Ethereum and Nasdaq Composite markets within a line graph, which is presented after a vertical dotted line; The forecast is stable.

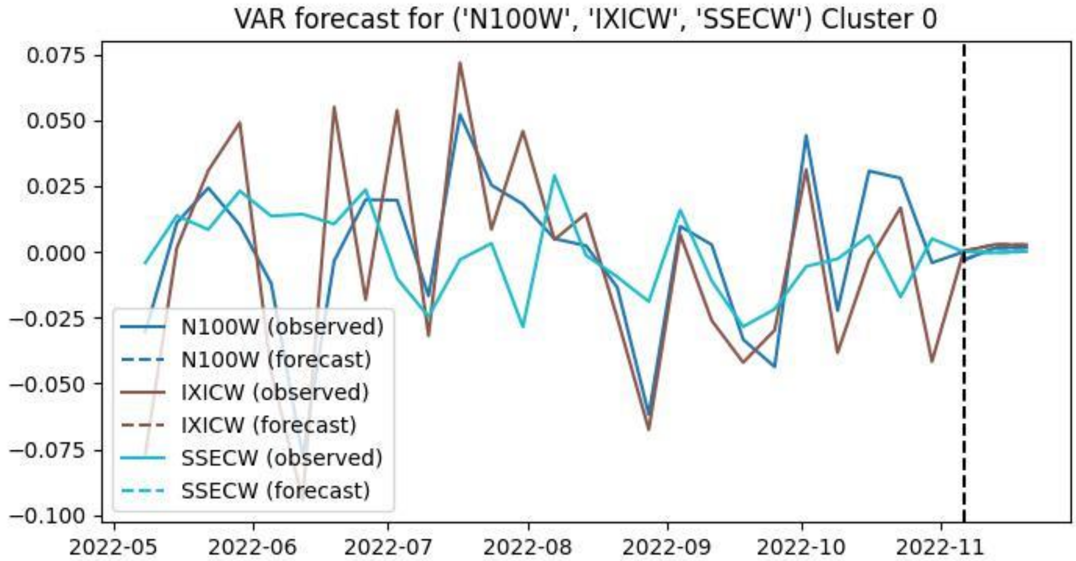

Figure 7 – Forecast for N100-IXIC-SSEC

Alt text: The forecast for Euronext 100, Nasdaq Composite and Shanghai Stock Exchange Composite markets within a line graph, which is presented after a vertical dotted line; The forecast is showing a slow upward trend.

