Towards Extracting Causal Graph Structures from Trade Data and Smart Financial Portfolio Risk Management

Ployplearn Ravivanpong[1], Till Riedel[1] and Pascal Stock[2]

[1]Karlsruhe Institute of Technology, Kaiserstraße 12, 76131 Karlsruhe, Germany
[2]Frankfurt School of Finance and Management, Adickesallee 32-34, 60322 Frankfurt am Main, Germany

Abstract
Risk managers of asset management companies monitor portfolio risk metrics such as the Value at Risk in order to analyze and to communicate the risks timely to portfolio managers, and to ensure regulatory compliance. They must investigate the possible causes if a portfolio risk significantly increases or breaches a regulatory limit. However, monitoring can quickly become overwhelming, time and labor-intensive as each risk manager has to deal with over a hundred portfolios, numerous daily market data, and hundreds of risk factors of the supervised portfolios and of their securities. Particularly, understanding the interrelations between incidents in different portfolios beyond high level indicators is important. However, analyzing these interrelations manually is one of the most difficult tasks. In this paper, we describe and demonstrate how automatically generating causal graphs can address the capacity problem of practitioners in risk management, who are facing more and more capital markets based risk data daily on the portfolio level alone. Based on a proof of concept implementation, we compare a pairwise causal-inference-based approach with a clustering-based construction approach. We discuss the advantages and disadvantages of both approaches, both computationally and based on the resulting structure. Based on our initial findings, we outline further challenges and research topics.

Keywords
risk management, causal inference, agglomerative hierarchical clustering, network visualization

1. Introduction

Despite the fact that the financial risk management domain is already matured with regard to employed econometric models, is well regulated and requires high transparency, we believe that it can benefit from machine learning to improve daily portfolio risk management. A risk manager in an asset management company must maintain an overview of 100 to 250 portfolios daily. Each of these portfolios includes an aggregate of over 160 risk factors. The most important risk factors are equity risk factors for stocks, interest rates and yield-curve risk factors for fixed-income securities, and foreign exchange risk factors. Without modern risk management software and aggregate risk measures like the Value-at-Risk (VaR) that is required by the European and recently US regulation, it would be impossible for one risk manager to manage and analyze over 16,000 portfolio risk factors manually on a daily basis. Nonetheless, if there is a regulatory risk limit breach, a deep causal analysis into portfolios is necessary to mitigate risk and potential losses. Such analyses are time intensive and requires manually examining on average 98 securities, each having over 160 risk factors. An intelligent automation of the analyses is thus needed to enable more efficient root cause analyses.

Published in the Workshop Proceedings of the EDBT/ICDT 2022 Joint Conference (March 29-April 1, 2022), Edinburgh, UK
ployplearn.ravivanpong@kit.edu (P. Ravivanpong); till.riedel@kit.edu (T. Riedel); p.stock@fs.de (P. Stock)

When allocating a part of the portfolio to assets with a lower risk profile, e.g. German government bonds, portfolio managers take care that these assets are minimally correlated to those assets in the portfolio with high risks, e.g. emerging market stocks and bonds. Yet a causal relationship, undetected by traditional methods or unaware of by risk and portfolio managers, between underlying assets can unexpectedly lead to a high correlation between several portfolios of different types during a small crisis. In one instance the VaR ratio of an energy portfolio and a South East Asian (SEA) small-market-capitalization portfolio increased significantly. To trace the possible causes, a risk manager examined the marginal losses of assets in both portfolios. The stocks of the state-own Venezuelan oil company were found to be the cause for the increased VaR ratio of the energy portfolio. Only through examining model parameters, it was found out two days after the incident that nine SEA companies supplied machines to the Venezuelan oil company. Both portfolios indirectly suffered from the US-embargo on the Venezuelan government. Any technology that helps to uncover such a "causal mechanism" will significantly reduce the immense time needed to complete the analysis.

Such root cause analysis of a change in risk profile or limit breach, can be addressed by employing causal graph, graph visualization, and graph anomaly detection. If causal relationships between assets and market factors can be reliably derived, we can monitor their dynamics and anomalous development using network analysis and visualization to obtain an overview across all portfolios, even though the network may only suggest the causal chains.

