Trimmed fuzzy clustering of financial time series based on dynamic time warping

Pierpaolo D'Urso[1] · Livia De Giovanni[2] · Riccardo Massari[1]

Abstract

In finance, cluster analysis is a tool particularly useful for classifying stock market multivariate time series data related to daily returns, volatility daily stocks returns, commodity prices, volume trading, index, enhanced index tracking portfolio, and so on. In the literature, following different methodological approaches, several clustering methods have been proposed for clustering multivariate time series. In this paper by adopting a fuzzy approach and using the Partitioning Around Medoids strategy, we suggest to cluster multivariate financial time series by considering the dynamic time warping distance. In particular, we proposed a robust clustering method capable to neutralize the negative effects of possible outliers in the clustering process. The clustering method achieves its robustness by adopting a suitable trimming procedure to identify multivariate financial time series more distant from the bulk of data. The proposed clustering method is applied to the stocks composing the FTSE MIB index to identify common time patterns and possible outliers.

Keywords Robust fuzzy C-medoids clustering · Trimming · Dynamic time warping · Multivariate financial time series · FTSE MIB index

1 Introduction

The literature on time series clustering has increased considerably over the last two decades, with a large range of empirical applications in many different fields such as economics, finance, econophysics, environmental sciences, neuroscience, social sciences, among others.

✉ Riccardo Massari
 riccardo.massari@uniroma1.it

 Pierpaolo D'Urso
 pierpaolo.durso@uniroma1.it

 Livia De Giovanni
 ldegiovanni@luiss.it

1 Department of Social Sciences and Economics, Sapienza - University of Rome, P.za Aldo Moro, 5, 00185 Rome, Italy

2 Department of Political Sciences, LUISS Guido Carli, Rome, Italy

As for financial analysis in particular, the clustering methods are useful and fruitful for clustering time series related to data on stocks quoted on the stock exchange. For instance, it is useful to obtain a clustering of a set of companies quoted in the Market Exchange by analysing the performances of the shares in a time interval; in this way, the results of the cluster analysis can provide precious information for suitably and promptly planning the policy and the strategic assessment and the behaviour of the operators that invest money in stocks. At the same time, since several pieces of information are available on the performance of each company's share, e.g., opening and closing prices, highest and lowest quotations, and volumes of shares traded on each trading day, then it is of some importance to consider them all in the clustering process.

In literature various clustering methods based on different methodological approaches have been suggested for classifying financial time series. Pattarin et al. (2004) proposed a procedure for clustering financial time series with an application to mutual funds style analysis. Dose and Cincotti (2005) proposed a method for clustering financial time series with application to index and enhanced index tracking portfolio. Basalto et al. (2007, 2008) adopted the Hausdorff distance in hierarchical clustering for financial time series. Caiado and Crato (2007) suggested a GARCH-based method for clustering of financial time series applied to international stock markets. Lai et al. (2009) clustered stock prices time series from Taiwan Stock Exchange Corporation combining fuzzy decision trees and genetic algorithm clustering technique. By adopting a non-parametric approach, Vilar et al. (2009) mined financial data by first identifying trends via local polinomial regression. De Gregorio and Iacus (2010) classified financial assets from NYSE/NASDAQ by introducing a new metric based on the Markov operator. Maharaj et al. (2010) suggested a wavelet-based fuzzy clustering method of time series and applied their technique to daily returns of sixteen emerging market indices and eighteen developed market indices. De Luca and Zuccolotto (2011) suggested a tail dependence-based dissimilarity measure for financial time series clustering. The authors further analysed the clustering of financial time series aimed at detecting similarities in the extreme events (De Luca and Zuccolotto 2017). Piccardi et al. (2011) borrowed from network analysis a method for clustering financial time series. D'Urso et al. (2013) proposed two GARCH-based fuzzy clustering methods for classifying financial time series related to dailies' returns of Euro exchange rates. Bastos and Caiado (2014) introduced a new metric for classify financial time series, which shows stableness of results despite the high volatility of the data at hand. Durante et al. (2014) proposed a method for clustering financial time series in risky scenarios. Dias et al. (2015) explored a large set of financial time series with a model-based approach, identifying different regimes switching dynamics. Menardi and Lisi (2015) monitored the mutual funds market with the double aim of dimensional reduction of data and to identify similar patterns between monthly data of mutual funds. D'Urso et al. (2016) proposed different robust fuzzy clustering methods based on GARCH models for clustering time series; the authors illustrated the usefulness of these suggested clustering methods by means of two applications to financial and economic time series of the international stock-market volatility daily index returns, and the volatility daily stocks returns in FTSE MIB index. A method which proved to scale well with large collections of financial time series was introduced by Ando and Bai (2017) Nair et al. (2017) introduced a recommender system for traders and laymen by clustering stock price time series by means of the Self Organizing Maps. Aslan et al. (2018) suggested a clustering method for time series via threshold autoregressive models with application to commodity prices. Rahmanishamsi et al. (2018) applied their clustering method time series of Euro exchange rates against twenty–eight international currencies. Vilar et al. (2018) proposed clustering methods based on quantile autocovariances for time series and illustrated their application to daily stocks

returns in IBEX-35 index. A copula-based approach to cluster financial instruments was introduced by Yang et al. (2018). Lafuente-Rego et al. (2019) proposed different robust fuzzy clustering methods based on quantile autocovariances for time series. The authors applied their methods to two financial time series, i.e., the daily returns of Euro exchange rates and the daily stocks returns in IBEX-35 index. Nakagawa et al. (2019) classified price fluctuation by means of a k-medoids clustering algorithm with dissimilarities between time series computed via the Dynamic Time Warping method.

Most of the literature regards univariate financial time series. Still, several features can be observed on the time pattern of a financial asset, as already pointed out. Hence, there is scope to generalize the cluster analysis to multivariate financial time series. In this paper, capturing the time-varying features of the multivariate financial time series by means of the Dynamic Time Warping and adopting the Partitioning Around Medoids (PAM) technique in a fuzzy framework, we propose a robust time series clustering method based on the "impartial" trimming procedure (García-Escudero and Gordaliza 1999; García-Escudero et al. 2003, 2010). Dynamic time warping can deal with multivariate time series with no need for specific statistical assumption on the data generating process. Combining dynamic time warping and the impartial trimming approach is then particularly useful for classifying multivariate financial time series in presence of possible outliers, i.e., multivariate time series whose dynamics deviate markedly from the rest of data, even for only one of their components.

The remaining of the paper is organized as follows. In Sect. 2 we illustrate the methodological aspects of the proposed fuzzy clustering method for multivariate financial time series. In Sect. 3, we apply our clustering method to the set of stocks composing the FTSE MIB index. Some conclusions are shown in Sect. 4.

2 Methodology

2.1 Comparing multivariate financial time series with dynamic time warping

The dynamic time warping (DTW) distance (Velichko and Zagoruyko 1970; Berndt and Clifford 1994; Ratanamahatana and Keogh 2004; Izakian et al. 2015) for multivariate time series stretches or compresses the patterns of two objects locally, in order to make their shape as similar as possible. DTW distance allows for differences in length and phase in the multivariate time series, thus proving to be most flexible than the Euclidean distance (Izakian et al. 2015). Moreover, it can be conveniently computed on raw data, since there is no need for *a-priori* assumptions on the data distribution.

In the following, we first show how to conveniently represent multivariate time series for clustering sake. Then, we briefly describe the DTW computation in the case of multivariate time series.