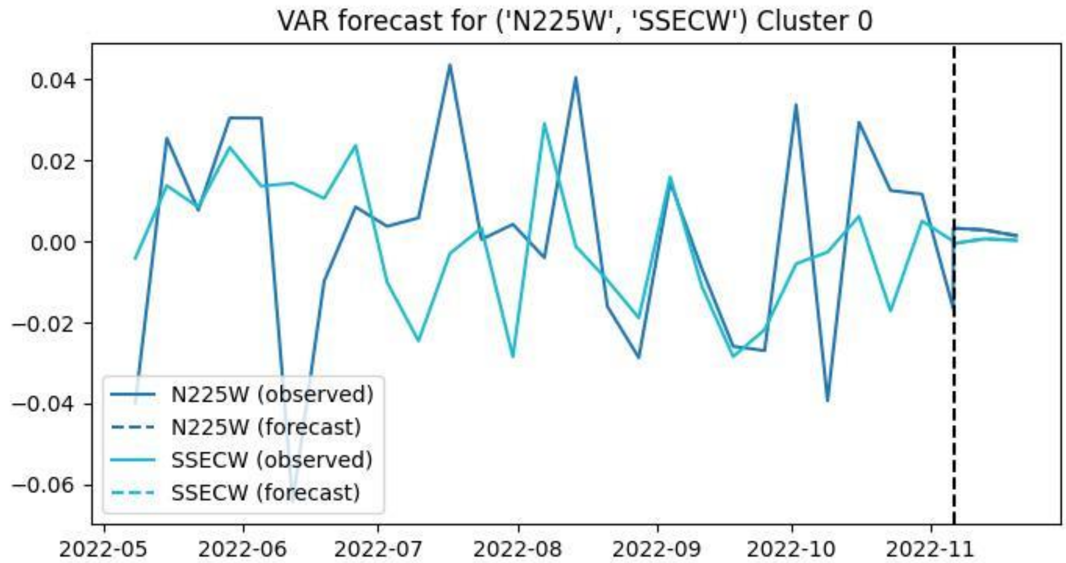

Figure 8 – Forecast for N225-SSEC

Alt text: The forecast for Nikkei 225 and Shanghai Stock Exchange Composite markets within a line graph, which is presented after a vertical dotted line; The forecast is stable for SSEC and shows a slow downward trend for N225.

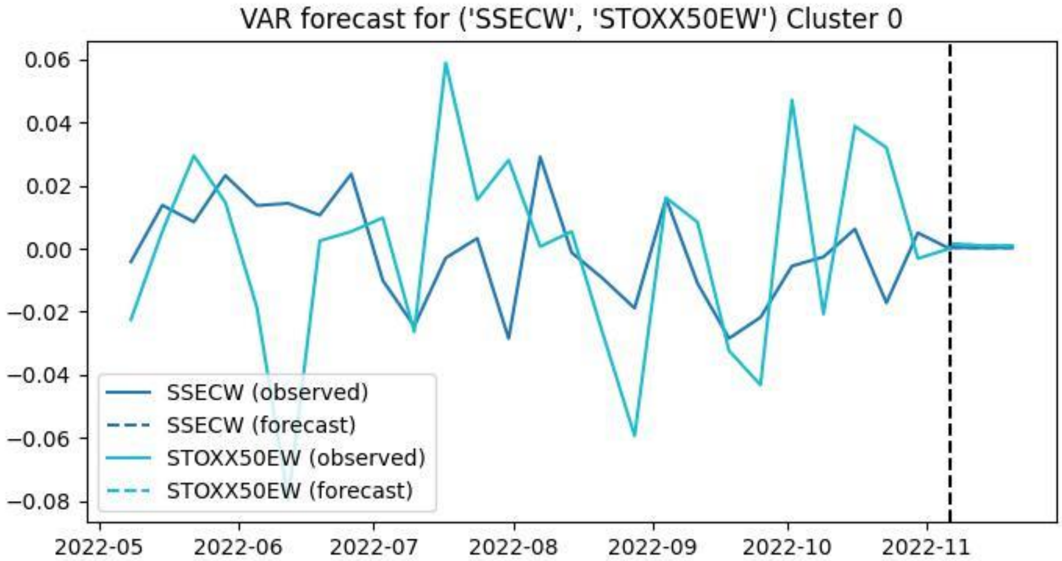

Figure 9 – Forecast for SSEC-STOXX50E

Alt text: The forecast for Shanghai Stock Exchange Composite and Euro Stoxx 50 markets within a line graph, which is presented after a vertical dotted line; The forecast is stable.