Within an exploratory research project with a German asset management company, we applied existing methods to derive a causal graph of securities based on a pairwise analysis of a subset of portfolios. Our preliminary results, however, did not meet the user requirements. Consequently, we employed agglomerative hierarchical clustering (AHC) of portfolio risk profiles with network visualization as a simpler and more practical alternative. In this paper, we describe application requirements, our approaches to the technical as well as requirement challenges, and demonstrate the use of AHC and network visualization to support investment portfolio risk management in practice. We discuss the advantages and shortcomings of the approaches, and finally outline our perspective on further research topics for the presented application.

2. Background and Related Work

A causal graph or a causal diagram is a directed graph that visualizes causal relationships between variables. A node represents a variable. A directed edge $A \longrightarrow B$ means that A is the direct cause of B, i.e. changing A will result in the change in B, all else being equal [1]. To understand the interaction between market actors and risk factors, the information reflecting the real activity of companies (e.g. net cash-flow) should be used to derive the causal graph[1]. Since it is infeasible to acquire and curate structured data of such information, trade data are usually used as a proxy, assuming that those information are reflected in security prices according to the efficient market hypotheses [2]. Typically, a 1-day return $r_t = \frac{p_t}{p_{t-1}} - 1$ is used, which is the percentage change of today's (closing) price from the previous trading day. In this case, a node in the resulting causal graph presents a security or an economic factor, such as the interest rate or the oil price. An edge $A \longrightarrow B$ can be viewed generally as the Granger causal direction, i.e. A contains information to forecast B [3]. Risk managers can trace the paths of the information flow to the security in question and use them as a basis to determine the actual causation.

The choice of an algorithm to derive a causal graph from time series depends on the assumptions[2] that can be made about the underlying causal structure: linear vs. non-linear dependency, acyclicity, causal sufficiency (no unmeasured confounder [5]), and contemporaneous relationship. Although risk models typically assume linear statistical dependency, the presence of non-linear dependency has also been detected [6][7]. A cyclical relationship or a direction switch due to structural changes, between economic factors has also been documented [8]. The securities that are not publicly traded, also known as over-the-counter (OTCs), can be viewed as potential unmeasured confounders in the data. Yet, information on the current condition of unobserved companies may be reflected in the market data and the prices of publicly traded securities; because portfolio managers incorporate them in their the asset allocation strategies and trading behavior [2]. Assuming at least semi-strong market efficiency thus allow us to assume causal sufficiency to a certain extent. We do not take contemporaneous relationship into account yet, in order to avoid noises created by hyper-traders and short-live panic trade-behavior. Keeping these conditions in mind, the suitable causal discovery methods must be capable of deriving causal relationships from a large number of time series, are non-parametric and do not strictly impose acyclicity.

Considering the above assumptions together with the technical requirements from the risk management, existing causal inference methods[3] that are readily applicable, satisfy the majority of the conditions, and theoretically scalable are Effective Transfer Entropy (ETE) [10] and Peter-Clark Momentary Conditional Independence (PCMCI) method [11].

3. Data and Requirements

The choice of approaches depends not only on the available data but also on the technical and user requirements. The available data for our study are the internal portfolio risk data and the Deutsche Börse Public Dataset (PDS). The internal data consist of portfolio metadata, such as the ISIN, class of the assets, and their aggregate portfolio risk measures (e.g. UCITS gross exposure, present values) and risk factors (e.g. change in oil prices, interest rates). The portfolio risk measures data is longitudinal. Each portfolio has several time series of risk measures and factors. Since portfolios contain different asset types, like only stocks or bonds, or a mixture of them, some risk measures do not exist for all portfolios, resulting in missing values. The PDS consists of the initial price, lowest price, highest price, final price, and volume of all securities traded on the Eurex and Xetra trading systems aggregated in minute-interval [12]. The internal data set can be linked with the PDS using the asset ISIN. However, the investigated portfolios also consist of currency,

[1]The graph is also known as a financial network.

[2]Since we assume the structural causal model exists in our case, we also assume *Markov condition*. Additionally, we also assume *faithfulness*, allowing us to infer dependencies from the resulting graph [4].

[3]We initially also considered Temporal Causal Discovery Framework (TCDF) [9] because of its ability to detect the presence of hidden variables (which is beneficial yet not strictly necessary). However, the method was rejected due to its use of a complex black-box Attentive Convolutional Neural Network.