2.1.1 Algebraic representation of multivariate time series

Multivariate time series are usually algebraically represented as a three-way data array (D'Urso 2000, 2004, 2005), or time data array, of

$$\mathbf{X} \equiv \{x_{ijt} : i = 1, \ldots, I; \ j = 1, \ldots, J; \ t = 1, \ldots, T\} \tag{1}$$

where i indicates the generic unit (object), j the variable, and t the generic time; x_{ijt} represents the j-th variable observed in the i-th unit at time t.

For classification task, the time data array $\mathbf{X}$ can be conveniently expressed as the set of the bi-dimensional matrices

$$\mathbf{X} \equiv \{\mathbf{X}_i \, : \, i = 1, \ldots, I\}. \tag{2}$$

The matrix $\mathbf{X}_i \equiv \{\mathbf{x}_{it} \, : \, t = 1, \ldots, T\}$ represents the i-th multivariate time trajectory, where $\mathbf{x}_{it} \equiv (x_{i1t}, \ldots, x_{ijt}, \ldots, x_{iJt}), \ i = 1, \ldots, I, \ t = 1, \ldots, T$.

2.1.2 DTW computation

Let $\mathbf{X}_i \equiv \{\mathbf{x}_{i1}, \ldots, \mathbf{x}_{it}, \ldots, \mathbf{x}_{iT}\}$ and $\mathbf{X}_{i'} \equiv \{\mathbf{x}_{i'1}, \ldots, \mathbf{x}_{i't'}, \ldots, \mathbf{x}_{i'T'}\}$ two multivariate time series, where T and T' need not be identical. In the DTW jargon, they are defined as the "query" (or test) and the "reference" object respectively.

The total distance between $\mathbf{X}_i$ and $\mathbf{X}_{i'}$ is then computed through the so called "warping curve", or "warping path", which ensures that each data point in $\mathbf{X}_i$ is compared to the "closest" data point in $\mathbf{X}_{i'}$.

Let

$$\Phi_l = (\varphi_l, \psi_l), \ l = 1, \ldots, L.$$

be a set of realigned indices $1, \ldots, T$ and $1, \ldots, T'$ under the following constraints:

1. boundary condition: $\Phi_1 = (1, 1), \ \Phi_L = (T, T')$;
2. monotonicity condition: $\varphi_1 \leq \ldots \leq \varphi_l \leq \ldots \leq \varphi_L$ and $\psi_1 \leq \ldots \leq \psi_l \leq \ldots \leq \psi_L$.

The total dissimilarity between the two "warped" multivariate time series is:

$$\sum_{l=1}^{L} d(\mathbf{x}_{i,\varphi_l}, \mathbf{x}_{i',\psi_l}) m_{l,\Phi}$$

where

- $m_{l,\Phi}$ is a local weighting coefficient which ensures that the accumulated cost function is comparable when two multivariate time series of different length are considered. Indeed, the ability of comparing multivariate time series of different length is one of the most remarkable feature of DTW, with respect to Euclidean distance or other standard distance measures for time series. In particular, in our case at each step the cost function is normalized by $(T + T')$.
- $d(., .)$ is the Euclidean distance for multivariate time series (Giorgino et al. 2009):

$$d(\mathbf{X}_i, \mathbf{X}_{i'}) = (\|\mathbf{x}_{it} - \mathbf{x}_{i't'}\|)^{\frac{1}{2}}. \tag{3}$$

Since there are several warping curves, the DTW distance is the one which correspond to the optimal warping curve. First, a $T \times T'$ "local cost" matrix which contains the distances between each pair of points is computed (Giorgino et al. 2009). Second, the DTW algorithm finds the path that minimizes the alignment between X_i and $X_{i'}$, starting at $d(1, 1)$ and finishing at $d(T, T')$, and aggregating the cost, i.e., the total distance. At each step, the algorithm finds the direction in which the distance increases the least, under the given constraints, yielding the optimal warping curve $\hat{\Phi}_l = (\hat{\varphi}_l, \hat{\psi}_l)_{l=1,\ldots,L}$ which minimizes the total dissimilarity between $\mathbf{X}_i$ and $\mathbf{X}_{i'}$ (see Fig. 1):

$$D(\mathbf{X}_i, \mathbf{X}_{i'}) = \sum_{l=1}^{L} d(\mathbf{x}_{i,\hat{\varphi}_l}, \mathbf{x}_{i',\hat{\psi}_l}) m_{l,\hat{\Phi}}. \tag{4}$$
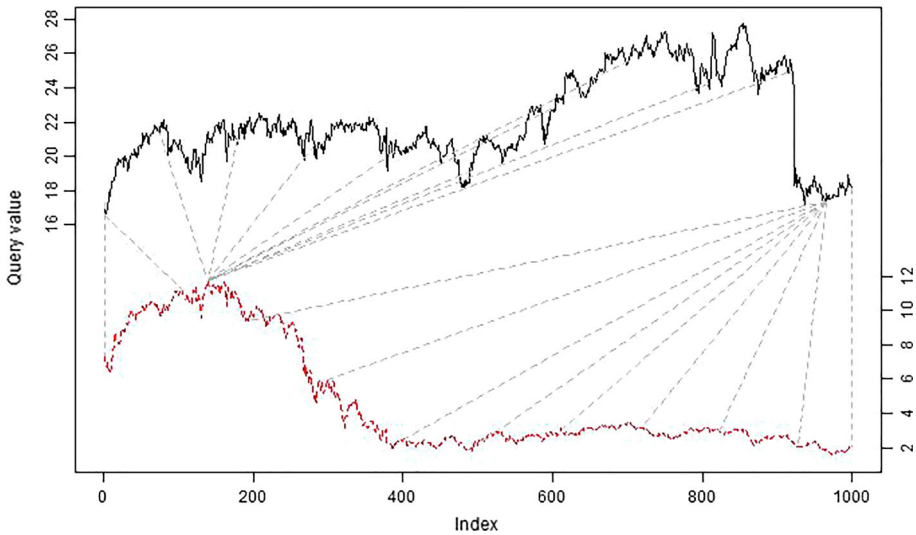
Fig. 1 Dynamic time warping distance

2.2 The clustering model

In this section, the DTW-based Trimmed Fuzzy C-Medoids Clustering model (DTW-Tr-FCMdC) is introduced, which is a robust version of the DTW-based Fuzzy C-Medoids Clustering model (DTW-FCMdC) introduced in D'Urso et al. (2018). The model achieves its robustness by adopting the "impartial trimming" procedure (García-Escudero and Gordaliza 1999; García-Escudero et al. 2010) to identify the time series more distant from the bulk of data. The procedure is said to be "impartial" because there is not a privileged direction in the trimming of the data, but the trimming is led by the data at hand. This approach is also suitable to detect both "outlying clusters" (outliers grouped in one small cluster) and "radial outliers" (isolated outliers) (García-Escudero et al. 2003).

Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_I)$ be a set of I multivariate time series, and $\widetilde{\mathbf{X}} = (\widetilde{\mathbf{X}}_1, \widetilde{\mathbf{X}}_2, \ldots, \widetilde{\mathbf{X}}_C)$ a subset of $\mathbf{X}$ with cardinality C. Given a trimming size α which ranges between 0 and 1, we solve the double minimization problem which yields to the DTW-Tr-FCMd clustering model:

$$
\begin{cases}
J(C, \alpha, u_{ic}, p) := \min_{Y} \min_{u_{ic}} : \sum_{i=1}^{H(\alpha)} \sum_{c=1}^{C} u_{ic}^p D^2(\mathbf{X}_i, \widetilde{\mathbf{X}}_c) \\
\text{s.t.: } \sum_{c=1}^{C} u_{ic} = 1, \quad u_{ic} \geq 0,
\end{cases}
\tag{5}
$$

where

- $D^2(\mathbf{X}_i, \widetilde{\mathbf{X}}_c)$ is the squared DTW-distance computed between the i-th multivariate time series and the c-th medoid;
- u_{ic} is the membership degree of the i-th time series to the c-the cluster;
- $p > 1$ is the fuzziness parameter—the greater the value of p the more fuzzy is the obtained partition;
- Y ranges on all the subsets of the I time series, containing $H(\alpha) = \lfloor I \cdot (1 - \alpha) \rfloor$ time series ($\lfloor . \rfloor$ is the integer part of a given value).

Using the above described trimming rule we allow for a proportion α of time series to be left unassigned (García-Escudero et al. 2010). Notice that (5) includes DTW-FCMdC as a limit case when $\alpha = 0$. Then, each non-trimmed time series is allocated into the cluster corresponding to its closest medoid time series, based on their pairwise DTW distance.

The local optimal solution is:

$$u_{ic} = \frac{1}{\sum_{c'=1}^{C} \left[\frac{D^2(\mathbf{X}_i, \widetilde{\mathbf{X}}_c)}{D^2(\mathbf{X}_i, \widetilde{\mathbf{X}}_{c'})} \right]^{\frac{1}{p-1}}}. \tag{6}$$

where i ranges on the subset of the non-trimmed time series.

The computational steps of the proposed clustering model are reported in Algorithm 1.

Algorithm 1 DTW-Tr-FCMdC algorithm

1: Fix C, the trimming level α, and the maximum number of iteration $max.iter$;
2: Set $iter = 0$;
3: Pick initial medoids: $\widetilde{\mathbf{X}} \equiv \{\widetilde{\mathbf{X}}_1, \ldots, \widetilde{\mathbf{X}}_c, \ldots, \widetilde{\mathbf{X}}_C\}$;
4: **repeat**
5:     Identify the subset Y made of the $H = \lfloor I \cdot (1 - \alpha) \rfloor$ time series closest to the medoids $\widetilde{\mathbf{X}}$;
6:     Store the current medoids: $\widetilde{\mathbf{X}}_{OLD} = \widetilde{\mathbf{X}}$;
7:     Compute u_{ic} using (6);
8:     Select the new medoids $\widetilde{\mathbf{X}} \equiv \{\widetilde{\mathbf{X}}_1, \ldots, \widetilde{\mathbf{X}}_c, \ldots, \widetilde{\mathbf{X}}_C\}$: $q = \underset{1 \leq i' \leq I}{\operatorname{argmin}} \sum_{i''=1}^{H} u_{i''c}^p \sum_{j=1}^{J} (\mathbf{X}_i - \mathbf{X}_{i'})^2$
9:     **return** $\widetilde{\mathbf{X}}_c = \widetilde{\mathbf{X}}_q$ $(c = 1, \ldots, C)$
10:     $iter \leftarrow iter + 1$;
11: **until** $\widetilde{\mathbf{X}}_{OLD} = \widetilde{\mathbf{X}}$ or $iter = max.iter$

Since Algorithm 1 is a recursive algorithm, it is not guaranteed that the global minimum is reached. Thus, more than one random start is suggested to obtain a stable solution.

The complexity of the DTW algorithm ($\mathcal{O}(TT')$) is counterbalanced by the PAM approach used for the proposed clustering algorithm. Indeed, the distance matrix is computed only at the start of the clustering process, and is invariant with the number of clusters.

Finally, it should be remarked that the fuzziness parameter p determines the shapes of the clusters and produces a fine tuning between the membership degrees close to 0 or 1 and those with intermediate values. The more m increases, the fuzzier the membership degrees are. Conversely, if m is close to 1, the resulting partition is hard (Wedel and Steenkamp 1989). In the fuzzy clustering literature, several heuristic procedures to select p have been proposed (see, among others, McBratney and Moore 1985; Okeke and Karnieli 2006; Wedel and Steenkamp 1989), but there is a lack of sounding theoretical basis to justify the selection of the fuzziness parameter. For this reason, as it is suggested in literature, we set $p = 1.5$ Kamdar and Joshi (2000).

2.3 Optimal trimming size and number of clusters

For a given value of the trimming size α, the number of clusters C can be determined by considering the usual internal validity indices for fuzzy clustering, like the Fuzzy Silhoutte (Campello and Hruschka 2006) and/or the Xie-Beni index (Xie and Beni 1991). However, the choice of the optimal value of C cannot be disentangled by the choice of the optimal value of α. Indeed, the robustness of the clustering process depends both on data and on the number of clusters (Hennig et al. 2008; García-Escudero et al. 2010).

García-Escudero et al. (2003) proposed an Exploratory Data Analysis (EDA)-based procedure which proved to be effective in the selection of α and C (D'Urso et al. 2017). The procedure can be summarized as follows. First, we retrieve the optimal values of the objective function (5), $\hat{J}(\alpha, C, u_{ic}, p)$ for different values of α and C. In particular, C ranges between 2 and $\lfloor \sqrt{I} \rfloor$ (Anderson et al. 2010), while $\alpha \in [0, 0.5]$, since with $\alpha > 0.5$ it is difficult to distinguish the real patterns in data from those generated by the outliers (Davé and Krishnapuram 1997).

1. For a given C we plot $\hat{J}(\alpha, C, u_{ic}, p)$ against increasing values of α.
2. If the rate of decrease of $\hat{J}(\alpha, C, u_{ic}, p)$ is smooth all over the range of α, then there are no outliers.
3. If there is an abrupt change of the rate of decrease in correspondence of a given value of α, this indicates the presence of $\alpha \cdot 100\%$ outliers.

To detect the correct number of clusters we also look at the trend of the numerical second derivative of the objective function, $\hat{J}''(\alpha, C, u_{ic}, p)$. If the number of clusters is improper, then there will be one or more "peaks" in the trend of the second derivative. If the number of clusters is correct we could observe two different behaviour of the second derivative curve, depending on whether there are outliers or not.

- If there are no outliers, the smallest C for which the trend is smooth for all value of α is the proper number of clusters.
- If there are outliers, we expect a fast decrease of the trend up to the value of α that identify the correct number of outliers and subsequently a smooth decay.

3 Clustering of FTSE MIB index stocks

In financial analysis the performance of a stock in terms of such variables as closing prices and trading volumes can be interpreted in term of the temporally relation with others stocks. Then there is scope for a clustering analysis of a set of stocks.

In this application we considered 40 of the most-traded stocks in the Italian Stock Exchange that make up the FTSE MIB (Financial Times Stock Exchange Milano Indice di Borsa) index. Stocks are referred to major Italian and foreign companies listed on the Italian Stock Exchange, and are selected to replicate the broad sector weights of the Italian stock market. The FTSE MIB is the primary benchmark Index for the Italian equity markets, and its constituent shares are selected from the from the universe of stocks trading on the Borsa Italiana (BIt) main equity market, to ensure the index best represents the Italian equity markets for investors—it represents approximately 80% of the domestic market capitalization. Each stock is analysed for size and liquidity, and the overall Index has appropriate sector representation. FTSE-MIB has been widely studied for the analysis of Italian Stock Market—among most recent contributions, see (Iglesias 2015; Degiannakis and Floros 2016; Liu and Tse 2017; Chang et al. 2018). Data are available on the finance section of the Yahoo website.[1] The available information for each stock are open and closing price, the highest and lowest price and the adjusted closing, other than the traded volume on a daily basis. Since there is a very high correlation (near 1) among the different definition of prices, here we considered only the adjusted closing price as a measure of the stock performance. Adjusted closing price corrects each stock's closing price from any distributions and corporate actions that occurred at any time before the next opening day. When examining historical returns, it is customary

---

[1] http://finance.yahoo.com.

to adopt the adjusted closing price (Rechenthin et al. 2013). However, for robustness' sake, a further analysis were conducted by considering all the available variables, which yielded to very similar results.[2] Furthermore, we also considered the traded volume, which is the total quantity of shares or contracts traded for any stock. The period of observation of the time series covers from January 2015 until November 2018.