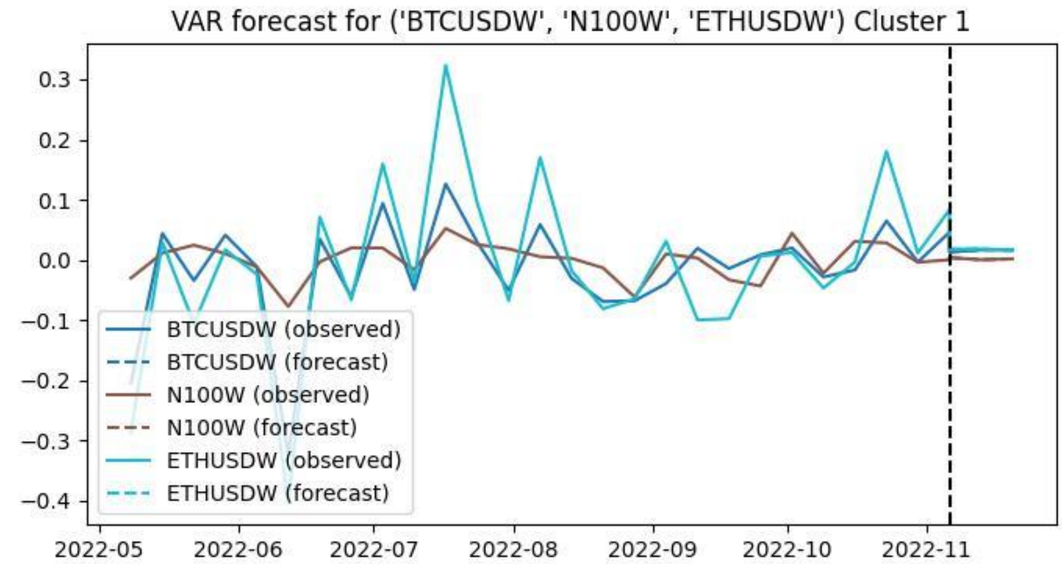

Figure 10 – Forecast for BTCUSD-N100-ETHUSD

Alt text: The forecast for Bitcoin, Euronext 100 and Ethereum markets within a line graph, which is presented after a vertical dotted line; The forecast is stable.

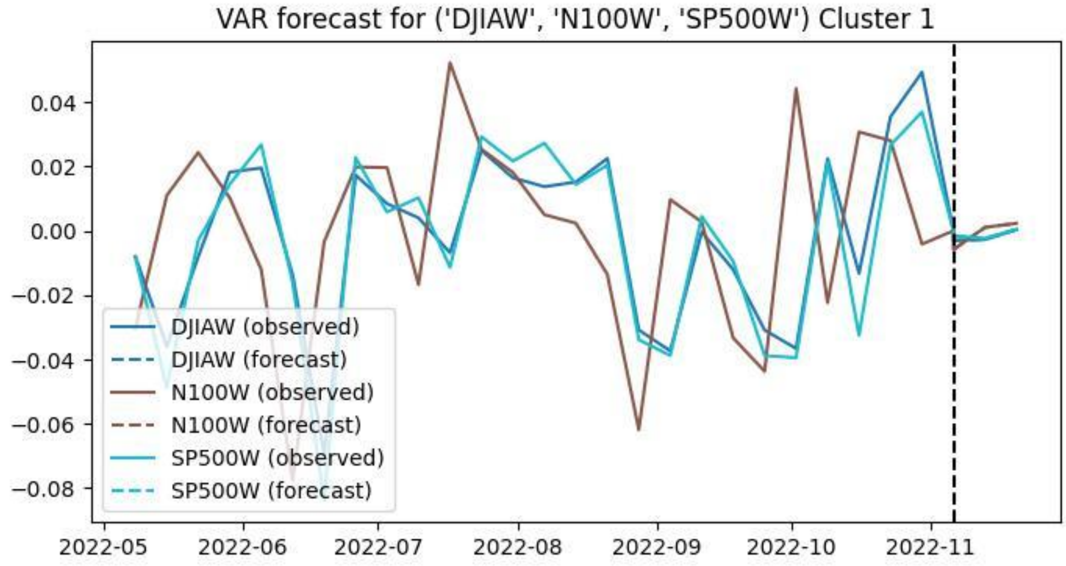

Figure 11 – Forecast for DJIA-N100-SP500

Alt text: The forecast for Bitcoin, Euronext 100 and Standard and Poor's 500 markets within a line graph, which is presented after a vertical dotted line; The forecast is showing an upward trend.