OTCs, and derivatives. These assets cannot be matched with the available trade data. Since the internal data are aggregated on a daily basis, only the daily closing price of the trade data is relevant.

Risk management requires a combination of high transparency, explainability, and timely detection as well as action. Transparency and explainability are required, not only for external communication to institutional investors and the regulator, like the German BaFin, but also for internal communication to portfolio managers. This currently means that all the variables that enter a model and the underlying assumptions (e.g. for proxy variables) must be known, while the reasons they are used and their effects on the models must be understood. Furthermore, the selected methods must enable a risk manager to derive and describe the mechanism of risk development to stakeholders and regulators. Hence, black-box models are rejected by default unless there is an acceptable justification while generated features must have economic or statistical meanings.

Given that it usually takes 48 hours for senior risk managers to manually find the root cause of a risk limit breach and that the complex traditional risk model calculation is finished overnight, the selected algorithm run-time must ideally be within 12 hours on the available compute resources (we assumed a workstation with 24 CPU cores as a realistic price tag).

4. Deriving Causal Graph of Underlying Assets

We initially chose Effective Transfer Entropy (ETE) to build the graph in our experiments. Transfer entropy (TE) is an information-theoretic measure that quantifies the information flow between two time series, while ignoring their static correlations due to common cause[13]. The method is non-parametric and thus can detect both linear and non-linear statistical dependencies. Moreover, it is well researched and widely applied in causality learning from financial data [6][14] [15][16][17]. We use the Rényi entropy as a basis for the TE because it uses a weighting parameter, which allows us to focus on different areas of distribution [10]. The ETE is the TE of the pair of original time series adjusted by the TE of the shuffled data. This enables the ETE method to be able to detect small effects, as a consequence of limited data and many variables, such is often the case of financial time series [6]. A significance test is performed on ETE of each security pair in each direction. Only edges with positive ETE and *p-value* less than a pre-defined threshold (e.g. 5%) are kept. The result is a causal graph[4].

The major drawback of the ETE is the number of hyperparameters to consider. This includes the number of lags, the discretization method, the Rényi entropy weighting parameter, the number of shuffles, and the number of bootstraps for statistical inference. A sensitivity analysis can be performed to get a robust result, but is computationally expensive. Also, it does not address the hidden variable problem and eventually suffers from the curse of dimensionality as the number of variables and lags gets larger [18]. Nevertheless, the ETE is chosen because the concept can be interpreted as a non-parametric Granger causality and its application is accepted in the financial domain.

The PCMCI is designed under the causal discovery framework and with parallelization in mind. It uses the PC algorithm to quickly first identify potential causes of an interested variable and prunes them using a conditional independence test. As a result, the algorithms can detect small effects, given a significance level. It also has fewer hyperparameters to consider. Apart from the number of lags and the number of variable combinations in the conditional independence test, users have only to specify the significance level to limit the false positive rates. By using the CMI non-parametric test, the method can identify both linear and non-linear dependencies [11]. Despite the algorithm assumptions of no hidden variable, acyclicity, and stationary time series, it is selected for our preliminary study because it is based on conditional independence test, a concept familiar in finance, can output a temporal causal graph, and is parallelizable.

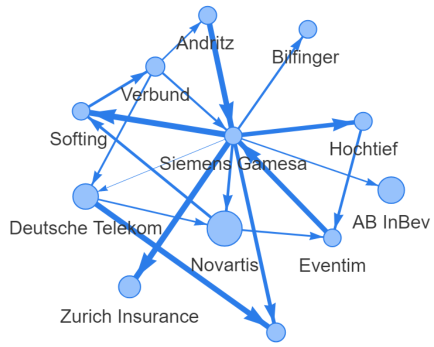

Figure 1: A subgraph of the causal graph identified by the ETE. The nodes are the stocks of the companies. Their size corresponds to market capitalization. The arrows represent the information flow. Its width indicates the size of ETE.