The time series of daily adjusted closing prices and of traded volumes are shown in Figs. 2 and 3.

Our objective is to find a proper partition of the stocks considered. Since data are not available for all stocks for the whole period considered, as observed in Sect. 2.1, DTW distance is suitably apt to produce pairwise comparison among data, without loss of information. On the contrary, with the Euclidean distance one has to render all time series of the same length, hence losing information on a sizeable part of the period considered.

First, we performed the analysis with the non-robust clustering model DTW-FCMdC (D'Urso et al. 2018). The optimal value of C is detected by performing the classification for increasing values of $C = 2, \ldots, C_{max}$, where $C_{max} = \lfloor \sqrt{I} \rfloor$ (Anderson et al. 2010). Then, the value of C which minimizes the Xie-Beni index, and/or maximizes the Fuzzy Silhouette coefficient is selected as the optimal value of C. Since $\lfloor \sqrt{40} \rfloor = 6$, we found that the optimal value of C in [2, 6] is $C^* = 3$, with both criterion. The membership degrees are reported in a ternary plot (Fig. 4).

The three medoids are "DiaSorin", an Italian multinational biotechnology company, "Stmicroelectronics", a French-Italian multinational electronics and semiconductor manufacturer, and "Telecom Italia", the most important Italian telecommunications company.

However, from the interpretative standpoint, this partition is of little use since 35 out of 40 companies were classified in the second cluster. This is probably due to the presence of outliers patterns in the dataset, i.e., bivariate time series in which one or both components have anomalous dynamic patterns with respect to the bulk of data.

Therefore, we adopted the proposed clustering model DTW-Tr-FCMdC. As already noted in Sect. (2.3), the choice of the optimal number of clusters and of the trimming size, i.e., the number of outliers to be drop during the optimization process, must be done at the same time. Since $\lfloor \sqrt{40} \rfloor = 6$, for each value of $C \in [2, 6]$ and of $\alpha \in [0, 0.5]$ (step = 0.05) we recorded the value of the objective function (5) and of its numerical second derivative. Results are illustrated in Fig. 5.

Based on this analysis, the optimal value for C is 3 and for α is 0.075, which correspond to three outliers. Indeed, the trend of the numeric second derivative of the objective function (5) becomes smooth after $\alpha = 0.075$, indicating that dropping more multivariate time series does not cause the removal of a large cluster.

The membership degrees are reported in Table 1, and in the ternary plot represented in Fig. 6. In the table, the medoids are highlighted in bold, while outliers are in italic. In the last column of the table, the closest hard partition is reported, i.e, the cluster to which each unit is allocated based on the highest membership degree. With the partial exception of "Brembo" and "Fca", which are fuzzy allocated to the first and the third cluster, all stocks are assigned with a low degree of uncertainty.

The three medoids' time series are illustrated in Fig. 7. In Fig. 8 the average values by year of the two variables, for the overall sample and for each cluster—weighted with the membership degrees of each stock to each cluster—, are reported.

The medoid of the first cluster is "Buzzi Unicem", an Italian company which produces construction materials. Overall, the fourteen stocks in this cluster show closing prices slightly
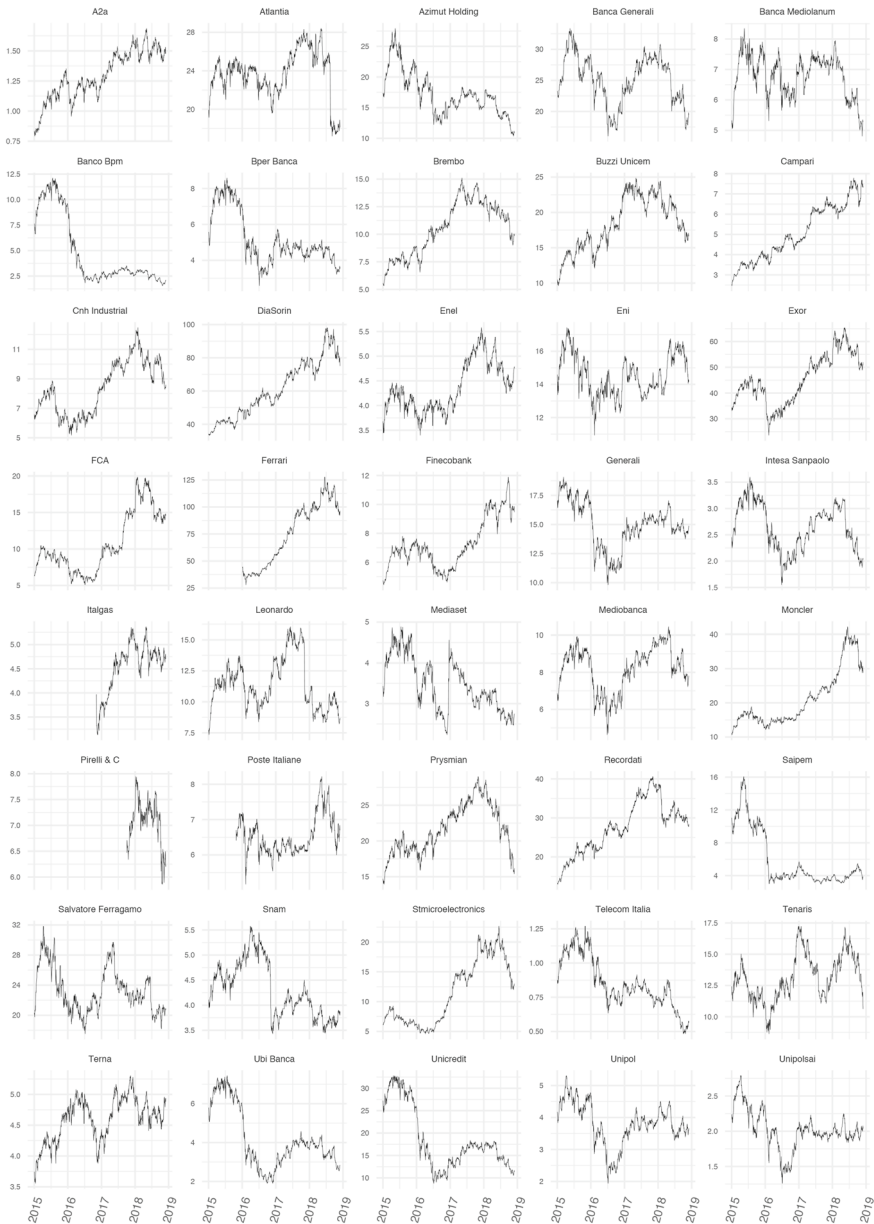

Fig. 2 Daily adjusted closing prices

higher than average, but with a trend similar to that observed for the whole sample. As for the volumes traded, the trend is declining, with lower values with respect to the rest of the stocks in FTSE MIB.

