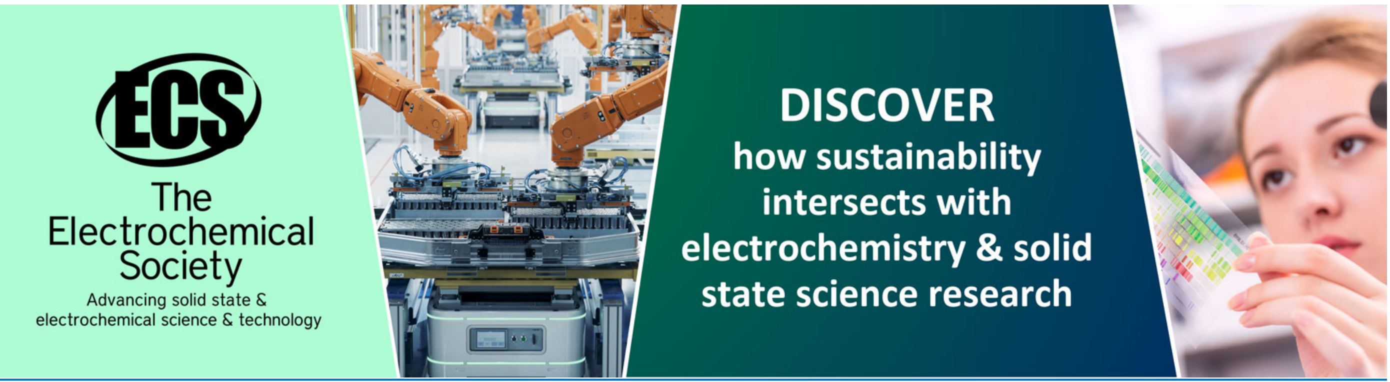
Journal of Physics: Conference Series

Detecting a currency's dominance using multivariate time series analysis



Detecting a currency's dominance using multivariate time series analysis

Nur Syahidah Yusoff [1] **and Shamshuritawati Sharif** [2]

[1]Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang, 26300, Gambang, Pahang, Malaysia
[2]School of Quantitative Sciences, UUM-College Arts and Science, Universiti Utara Malaysia, Sintok, Kedah, Malaysia

E-mail: wnsyahidah@ump.edu.my, shamshurita@uum.edu.my

Abstract. A currency exchange rate is the price of one country's currency in terms of another country's currency. There are four different prices; opening, closing, highest, and lowest can be achieved from daily trading activities. In the past, a lot of studies have been carried out by using closing price only. However, those four prices are interrelated to each other. Thus, the multivariate time series can provide more information than univariate time series. Therefore, the enthusiasm of this paper is to compare the results of two different approaches, which are mean vector and Escoufier's RV coefficient in constructing similarity matrices of 20 world currencies. Consequently, both matrices are used to substitute the correlation matrix required by network topology. With the help of degree centrality measure, we can detect the currency's dominance for both networks. The pros and cons for both approaches will be presented at the end of this paper.

1. Introduction

The study of the similarity measure and correlation between the currency's price is an important subject in a multivariate data setting. In this paper, we focus on a large set of currency exchange rate data set where each of currencies has four different daily prices; closing price, opening price, highest price and lowest price. This condition called for a multivariate time series approach instead of a univariate time series of one price only [1, 2].

The multivariate time series will provide more information than univariate time series since those four prices are interrelated to each other. Furthermore, a research work by Liangyue [3] found that a multivariate time series approach offers good prediction results. This consideration would be substantial advantages in modelling and prediction multivariate time series.