We applied both ETE and PCMCI on the 1-day returns of securities in the portfolios. Missing values due to non-trading activity are filled with the last closing price for simplicity. Seldom traded securities, whose data contain mostly missing values, are removed. The results were of

[4] [4] cautions that the size of TE should not be viewed as quantitative causal strength. Therefore, it can be interpreted at best as stronger or weaker dependencies, analogous to the interpretation of correlation.

mixed success. First, we encountered a computational challenge when applying both methods using their original implementation. Although both methods have been applied to financial data as a proof of concept, none of the experiments involve over seven hundred time series. Even though we managed to improve the parallelization of ETE, it still took almost six hours on 24 CPUs and 72 GB RAM for analyzing roughly 700 time-series of over two years daily returns of the investigated securities that are part of the given sample of 50 portfolios. The long run time for ETE was due to shuffling and bootstrapping in combination with the pairwise test. The CMI, which is a fully non-parametric test, in the second part of PCMCI has a long run time and does not scale well for a high number of time series. With the computational resources, that could be realistically be made available at the time, it could not be ensured that the run time will not significantly exceed 12 hours when analyzing additional thousands of assets and economic factors. As the algorithm is subject to quadratic scaling, we saw it as impractical given that we only had analyzed a relatively small subset. Also, the resulting causal graph, in which each asset is presented as a node, is complex to interpret both manually and automatically. An accuracy evaluation was infeasible without access to knowledge of existing risk models.

5. Agglomerative Hierarchical Clustering of Risk Profiles

Due to the complexities of asset-level causal inference as a basis for identifying risk drivers and their interrelationship, it makes sense to look for alternative approaches. One major requirement should be that such an approach: *(i)* intuitively shows the relationship to the monitored risk indicators; *(ii)* can be efficiently calculated automatically on a continuous basis; and *(iii)* creates structures that expose causal relationships beyond already established classification schemes.

In contrast to an eager bottom-up approach as presented above, an alternative might lie in a top-down approach that examines the indicator in question. The interrelationship between different portfolios should already be captured in the risk models, and thus be observable at the indicator level. Indicators have the advantage that they are already normalized and are designed for comparison. However, they are typically not used to structure information. This can be done by employing clustering analysis on the time series created by those indicators. While measuring distance between totally different kinds of portfolios in all market situations might be of limited value, subspace clustering approaches have the advantage that they can build hierarchies based on local similarities within subgroups. Our assumption is that the subgroups and their re-formation, given different market situations, might expose common risk drivers, without having to analyze the underlying assets.

For an exploratory study to uncover an alternative grouping of portfolios, a suitable clustering method must not require users to specify the number of clusters. Among them, agglomerative hierarchical clustering (AHC) is the most practical. The nearest-neighbor chain algorithm, which is available in most statistical software, is fast and always yields the same clustering if none of the instance pairs have the same minimum distance.The clustering result can be visualized as a dendrogram, regardless of the number of features, showing the class hierarchy. Due to these advantages, especially for the exploration of new categorization, we select the AHC for our application.

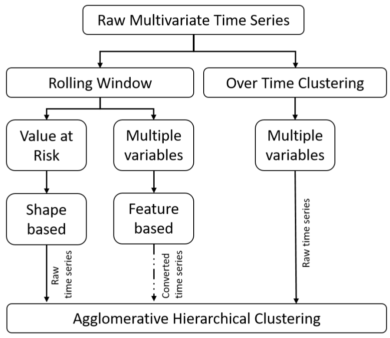

Figure 2: Overview of applicable approaches to multivariate time series clustering in risk management context.

The AHC method is originally developed for cross-sectional data. The adaption for longitudinal data clustering is a three-step decision process. First, one decides if the clustering should be performed along the time axis (*whole time series clustering* or *rolling window clustering*) or on all features, at each time point while ignoring the time axis (*over-time clustering*). The second step is to choose the similarity measure. Four approaches are usually found in practice [19][20]: shape-based, feature-based, model-based, and compression-based. In the last step, one chooses the clustering algorithm (*linkage method*). Since model-based and compression-based approaches generate complex features and transparency plays a major role in our case, we are limited to shape-based and feature-based similarity measures. Figure 2 summarizes the two main approaches for our use-case. Each of them presents different views about risk similarity.