Stocks in the second cluster—whose medoid is "DiaSorin"— display an upward trend for closing price, especially in the last two years. The cluster is a "niche" cluster composed by three stocks— "DiaSorin", "Exor" and "Ferrari"— whose closing prices are the largest in the
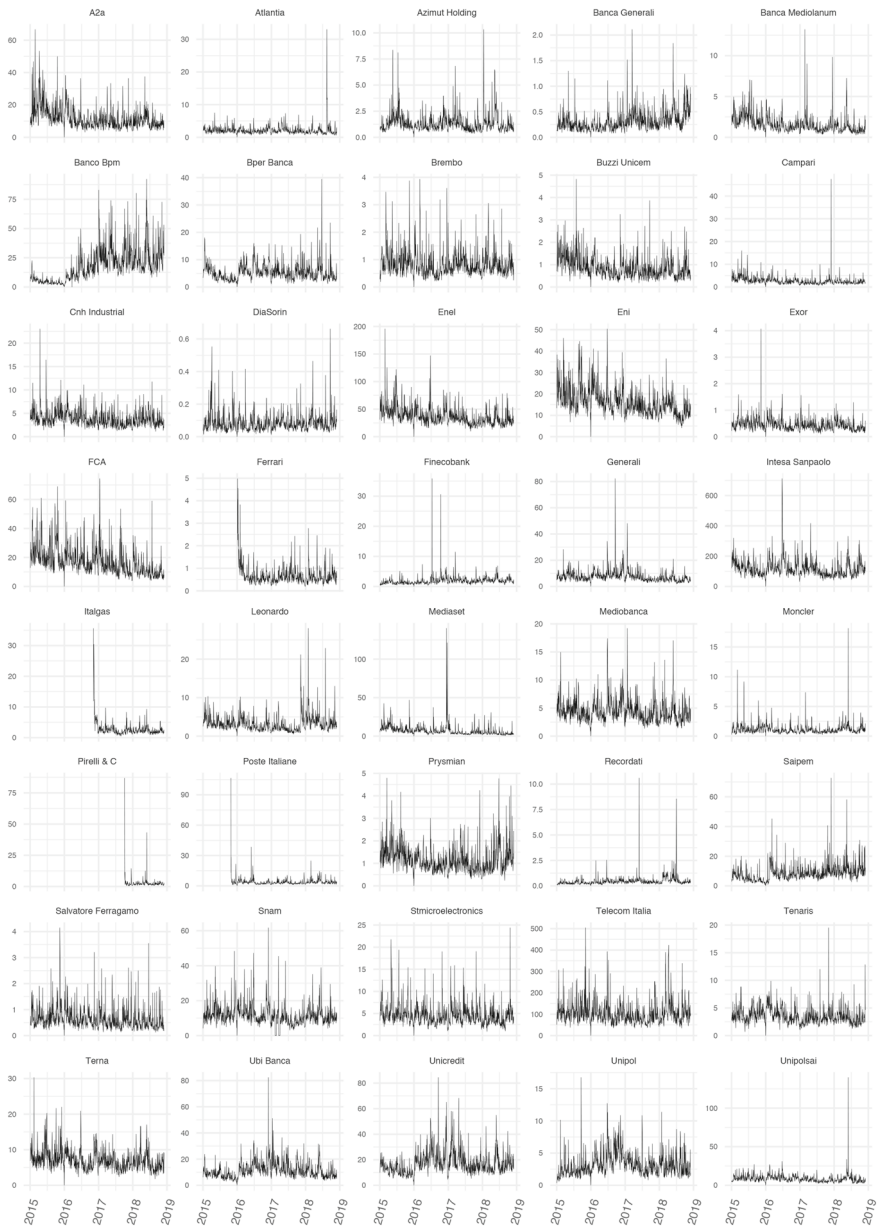
Fig. 3 Daily traded volumes (1,000,000 ops.)

sample considered (see also Fig. 2). On the contrary, volume traded are among the lowest. Notice that the second cluster of this robust partition coincides with the first cluster of the non-robust partition.

Finally, the medoid for the third cluster is "Terna", a transmission system operator (TSO). This cluster is characterized by twenty stocks whose closing prices are approximately stable during the period observed, with values that are lower than those observed in the whole
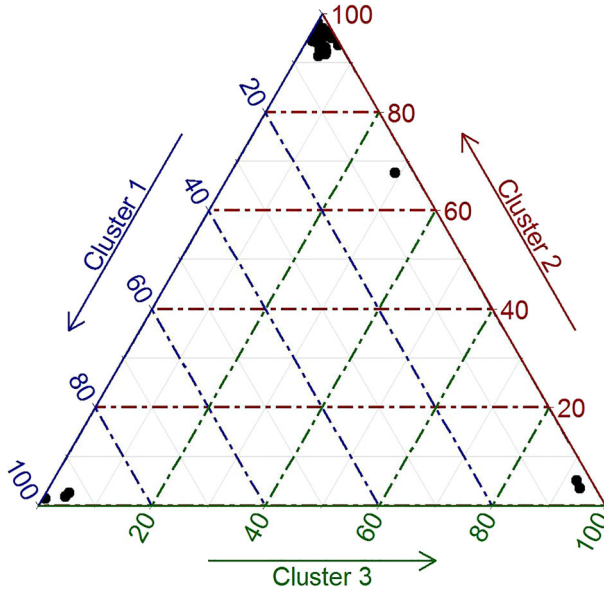
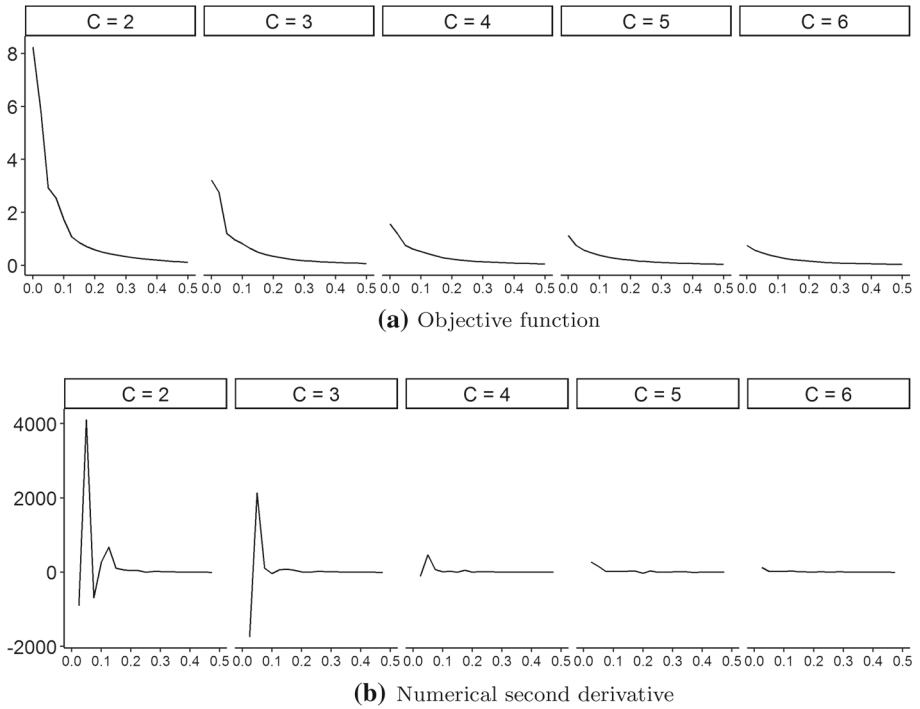**Fig. 4** Membership degrees—DTW-FCMdC clusteing model