From comprehensive literature review, there are only two common strategies to measure the association between a set of variables in the case of multivariate time series, which are Canonical Correlation Analysis (CCA) and RV-coefficient [4]. The first strategy is proposed by Hotelling [5] in the year 1936. It seeks a linear combination of one set of variables that maximally correlated with a linear function of the other set of variables. While, the second strategy is Escoufier's RV coefficient (ERVC). This strategy is present by Escoufier [6] in year 1973 to measure of similarity between positive semidefinite matrices [7]. The coefficients are similar to the Pearson correlation coefficient.
In this paper, the first strategy is not appropriate because CCA discover a set of variables that are interrelated across sets, but not interrelated within the set. Therefore, the second strategy is used in our

IOP Conf. Series: Journal of Physics: Conf. Series **890** (2017) 012125 doi:10.1088/1742-6596/890/1/012125

study because each currency is related to each other currencies, and all the prices interrelated across four different prices.

To illustrate the pros and cons of the ERVC, the results of the analysis based on the mean vector (MV) approach are performed to show the performance of ERCV. Commonly, MV is the basic strategy that used for analysing the multivariate data. The computation of MV is actually the average of those four different types of prices, i.e., closing, opening, highest and lowest prices. Thus, both matrices from ERVC and MV are used to substitute the correlation matrix required by network topology.

A Matrix of correlation able to examine the similarity between pairs of datasets in a simple and comprehensive way. Generally, the correlation may take values between negative one and positive one. Matrix correlations have already a long history in multivariate analysis [8, 9]. Furthermore, with the help of a degree centrality measure, the currency's dominance for both networks can be identified. The rest of the paper is organised as follows. In the Section 2, we present the related methodology, followed by results and discussion of corresponding example in Section 3. At the end, this paper is closed with a conclusion in the last section.

2. Methodology

This section focused on the theories of ERVC and MV. However, to detect the currency's dominance of the ERVC and MV matrices is not an easy task. The complexity of multivariate time series analysis, and the cross relationship among currencies make it difficult for the analysis. This motivates us to introduce a network topology which is based on ERVC and MV as a further analysis to detect the currency's dominance.

In this paper, the currency exchange rate time series are analysed using 20 Western Europe currencies. List of 20 currencies is presented in table 1. The currencies data set is retrieved from Pacific Exchange Rate Service (http://fx.sauder.ubc.ca/EUR/analysis.html) started from February 2015 until April 2015. The analysis for the corresponding ERVC and MV are performed after the logarithm of the price of currency is computed from the original data. Therefore, the used data are independent and stationary [10].

Table 1. 20 Western Europe Currencies.

ALL	Albanian Lek
AMD	Armenian Dram
ANG	Dutch Guilder
BGN	Bulgarian Lev
BYR	Belarusian Ruble
CHF	Swiss Franc
CZK	Czech Koruna
DKK	Danish Krone
GBP	British Pound
GIP	Gibraltar Krone
HRK	Croatian Kuna
HUF	Hungarian Forint
ISK	Icelandic Krona
MKD	Macedonian Denar
NOK	Norwegian Krone
PLN	Polish Zloty
RON	Romanian New Leu
RUB	Russian Ruble
SEK	Swedish Krona
UAH	Ukrainian Hryvnia

2.1. Escoufier's RV coefficient (ERVC)

ERVC matrix starts with sample covariance matrix of the data. Let X be a n × p matrix and Y be a n × q matrix corresponding to two sets of variables defined for the same n individuals. Therefore, the ERVC [8, 11] can be defined as

$$RV_{\boldsymbol{XY}} = \frac{Tr\left(S_{\boldsymbol{XY}} S_{\boldsymbol{YX}}\right)}{\sqrt{Tr\left(S_{\boldsymbol{XX}}^2\right)}\sqrt{Tr\left(S_{\boldsymbol{YY}}^2\right)}}. \tag{1}$$

It is shown that the ERVC can be used as a measure of similarity of the two variables [12]. Furthermore, according to Zhang et al. [13], the statistic of ERVC is a good substitute for the Pearson correlation coefficient to measure the similarity of two variables. With this point of view, those relationships are eligible to be measured by using ERVC among variables.

2.2. Mean vector (MV)

MV is the basic step to do in analysing the multivariate data since by monitoring the structure of MV, the presence of special-causes of variation in that matrix can be detected. By adopting the theory in Anderson [14], the MV is defined as,

$$\bar{\boldsymbol{X}}_a = \frac{1}{k}\sum_{i=1}^{k} x_i \tag{2}$$

where

a = 1, 2, …, n; n represents number of day and
i = 1, 2, …, k; k represents four different prices per day, i.e., closing, opening, highest and lowest prices.