5.1. Rolling window clustering

Rolling window clustering is a variation of *whole time series clustering*, which groups a set of similar time series together [19]. Since the risk and market indicators are updated almost daily and their previous year values should contribute little to their current trend, we only need to cluster the recent values and update the clustering with rolling windows. Initially, both shaped-based and feature-based approaches are under consideration. Even though the feature-based approach is more robust against missing values and is suitable for multivariate time series, it was also found during the preliminary analysis that the generated features add more interpretation difficulty and reduces quick communication. Fortunately, we can avoid the multivariate time series conversion and the missing value problems in the shape-based approach by using the VaR exposure ratio. The VaR exposure ratio is calculated from all relevant risk factors of a portfolio and scaled by its reference portfolio. Due to the strict regulatory requirements laid down in the German Derivative Regulation, it must be calculated for all regulated portfolios and may not contain missing values. Hence, the shape-based approach is preferred. The Ward's linkage method is selected for its ability to identify distinct clusters. Consequently, we are limited to the Euclidean distance.

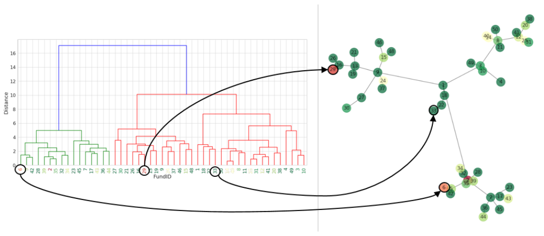

Figure 3: The dendrogram shows the clusters of portfolios in June 2017 based on the last 90-trading days VaR exposure ratio. The deep green portfolios always remain within while the deep red portfolios usually exceed the limit.

An example of the rolling window clustering of portfolios based on their VaR exposure ratio of the last 90 trading days results in groups of portfolios as shown in Figure 3. The network visualization is created by applying the minimum spanning tree (MST) algorithm to the cophenetic distance matrix [20]. The force-based Fruchterman-Rhiengold layout is the best suited for visualization as it put cluster members as to each other while separating the clusters from one another as much as possible. The traffic light color scheme corresponding to risk limit breach frequency helps risk managers to spot high-risk portfolios. A qualitative evaluation shows that portfolios of the same theoretical type are correctly grouped together, but with some exceptions. For example, the German equity portfolios lied close together in a cluster as expected. Yet, one of the portfolio manager's most important defensive multi-asset fund for retail investors and private wealth management was also among the cluster members. Such a close association between a defensive multi-asset fund (invested only 30% global stocks and 70% in bonds) and several German equity flagship funds (100% German stocks), was not obvious at first glance. A look into their shared risk factors within their cluster, which took a few minutes, showed the newly listed stock of Siemens Healthineers was the common cause; because it experienced an increase in volatility on its first trading days and all the portfolios were heavily invested in shares of the Siemens AG, the parent company. Such analysis that would have taken over an hour or more, even though the IPO of the spin-off was well known, became visible and obvious in a few minutes for the risk managers using the AHC and network visualization.

5.2. Over-time clustering

The result of the rolling window clustering is affected by the window size. Besides, by using only one aggregated indicator, the method ignores additional information that may be contained in other risk measures. In order to take into account multiple variables while keeping feature engineering at a minimum, one can treat the data at each time point as cross-sectional data and perform a cluster analysis on them. The approach is known as *over-time clustering* [21]. This means we make an assumption that all the data at each time point already contain all information from its lags, allowing us to ignore the time feature in the clustering. By keeping a fixed set of variables to be clustered, one can analyze the cluster dynamics over time and thus identify anomalous development [21].

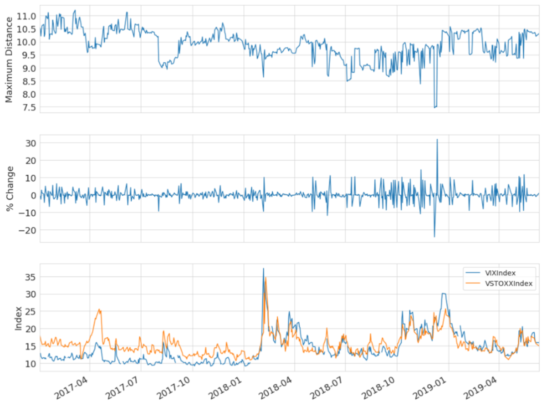

Figure 4: The overall distance at each time point (top), its daily percentage change (middle) and the VSTOXX and VIX volatility indices (bottom).

We applied the overtime clustering using Euclidean distance and Ward's linkage on around 30 risk measures.