(a) Objective function

(b) Numerical second derivative

Fig. 5 Objective function and numerical second derivative for different values of C

Table 1 Membership degrees

Stock	Cluster 1	Cluster 2	Cluster 3	Closest hard partition
A2a	0.0024	0.0000	0.9976	3
Atlantia	0.9999	0.0000	0.0001	1
Azimut Holding	0.9995	0.0000	0.0005	1
Banca Generali	0.9998	0.0000	0.0002	1
Banca Mediolanum	0.0019	0.0000	0.9981	3
Banco Bpm	0.0146	0.0001	0.9854	3
Bper Banca	0.0002	0.0000	0.9998	3
Brembo	0.6652	0.0001	0.3348	1
Buzzi Unicem	1.0000	0.0000	0.0000	1
Campari	0.0004	0.0000	0.9996	3
Cnh Industrial	0.0856	0.0000	0.9144	3
Diasorin	0.0000	1.0000	0.0000	2
Enel	0.0078	0.0002	0.9921	3
Eni	0.8211	0.0007	0.1781	1
Exor	0.0067	0.9926	0.0007	2
FCA	0.3934	0.0003	0.6063	3
Ferrari	0.0000	0.9999	0.0000	2
Finecobank	0.0362	0.0000	0.9638	3
Generali	0.9269	0.0001	0.0730	1
Intesa San Paolo	*Outlier*			
Italgas	0.0001	0.0000	0.9999	3
Leonardo	0.8848	0.0000	0.1152	1
Mediaset	0.0002	0.0000	0.9998	3
Mediobanca	0.0141	0.0000	0.9859	3
Moncler	0.9977	0.0004	0.0019	1
Pirelli & C	0.0066	0.0000	0.9934	3
Poste Italiane	0.0012	0.0000	0.9988	3
Prysmian	1.0000	0.0000	0.0000	1
Recordati	0.9970	0.0009	0.0021	1
Saipem	0.0053	0.0000	0.9947	3
Salvatore Ferragamo	0.9999	0.0000	0.0001	1
Snam	0.0001	0.0000	0.9999	3
Stmicroelectronics	0.9397	0.0000	0.0602	1
Telecom Italia	*Outlier*			
Tenaris	0.9844	0.0000	0.0156	1
Terna	0.0000	0.0000	1.0000	3
Ubi Banca	0.0004	0.0000	0.9996	3
Unicredit	*Outlier*			
Unipol	0.0001	0.0000	0.9999	3
Unipolsai	0.0010	0.0000	0.9990	3
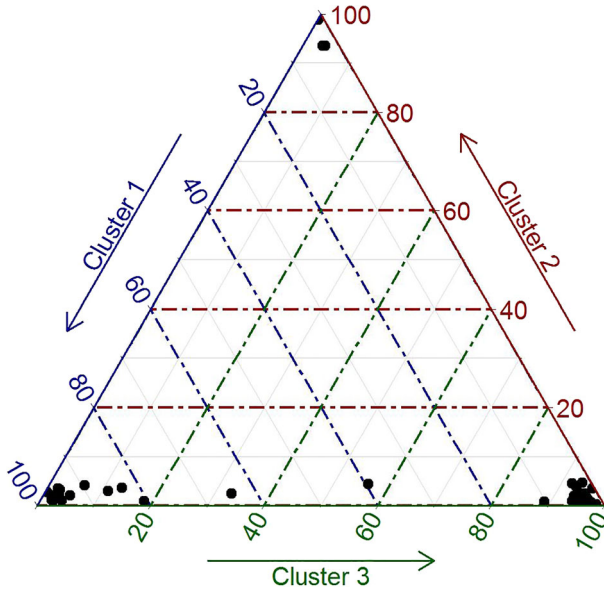
Fig. 6 Membership degrees—DTW-Tr-FCMdC

(a) Daily adjusted closing prices

(b) Daily traded volumes (1,000,000 ops.)

Fig. 7 Medoids

(a) Daily adjusted closing prices
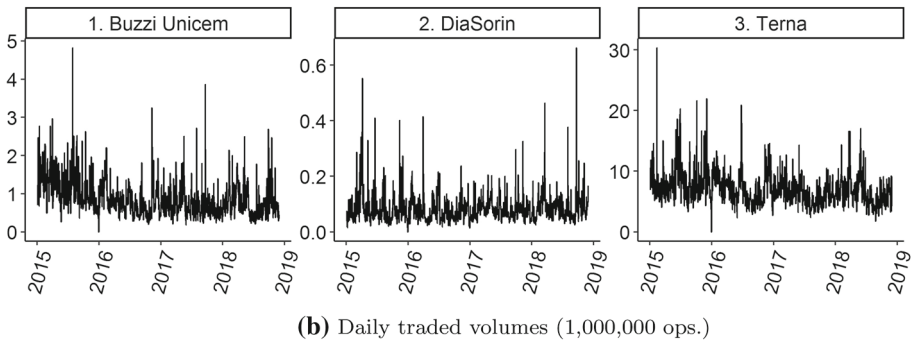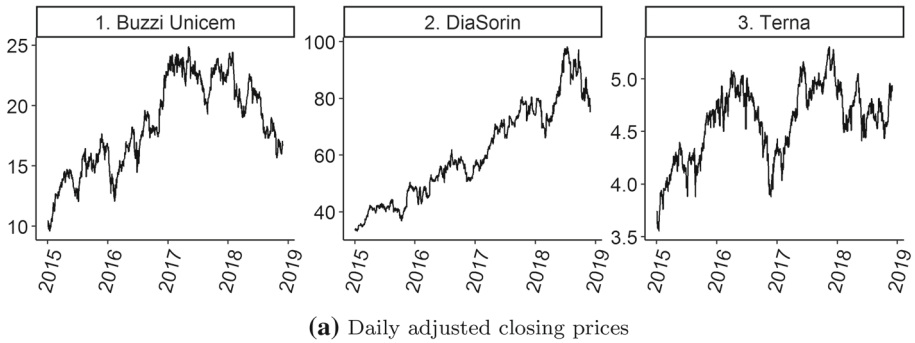

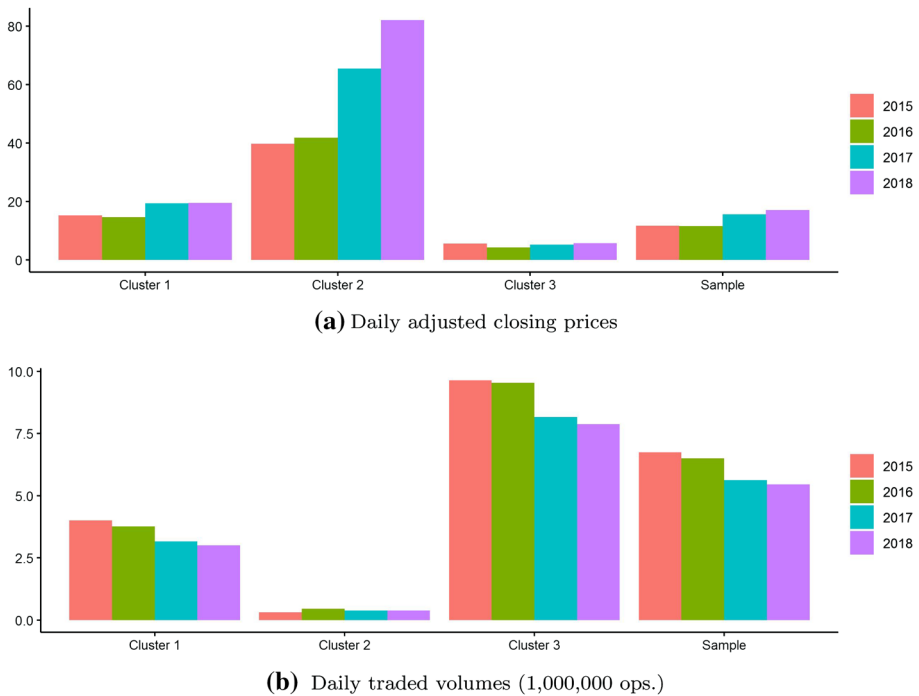**(b)** Daily traded volumes (1,000,000 ops.)