2.3. Network topology

In this study, network topology starts with ERVC and MV matrices with 20 x 20 orders, followed by transforming them into a distance matrix, D [10]. The element in the i-th row and j-th column of D is

$$d_{ij} = \sqrt{2(1-a_{ij})}\ ;\ 0 \le d_{ij} \le 2 \tag{3}$$

where

$a_{ij} = rv_{ij}$ for ERVC and
$a_{ij} = mv_{ij}$ for MV.

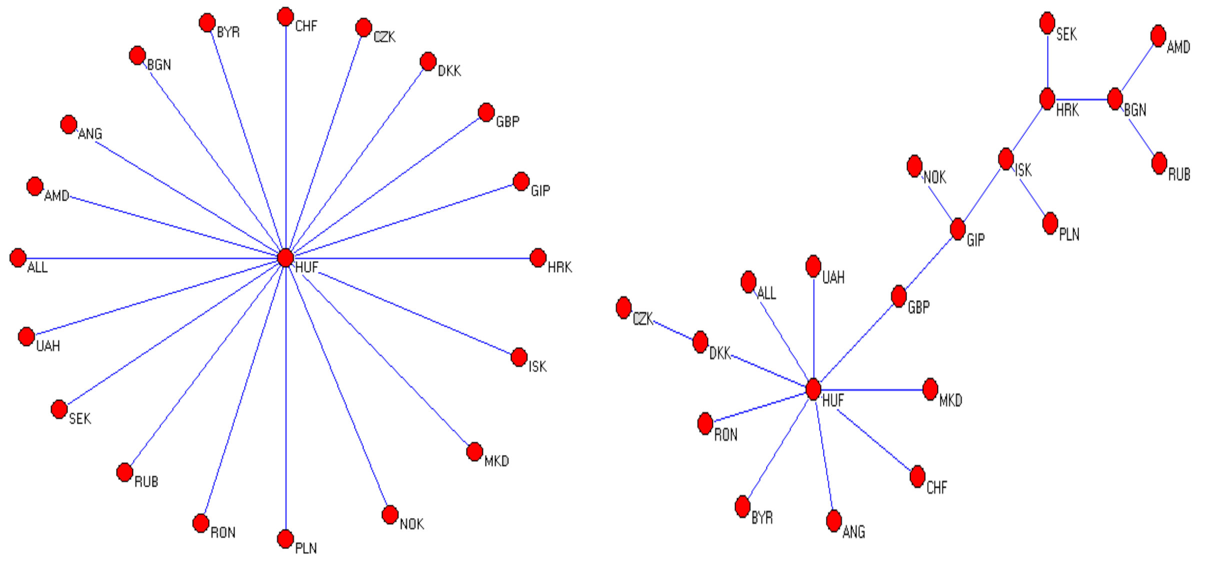
Small value of d_{ij} imply strong correlation between currencies. In this case, the matrix D represents a correlation network among variables. From this matrix, we construct a minimum spanning tree (MST) by using Kruskal's algorithm [15] provided in Matlab version 7.8.0 (R2009a). From MST, we construct the network topology of all currencies. This is a simplification of the high dimensional ERVC and MV matrices of currencies which will be used to summarise the most important information. An open source called 'Pajek' [1, 11, 16] will be used to visualise the network topology. Next, centrality measure can be used to enrich the interpretation of that network.

Conceptually, centrality is used to measure how central an individual is located in network [17]. Examples of commonly used measures of node importance include degree centrality, closeness centrality ([18-20]), betweeness centrality [18], eigenvector centrality, information centrality, flow betweeness centrality, the rush index centrality, local centrality [21], lobby index centrality [22], and evidential centrality [23]. In summary, the interpretation of that network is delivered by using the simplest centrality measure that is degree centrality measure [2, 24, 25].