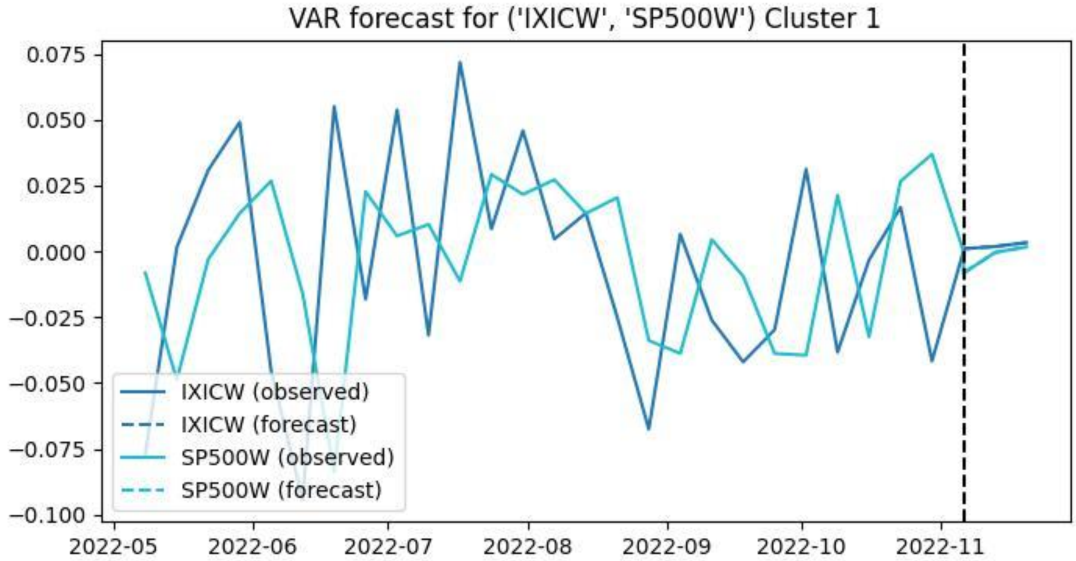

Figure 12 – Forecast IXIC-SP500

Alt text: The forecast for Nasdaq Composite and Standard and Poor's 500 markets within a line graph, which is presented after a vertical dotted line; The forecast is stable for IXIC and shows an upward trend for SP500.

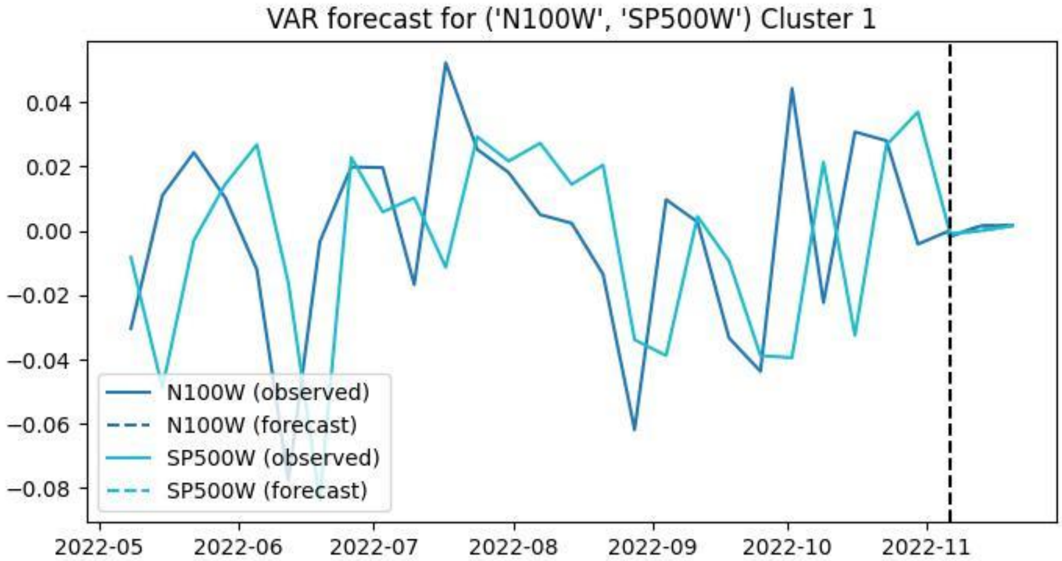

Figure 13 – Forecast for N100-SP500

Alt text: The forecast for Euronext 100 and Standard and Poor's 500 markets within a line graph, which is presented after a vertical dotted line; The forecast shows a slow upward trend.