Fig. 8 Yearly average values

sample. As for the volumes traded, the overall trend is similar to that observed for the whole sample, but with larger values. Hence, not considering the outliers, stocks in this cluster are among the most trade during the period observed.

The three outliers are "Intesa San Paolo" and "Unicredit"—two of the major banking groups in Italy and in Europe— and "Telecom Italia" (Fig. 9). Actually, neither of the trends observed, both for closing prices and volumes traded, match with that registered for the three clusters. The volumes traded for the three outliers are the highest in the sample observed, while the closing prices are among the lowest for "Intesa San Paolo" and "Telecom Italia". Still, there are substantial differences in trends among these three company, which do not allow to consider them as an outlying cluster on its own. Finally, notice that "Telecom Italia" in the non-robust partition was identified as the medoid of the third cluster, thus reinforcing the suspect that the partition obtained with DTW-FCMdC were affected by the presence of outliers.

4 Final remarks

In this paper, by capturing the time-varying features of multivariate time series by means of the dynamic time warping, we introduced a robust fuzzy clustering method for classifying multivariate financial time series. The proposed clustering method neutralizes the disruptive effects of possible outlying objects by adopting a trimming procedure capable to identify the time series more distant from the bulk of time series dataset. In particular, by adopting the so-called "impartial trimming" procedure suggested by García-Escudero and Gordaliza
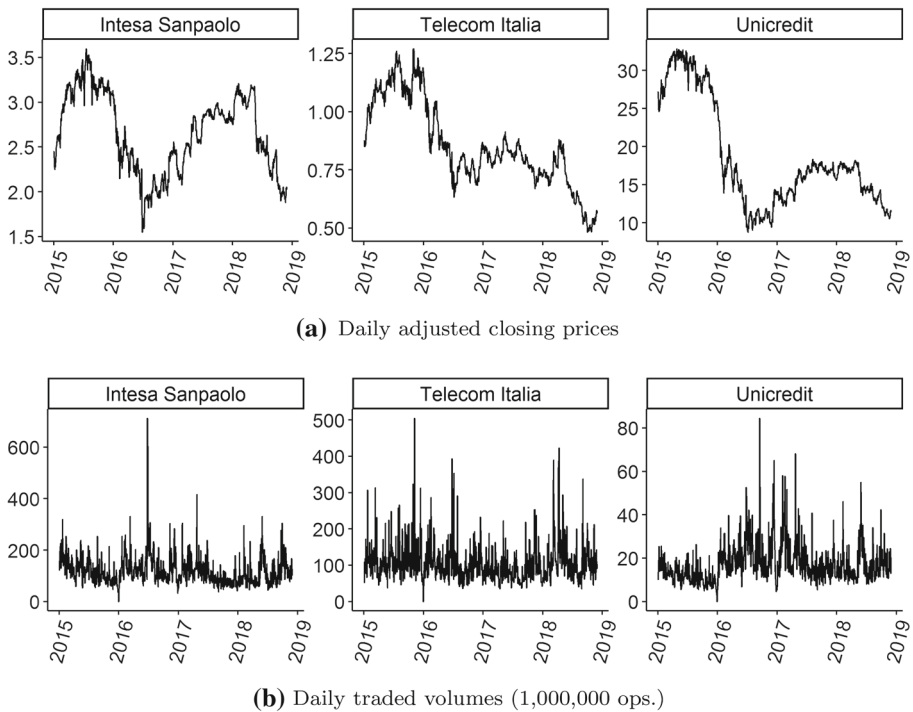
(a) Daily adjusted closing prices

(b) Daily traded volumes (1,000,000 ops.)

Fig. 9 Outliers

(1999) and García-Escudero et al. (2010), our fuzzy clustering method is capable to detect "outlying clusters of multivariate time series" grouped in one small cluster) and "radial outlier time series" (isolated outlying multivariate time series).

The usefulness and fruitfulness of the suggested clustering method has been shown in a finance framework by classifying the stocks composing the FTSE MIB index. In the analysis, we found evidence of the presence of two main clusters, one "niche" clusters and three radial outliers whose patterns do not match with any identified clusters.

In the future, we will explore other robust clustering approach for financial time series, e.g., the metric approach (see, for instance, D'Urso et al. 2018), and the noise cluster approach (see, for instance, Lafuente-Rego et al. 2019), by considering suitable features of the financial time series (feature-based approach) or the model generating the financial time series (model-based approach) (see Maharaj et al. 2019).

References

Anderson, D. T., Bezdek, J. C., Popescu, M., & Keller, J. M. (2010). Comparing fuzzy, probabilistic, and possibilistic partitions. *IEEE Transactions on Fuzzy Systems*, *18*(5), 906–918.

Ando, T., & Bai, J. (2017). Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures. *Journal of the American Statistical Association*, *112*(519), 1182–1198.

Aslan, S., Yozgatligil, C., & Iyigun, C. (2018). Temporal clustering of time series via threshold autoregressive models: Application to commodity prices. *Annals of Operations Research*, *260*(1–2), 51–77.

Basalto, N., Bellotti, R., De Carlo, F., Facchi, P., Pantaleo, E., & Pascazio, S. (2007). Hausdorff clustering of financial time series. *Physica A: Statistical Mechanics and its Applications*, *379*(2), 635–644.

Basalto, N., Bellotti, R., De Carlo, F., Facchi, P., Pantaleo, E., & Pascazio, S. (2008). Hausdorff clustering. *Physical Review E*, *78*(4), 046112.

Bastos, J. A., & Caiado, J. (2014). Clustering financial time series with variance ratio statistics. *Quantitative Finance*, *14*(12), 2121–2133.

Berndt, D.J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *Proceedings of the AAAI-94 workshop knowledge discovery in databases* (pp. 359–370). Seattle, WA.

Caiado, J., & Crato, N. (2007). A GARCH-based method for clustering of financial time series: International stock markets evidence. In C. Skiadas (Ed.), *Recent Advances in Stochastic Modeling and Data Analysis* (pp. 542–551). Singapore: World Scientific.

Campello, R. J. G. B., & Hruschka, E. R. (2006). A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems*, *157*, 2858–2875.

Chang, S.-L., Chien, C.-Y., Lee, H.-C., & Lin, C. (2018). Historical high and stock index returns: Application of the regression kink model. *Journal of International Financial Markets, Institutions and Money*, *52*, 48–63.

Davé, R. N., & Krishnapuram, R. (1997). Robust clustering methods: A unified view. *IEEE Transactions on Fuzzy Systems*, *5*(2), 270–293.

De Gregorio, A., & Iacus, S. M. (2010). Clustering of discretely observed diffusion processes. *Computational Statistics & Data Analysis*, *54*(2), 598–606.