3. Results and Discussion

ERVC and MV matrices consisting of 20 currencies as nodes connected by $([20-1]\times[20/2])=190$ links each of which corresponds to the ERVC and MV between two different nodes, respectively. However, by using the MST we only have to consider [20-1] = 19 links. MST is a subgraph that connects all the currencies (nodes) whose total weight, i.e., total distance is minimal.

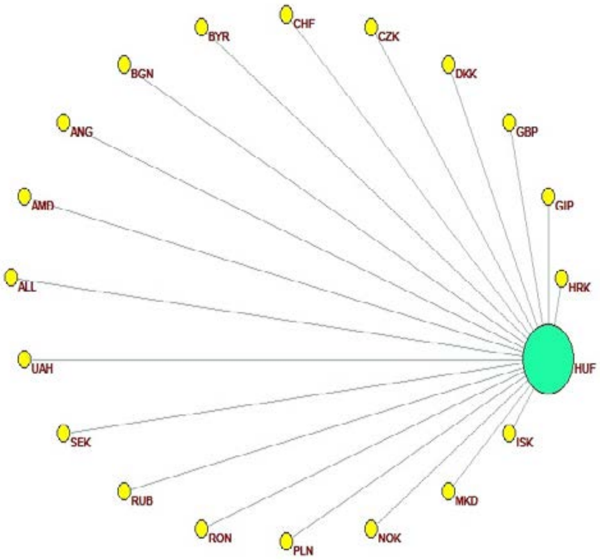
Figure 1 shows the corresponding MST for both approaches. This figure shows the most important relationship, i.e., the interconnectivity among all currencies in terms of MST. The larger the number of links is the more dominance of that particular node than the other. Based on MST, we learn that the most dominance currencies for both approaches is *HUF*.



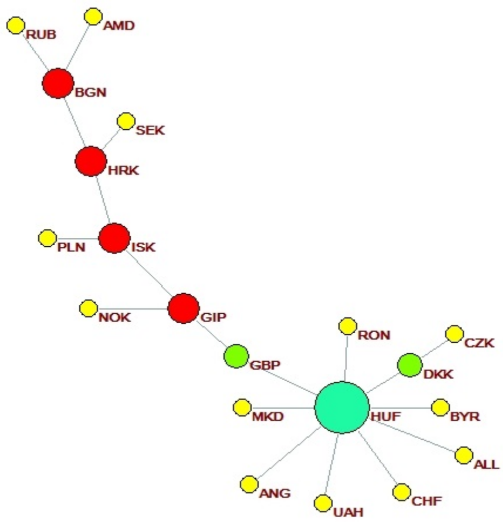
(b)

(a)

Figure 1. MST for ERVC (a) and MV (b).



(a)



(b)

Figure 2. Degree centrality measure for ERVC (a) and MV (b).

To elaborate the above finding more clearly, other information is presented using the degree centrality measure. This measure is defined as

$$C_{Degree}(N_i) = \sum_{j=1}^{p} a_{ij} \tag{4}$$

where a_{ij} is the element in *i*-th row and *j*-th column of an adjacent matrix and N_i is the *i*-th node.

In figure 2, the network topology where the colour of the node represents the rank of importance based on degree centrality is presented. The colours used in this analysis, ordered decreasingly in terms of the rank of importance: olive green, red, green and yellow. The higher the score of the centrality measures of a particular node, the more dominance that node is.

From figure 2 (a), for ERVC, *HUF* has the highest (olive green node) number of links, i.e., 19 links the in network. It plays the most important role in the network. This means that *HUF* is strongly influencing the others currencies. While for MV, *HUF* has the highest (olive green node) number links, i.e., 9 links, followed by *GIP*, *ISK*, *HRK* and *BGN* (red nodes), and *GBP* and *DKK* (green nodes). See figure 2 (b).

4. Conclusion

An analysis based on MST in Figure 1, we learn that all the currencies in ERVC are influenced by *HUF*. While, for MV, although the most dominance currencies is *HUF* but the others currencies still being influenced by *GIP*, *ISK*, *HRK*, *BGN*, *GBP* and *DKK*. Further analysis based on degree centrality measure leads us to the following conclusions.