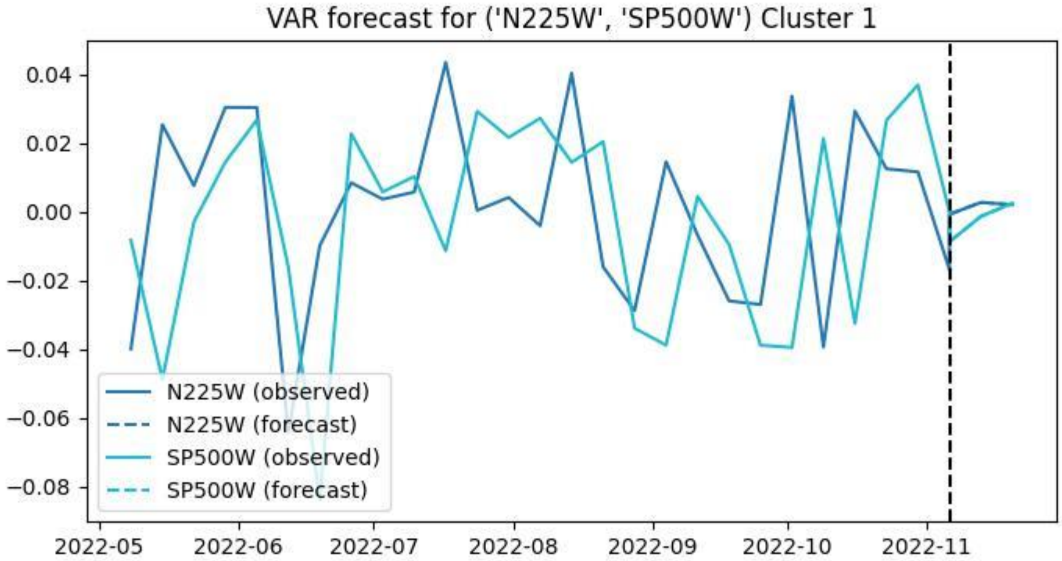

Figure 14 – Forecast for N225-SP500

Alt text: The forecast for Nikkei 225 and Standard and Poor's 500 markets within a line graph, which is presented after a vertical dotted line; The forecast is showing an upward trend.