Since we are interested in the dynamics, only variables that exist for all portfolios that seldom have missing values, excluding the VaR exposure ratio, are used. Figure 4 compares the final Ward's linkage distance or overall distance at each time point (top plot) to the VSTOXX and VIX volatility indices (bottom plot). They measure the implied volatility of the stocks in the EuroSTOXX50 index and the S&P500 index. These indices are commonly used as indicators of the overall stock market volatility, and thus the implied risk. The lower overall distance reflects the higher similarity between portfolio risk profiles, indicating that some common risk factors are exerting their influence across the board. A steep decrease in the overall distance is found to coincide with high market volatility. Our Granger causality test of information flow showed that the overall distance sometimes leads the volatility indices. However, since the quality of over-time clustering depends heavily on the set of selected variables, and maintaining the data quality and delivering them timely demands high effort, the overall distance is yet to be a useful addition to the traditional early warning indicator such as volatility indices.

6. Discussion

The applied research has shown that the use of AHC and network visualization helps risk managers to quickly get an overview over hundreds of portfolios. The most important insight from a practical perspective was the ability to see the effect of risk factors on portfolios across traditional asset classes like equities, bonds, foreign currencies, etc.. The traditional set-up of equity, multi-asset, and bonds portfolios, including overlay strategies using more complicated derivative strategies, did not allow a quick analysis of risk factors across asset classes and thus across different portfolio types. Focusing on the portfolio with the highest risk concentrations and the (theoretically, unlikely) neighboring portfolios in its cluster, allows risk managers to perform a more targeted root cause analysis before risk limit breaches occur. An analysis that would have taken over an hour can be carried out in a few minutes with the combination of AHC and network visualization.

The combination of AHC analysis on VaR exposure ratio and network visualization clearly supersedes the use of a causal graph regarding the practicality and technical requirements. Yet, the method is not without shortcomings. The window size affects the clustering results. Although sensitivity analysis can be performed and the results qualitatively evaluated, the window size cannot be systematically optimized; because we cannot quantitatively measure which clustering results are more reasonable in an exploratory analysis. Besides, the rolling window AHC is based only on the VaR exposure ratio, which is based on risk models that assume linear dependency and are handcrafted from expert knowledge. Thus the AHC only summarizes existing information and presents them in a more structured way to risk and portfolio managers. As such, it still misses signals that experts are unaware of and little new knowledge is gained. Since the resulting network is undirected, the AHC is unable to suggest the possible causal paths that could further reduce deep dive analysis effort should a risk limit breach occur. It is also debatable how early the AHC can visualize the weak effect of some common causes between portfolios before the effect becomes apparent. The AHC may be a promising and practical solution in the medium term. But a practical causal graph is desirable in the long term.

7. Conclusion and Outlook

If regulatory risk limit breaches or significant changes in the risk profile occur, a risk manager must manually analyze over ten thousand potential statistical sources, in order to support portfolio managers and institutional investors. In this paper we explored, how to extract supportive graph structures for this task based on real data provided to us by a large asset management company.

Our initial bottom up scheme to build such a causal graph based on public trade and portfolio data using ETE and PCMCI encountered a long run-time, and the resulting network that is too complex for human understanding. The absence of a reliable evaluation scheme prevented the approach to be readily used in practice. The combination of AHC of portfolio risk profiles based on their VaR exposure ratio and the network visualization of clustering results, is an easier deployable method and presents a more practical solution. While not being able to identify the causal chains directly, it fulfills the demand on transparency and gives risk managers a better overview to numerous statistical sources beyond existing categorization schemes and analysis strategies.

While we only present the results and experiences from a very preliminary proof of concept implementation, we believe that in the future the value of machine learning over manual analysis may increase if sources other than numerical trade data can be included into the analysis. Three challenges currently hinder practical application of causal discovery in risk management: scalability, visualization of a complex graph, and systematic evaluation of the resulting knowledge graph with minimal reliance on experts. We are currently investigating if recent developments like the two-phase ensemble algorithm that enables PCMCI to better utilize distributed systems [22] may address scalability. The problem of overly complex visualization needs to be addressed with graph-reduction techniques, showing to risk managers only subgraphs with unusual development. We plan to extend our work to use Causal NLP, a technique to learn causal graphs from text. It may be used to evaluate the causal graph derived from trade and economic data. Providing human-readable annotations might also solve problems of human oversight when dealing with a large data set.