De Luca, G., & Zuccolotto, P. (2011). A tail dependence-based dissimilarity measure for financial time series clustering. *Advances in Data Analysis and Classification*, *5*(4), 323–340.

De Luca, G., & Zuccolotto, P. (2017). A double clustering algorithm for financial time series based on extreme events. *Statistics & Risk Modeling*, *34*(1–2), 1–12.

Degiannakis, S., & Floros, C. (2016). Intra-day realized volatility for European and USA stock indices. *Global Finance Journal*, *29*, 24–41.

Dias, J. G., Vermunt, J. K., & Ramos, S. (2015). Clustering financial time series: New insights from an extended hidden Markov model. *European Journal of Operational Research*, *243*(3), 852–864.

Dose, C., & Cincotti, S. (2005). Clustering of financial time series with application to index and enhanced index tracking portfolio. *Physica A: Statistical Mechanics and its Applications*, *355*(1), 145–151.

Durante, F., Pappadà, R., & Torelli, N. (2014). Clustering of financial time series in risky scenarios. *Advances in Data Analysis and Classification*, *8*(4), 359–376.

D'Urso, P. (2000). Dissimilarity measures for time trajectories. *Statistical Methods & Applications*, *9*(1–3), 53–83.

D'Urso, P. (2004). Fuzzy C-Means clustering models for multivariate time-varying data: Different approaches. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *12*(03), 287–326.

D'Urso, P. (2005). Fuzzy clustering for data time arrays with inlier and outlier time trajectories. *IEEE Transactions on Fuzzy Systems*, *13*(5), 583–604.

D'Urso, P., Cappelli, C., Di Lallo, D., & Massari, R. (2013). Clustering of financial time series. *Physica A: Statistical Mechanics and its Applications*, *392*(9), 2114–2129.

D'Urso, P., De Giovanni, L., & Massari, R. (2016). GARCH-based robust clustering of time series. *Fuzzy Sets and Systems*, *305*, 1–28.

D'Urso, P., De Giovanni, L., & Massari, R. (2018). Robust fuzzy clustering of multivariate time trajectories. *International Journal of Approximate Reasoning*, *99*, 12–38.

D'Urso, P., Massari, R., Cappelli, C., & De Giovanni, L. (2017). Autoregressive metric-based trimmed fuzzy clustering with an application to PM_{10} time series. *Chemometrics and Intelligent Laboratory Systems*, *161*, 15–26.

García-Escudero, L. Á., & Gordaliza, A. (1999). Robustness properties of k means and trimmed k means. *Journal of the American Statistical Association*, *94*, 956–969.

García-Escudero, L. A., Gordaliza, A., & Matrán, C. (2003). Trimming tools in exploratory data analysis. *Journal of Computational and Graphical Statistics*, *12*, 434–449.

García-Escudero, L. A., Gordaliza, A., Matrán, C., & Mayo-Iscar, A. (2010). A review of robust clustering methods. *Advances in Data Analysis and Classification*, *4*, 89–109.

Giorgino, T., et al. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, *31*(7), 1–24.

Hennig, C., et al. (2008). Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis*, *99*(6), 1154–1176.

Iglesias, E. M. (2015). Value at Risk and expected shortfall of firms in the main European Union stock market indexes: A detailed analysis by economic sectors and geographical situation. *Economic Modelling*, *50*, 1–8.

Izakian, H., Pedrycz, W., & Jamal, I. (2015). Fuzzy clustering of time series data using dynamic time warping distance. *Engineering Applications of Artificial Intelligence*, *39*, 235–244.

Kamdar, T., & Joshi, A. (2000). On creating adaptive Web servers using Weblog Mining. Technical report TR-CS- 00-05, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County.

Lafuente-Rego, B., D'Urso, P., & Vilar, J. (in press 2019). Robust fuzzy clustering based on quantile autocovariances. Statistical Papers.

Lai, R. K., Fan, C.-Y., Huang, W.-H., & Chang, P.-C. (2009). Evolving and clustering fuzzy decision tree for financial time series data forecasting. *Expert Systems with Applications*, *36*(2), 3761–3773.

Liu, Q., & Tse, Y. (2017). Overnight returns of stock indexes: Evidence from ETFs and futures. *International Review of Economics & Finance*, *48*, 440–451.

Maharaj, E. A., D'Urso, P., & Caiado, J. (2019). *Time series clustering and classification*. Boca Raton: CRC Press.

Maharaj, E. A., D'Urso, P., & Galagedera, D. U. (2010). Wavelet-based fuzzy clustering of time series. *Journal of Classification*, *27*(2), 231–275.

McBratney, A., & Moore, A. (1985). Application of fuzzy sets to climatic classification. *Agricultural and Forest Meteorology*, *35*(1–4), 165–185.

Menardi, G., & Lisi, F. (2015). Double clustering for rating mutual funds. *Electronic Journal of Applied Statistical Analysis*, *8*(1), 44–56.

Nair, B. B., Kumar, P. S., Sakthivel, N., & Vipin, U. (2017). Clustering stock price time series data to generate stock trading recommendations: An empirical study. *Expert Systems with Applications*, *70*, 20–36.

Nakagawa, K., Imamura, M., & Yoshida, K. (2019). Stock price prediction using k-medoids clustering with indexing dynamic time warping. *Electronics and Communications in Japan*, *102*, 3–8.

Okeke, F., & Karnieli, A. (2006). Linear mixture model approach for selecting fuzzy exponent value in fuzzy c-means algorithm. *Ecological Informatics*, *1*(1), 117–124.

Pattarin, F., Paterlini, S., & Minerva, T. (2004). Clustering financial time series: An application to mutual funds style analysis. *Computational Statistics & Data Analysis*, *47*(2), 353–372.

Piccardi, C., Calatroni, L., & Bertoni, F. (2011). Clustering financial time series by network community analysis. *International Journal of Modern Physics C*, *22*(01), 35–50.

Rahmanishamsi, J., Dolati, A., & Aghabozorgi, M. R. (2018). A copula based ICA algorithm and its application to time series clustering. *Journal of Classification*, *35*(2), 230–249.

Ratanamahatana, C. A., & Keogh, E. (2004). Everything you know about dynamic time warping is wrong. In *Third workshop on mining temporal and sequential data*. Citeseer.

Rechenthin, M., Street, W. N., & Srinivasan, P. (2013). Stock chatter: Using stock sentiment to predict price direction. *Algorithmic Finance*, *2*(3–4), 169–196.

Velichko, V., & Zagoruyko, N. (1970). Automatic recognition of 200 words. *International Journal of Man-Machine Studies*, *2*(3), 223–234.

Vilar, J. A., Lafuente-Rego, B., & D'Urso, P. (2018). Quantile autocovariances: A powerful tool for hard and soft partitional clustering of time series. *Fuzzy Sets and Systems*, *340*, 38–72.

Vilar, J. M., Vilar, J. A., & Pértega, S. (2009). Classifying time series data: A nonparametric approach. *Journal of classification*, *26*(1), 3–28.

Wedel, M., & Steenkamp, J. (1989). A fuzzy clusterwise regression approach to benefit segmentation. *International Journal of Research in Marketing*, *6*(4), 241–258.

Xie, X. L., & Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *13*(8), 841–847.

Yang, C., Jiang, W., Wu, J., Liu, X., & Li, Z. (2018). Clustering of financial instruments using jump tail dependence coefficient. *Statistical Methods & Applications*, *27*(3), 491–513.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.