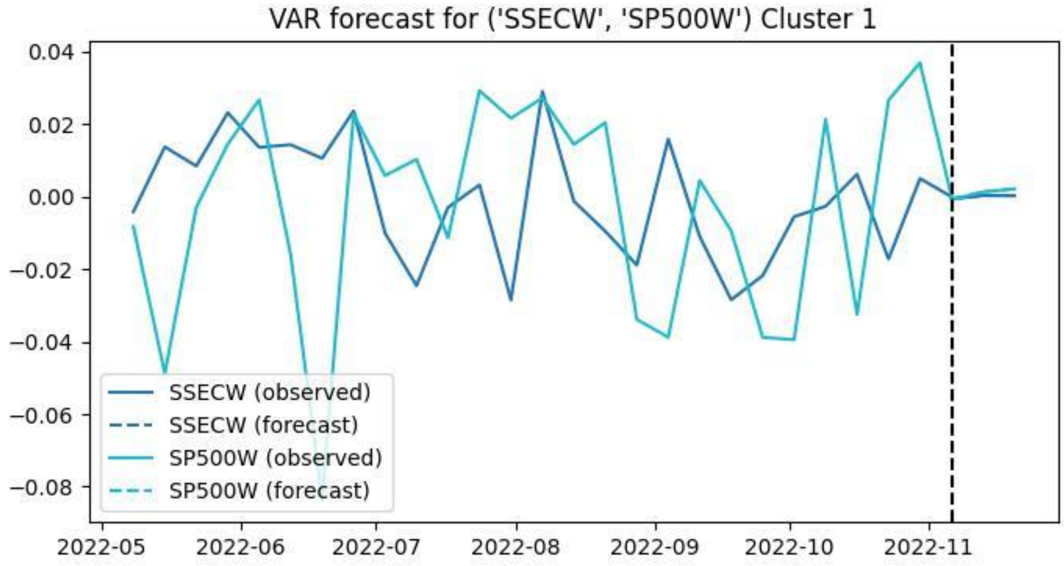

Figure 15 – Forecast for SSEC-SP500

Alt text: The forecast for Shanghai Stock Exchange Composite and Standard and Poor's 500 markets within a line graph, which is presented after a vertical dotted line; The forecast is stable for SSEC and shows a slow upward trend for SP500.

Conclusion

In this article, we explored the use of spectral co-clustering and vector autoregression (VAR) modeling to analyze multivariate time series data. The study applied these techniques to a dataset of asset prices from various markets, with a particular focus on identifying relationships between traditional and digital assets.

The results of the analysis showed that spectral co-clustering was effective in identifying subgroups of assets that exhibited similar patterns over time. VAR modeling was then applied to each subgroup to identify relationships between variables within the subgroup. The Granger causality test was used to assess the strength of these relationships, and the results were presented in the form of heatmaps.

The analysis revealed several interesting findings, including that Bitcoin and Ethereum exhibited a strong causal relationship within the digital asset subgroup, and that the US stock market had a significant causal effect on many other markets. Additionally, the study found evidence of contagion effects among the Asian markets during periods of financial crisis.

Overall, the study demonstrates the value of using spectral co-clustering and VAR modeling to analyze complex multivariate time series data. The techniques can help to uncover hidden relationships within the data and can provide valuable insights for investors and policymakers. However, as with any statistical analysis, it is important to interpret the results with caution and to consider the limitations and assumptions of the methods used.

With regards to the crypto assets, we can clearly observe that they are correlated between them as they does not pass the granger causality test.

References

Bennett S., Cucuringu M., Reinert G., (2022). Lead–lag detection and network clustering for multivariate time series with an application to the US equity market. *Mach Learn 111*, 4497–4538 (2022). Retrieved from https://doi.org/10.1007/s10994-022-06250-4

Chen Y., Yang H., (2016). A Novel Information-Theoretic Approach for Variable Clustering and Predictive Modeling Using Dirichlet Process Mixtures. Sci Rep 6, 38913. Retrieved from https://doi.org/10.1038/srep38913

Cucuringu M., Li H., Sun H., & Zanetti L. (2020). Hermitian matrices for clustering directed graphs: Insights and applications. *AISTATS*, pp 1–19. arXiv:1908.02096

Guangwei S., Liying R., Zhongchen M., Jian G., Yanzhe C., Jidong L., (2018). Discovering the Trading Pattern of Financial Market Participants: Comparison of Two Co-Clustering Methods. IEEE Access, *Special section on sequential data modeling and its emerging applications*. DOI 10.1109/ACCESS.2018.2801263

Nagy László, Ormos Mihály, (2018). Friendship of Stock Market Indices: A Cluster-Based Investigation of Stock Markets. *Journal of Risk and Financial Management*. 11. 88. 10.3390/jrfm11040088

Saliner JMG, (2019). Learning graphs from data and spectral clustering with applications to finance. Department of Electronic & Computer Engineering, Hong Kong University of Science and Technology

Sorensen K., (2014). Clustering in Financial Markets – A network theory approach. Royal Institute of Technology, School of Engineering Sciences. Retrieved from https://www.diva-portal.org/smash/get/diva2:754673/FULLTEXT01.pdf

Yosra B. S., Julien J., Sylvain A., (2022). Co-clustering for binary and functional data, *Communications in Statistics - Simulation and Computation*, Volume 51, 2022 - Issue 9, DOI 10.1080/03610918.2020.1764033