(i) For ERVC, the most dominance currency is *HUF*.
(ii) For MV, the most dominance currencies, ordered decreasingly in terms of importance are *HUF*, *GIP*, *ISK*, *HRK*, *BGN*, *GBP* and *DKK*.

Therefore, according to these findings, we conclude that the most dominance currencies in ERVC are *HUF* for MV, the most dominance currencies are *HUF*, *GIP*, *ISK*, *HRK*, *BGN*, *GBP* and *DKK*. Consequently, these currencies should give special attention in Currency Exchange Rate.

Based on those analyses, it is important to state that ERVC can illustrate directly which one is the most dominance currencies undoubtedly compared to MV.

Acknowledgements
We acknowledge financial support from the Ministry of Higher Education, via RACE vote number RDU141306. The authors would like to thank Universiti Malaysia Pahang and Universiti Utara Malaysia for the opportunity to do this research. Special thanks go to the reviewers for the comments and suggestions.

References

[1] Yusoff N S and Sharif S 2015 *AIP Conf. Proc.* **1643** 335-340
[2] Sharif S, Yusoff N S and Djauhari M A 2012 *Modern Applied Science* **6** 35-43
[3] Liangyue C, Alistair M and Kevin J 1998 *Physica D* **22** 75-88
[4] Josse J, Pages J and Husson F 2008 *Computational Statistics and Data Analysis* **53** 82-91.
[5] Hotelling H 1936 *Biometrika* **28** 321-377
[6] Escoufier Y 1973 *Biometrics* **29** 751-760
[7] Herve A 2007 *Encyclopedia of measurement and statistics* 849-853
[8] Robert P and Escoufier Y 1976 *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **25** 257–265
[9] Yanai H 1974 *Journal of Behaviormetrics* **1** 45-54
[10] Mantegna R N and Stanley H E 2000 *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, United Kingdom)
[11] Yusoff N S and Sharif S 2016 *Adv. Sci. Lett.* **22** 4028-4031
[12] Holmes S 2006 *Probability and Statistics: Essays in Honor of David A. Freedman* **2** 219-233
[13] Zhang H, Tian J, Li J and Zhao J 2009 *Biomedical Applications in Molecular, Structural and Functional* **7262** 726222-1
[14] Anderson T W 2003 *An Introduction to Multivariate Statistical Analysis* vol 3 (John Wiley and Sons, Inc., New York)
[15] Kruskal Jr J B 1956 *Proceedings of the American Mathematical Society* **7** 48-50
[16] De Nooy M, Mvrar A and Batagelj V 2004 *Exploratory Social Network Analysis with Pajek*. (Cambridge: Cambridge University Press)
[17] Okamoto K, Chen W and Li X Y 2008 *Springer-Verlag Berlin Heidelberg* **5059** 186-195
[18] Freeman L C 1978 Centrality in social networks: Conceptual clarification. *Social Networks 1*, 215-239
[19] Wasserman S and Faust K 1994 *Social Network Analysis: Methods and Applications* (Cambridge University Press, New York, NY)
[20] Opsahl T, Agneessens F and Skvoretz, J 2010 *Social Networks* **32** 245-251
[21] Chen D, Lü L, Shang M S, Zhang Y C and Zhou T 2012 *Physica a: Statistical mechanics and its applications* **391** 1777-1787
[22] Campiteli M G, Holanda A J, Soares L D, Soles P R and Kinouchi O 2013 *Physica A: Statistical Mechanics and its Applications* **392** 5511-5515
[23] Wei D, Deng X, Zhang X, Deng Y and Mahadevan S 2013 *Physica A: Statistical Mechanics and its Applications* **392** 2564-2575
[24] Borgatti S P and Everett M G 2006 *Social Networks* **28** 466-484
[25] Sharif S, Ap N C and Ruslan N 2017 *AIP Conf. Proc.* **1847** 020013