Acknowledgments

This work was supported by the German Federal Ministry of Research (BMBF) as part of the Smart Data Innovation Challenge (01IS19030A)

References

[1] J. Pearl, Causality, 2 ed., Cambridge University Press, United Kingdom, 2009.

[2] E. F. Fama, Efficient capital markets: A review of theory and empirical work, The Journal of Finance 25 (1970) 383–417. URL: http://www.jstor.org/stable/2325486.

[3] C. W. J. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438. URL: http://www.jstor.org/stable/1912791.

[4] J. Peters, D. Janzing, B. Schlkopf, Elements of Causal Inference: Foundations and Learning Algorithms, The MIT Press, USA, 2017.

[5] R. Scheines, An Introduction to Causal Inference, University of Notre Dame Press, Notre Dame, 1997, pp. 185–200.

[6] R. Marschinski, H. Kantz, Analysing the information flow between financial time series: An improved estimator for transfer entropy, The European Physical Journal B 30 (2002) 275–281.

[7] A. Levin, A. Tchernitser, Chapter 11 - multifactor stochastic variance models in risk management: Maximum entropy approach and lévy processes*, in: S. T. Rachev (Ed.), Handbook of Heavy Tailed Distributions in Finance, volume 1 of *Handbooks in Finance*, North-Holland, Amsterdam, 2003, pp. 443–480.

[8] M. Beine, G. Capelle-Blancard, H. Raymond, International nonlinear causality between stock markets, The European Journal of Finance 14 (2008) 663–686.

[9] M. Nauta, d. Bucur, C. Seifert, Causal discovery with attention-based convolutional neural networks, Machine Learning and Knowledge Extraction 1 (2019) 312–340.

[10] S. Behrendt, T. Dimpfl, F. J. Peter, D. J. Zimmermann, Rtransferentropy — quantifying information flow between different time series using effective transfer entropy, SoftwareX 10 (2019) 100265.

[11] J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, D. Sejdinovic, Detecting and quantifying causal associations in large nonlinear time series datasets, Science Advances 5 (2019) eaau4996.

[12] D. B. Group, Deutsche börse public dataset (DBG PDS), 2021. URL: https://github.com/Deutsche-Boerse/dbg-pds, original-date: 2017-08-24T10:22:58Z.

[13] T. Schreiber, Measuring information transfer, Phys. Rev. Lett. 85 (2000) 461–464.

[14] T. Dimpfl, F. J. Peter, Using transfer entropy to measure information flows between financial markets, Studies in Nonlinear Dynamics and Econometrics 17 (2013) 85–102.

[15] P. Jizba, H. Kleinert, M. Shefaat, Rényi's information transfer between financial time series, Physica A: Statistical Mechanics and its Applications 391 (2012) 2971–2989.

[16] L. Junior, A. Mullokandov, D. Kenett, Dependency relations among international stock market indices, Journal of Risk and Financial Management 8 (2015) 227–265.

[17] S. K. Stavroglou, K. Soramaki, K. Zuev, Causality networks of financial assets, The Journal of Network Theory in Finance 3 (2017) 17–67. doi:10.21314/JNTF.2017.029.

[18] J. Runge, J. Heitzig, V. Petoukhov, J. Kurths, Escaping the curse of dimensionality in estimating multivariate transfer entropy, Physical Review Letters 108 (2012) 258701.

[19] S. Aghabozorgi, A. Seyed Shirkhorshidi, T. Ying Wah, Time-series clustering – a decade review, Information Systems 53 (2015) 16–38.

[20] J. Papenbrock, Asset Clusters and Asset Networks in Financial Risk Management and Portfolio Optimization, Ph.D. thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany, 2011.

[21] M. Tatusch, G. Klassen, M. Bravidor, S. Conrad, Show me your friends and i'll tell you who you are. finding anomalous time series by conspicuous cluster transitions, in: T. D. Le, K.-L. Ong, Y. Zhao, W. H. Jin, S. Wong, L. Liu, G. Williams (Eds.), Data Mining, volume 1127, Springer Singapore, Singapore, 2019, pp. 91–103.

[22] P. Guo, Y. Huang, J. Wang, Scalable and flexible two-phase ensemble algorithms for causality discovery, Big Data Research 26 (2021) 100252.