International Journal of Approximate Reasoning 187 (2025) 109552

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

FCPCA: Fuzzy clustering of high-dimensional time series based on common principal component analysis

Ziling Ma *, Ángel López-Oriona, Hernando Ombao, Ying Sun

King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division. Thuwal 23955-6900, Saudi Arabia

ARTICLE INFO

Keywords:
Fuzzy clustering
Multivariate time series
Principal component analysis
Reconstruction criteria
Electroencephalogram signals

ABSTRACT

Clustering multivariate time series data is a crucial task in many domains, as it enables the identification of meaningful patterns and groups in time-evolving data. Traditional approaches, such as crisp clustering, rely on the assumption that clusters are sufficiently separated with little overlap. However, real-world data often defy this assumption, showing overlapping distributions or overlapping clouds of points and blurred boundaries between clusters. Fuzzy clustering offers a compelling alternative by allowing partial membership in multiple clusters, making it well-suited for these ambiguous scenarios. Despite its advantages, current fuzzy clustering methods primarily focus on univariate time series, and for multivariate cases, even datasets of moderate dimensionality become computationally prohibitive. This challenge is further exacerbated when dealing with time series of varying lengths, leaving a clear gap in addressing the complexities of modern datasets. This work introduces a novel fuzzy clustering approach based on common principal component analysis to address the aforementioned shortcomings. Our method has the advantage of efficiently handling high-dimensional multivariate time series by reducing dimensionality while preserving critical temporal features. Extensive numerical results show that our proposed clustering method outperforms several existing approaches in the literature. An interesting application involving brain signals from different drivers recorded from a simulated driving experiment illustrates the potential of the approach.

1. Introduction

Clustering multivariate time series (MTS) is crucial across domains such as finance [50], bioscience [3], and environmental sciences [79], where uncovering latent patterns yields actionable insights. Yet the complexity of MTS poses serious methodological hurdles [26,85].

High-dimensional signals, such as electroencephalogram (EEG) recordings, are a prototypical example: they contain many inter-dependent channels [51,5] and often vary in length [60]. Real-world series exhibit intricate cross-variable correlations [31] and may differ in duration or sampling frequency, exacerbating the limitations of conventional clustering algorithms [35,15,7].

Furthermore, MTS rarely definitively jumps from one regime to another. Instead, they drift through intermediate or mixed states. In EEG, for example, transitions between alertness and drowsiness, or between successive sleep stages, unfold gradually. Thus, a short

* Corresponding author.
E-mail addresses: ziling.ma@kaust.edu.sa (Z. Ma), angel.lopezoriona@kaust.edu.sa (Á. López-Oriona), hernando.ombao@kaust.edu.sa (H. Ombao), ying.sun@kaust.edu.sa (Y. Sun).

https://doi.org/10.1016/j.ijar.2025.109552
Received 11 April 2025; Received in revised form 9 August 2025; Accepted 14 August 2025

Available online 29 August 2025
0888-613X/© 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

segment can legitimately reflect more than one brain state [81,59,4]. Hard (crisp) clustering forces such ambiguous observations into a single class, erasing this continuum [64,48]. Fuzzy clustering, on the contrary, allows data points to belong to multiple clusters with varying degrees of membership and effectively represents the inherent uncertainty and ambiguity in classifying data into different clusters [56,2]. As noted in [28], "vague brain-state transitions are naturally treated by assigning each temporal pattern to one or more fuzzy clusters". This underscores the suitability of fuzzy methods for domains like neuroscience, medicine, and finance, where gradual changes occur.

Unlike univariate time series (UTS), MTS data encompass multiple interdependent variables, often exhibiting both temporal dependencies and complex cross-variable relationships, making the clustering task very challenging [63]. Additionally, real-world MTS datasets frequently feature variable sequence lengths and high dimensionality, further hindering the direct application of traditional clustering algorithms [24]. These challenges are exacerbated in fuzzy clustering, where the aim is not only to group similar MTS but also to capture partial memberships, producing an additional layer of computational and methodological complexity [64].

Although MTS has been studied in many domains and attracted growing interest, most fuzzy clustering methods still focus on univariate time series (UTS). In the UTS context, [53] proposed a distance-based method for short time series. However, their approach is only applicable to small-sample time series and often yields cluster solutions that are difficult to interpret meaningfully when applied to longer sequences. [22] introduced an autocorrelation-based fuzzy C-means clustering method. [46] developed a cepstral-based fuzzy C-means algorithm and demonstrated its advantages over the wavelet-based fuzzy method proposed by [47] through simulations. Later, [20] proposed a robust fuzzy C-medoids algorithm which represents each univariate series by a low-dimensional vector of autoregressive (AR) coefficients, although it relies on a correctly specified AR model. [34] developed a shape-based fuzzy clustering method by incorporating dynamic time warping (DTW) into the standard fuzzy C-means framework. Nevertheless, this approach is computationally expensive and scales poorly to datasets with many time series or with long time series. More recently, [71] proposed a quantile autocovariance-based fuzzy method, and later introduced robust versions [38].

With respect to fuzzy clustering of MTS, the existing literature remains decidedly limited. [17] compared several dissimilarity metrics for multivariate temporal trajectories and grouped them into three distinct classes based on their defining characteristics. Subsequently, [18] introduced a wavelet-based fuzzy clustering framework for MTS. [40] proposed a feature-weighted method that combines multivariate DTW and a shape-based distance, with medoids selected via density peaks; however, the method suffers from high computational complexity. Two robust extensions of DTW-based fuzzy clustering have been proposed: one incorporates impartial trimming [19], while the other employs an exponential DTW metric with entropy regularization for multivariate series [21]. [44] transformed each MTS into a vector of quantile cross-spectral density estimates, applied PCA to reduce the feature space, and then performed fuzzy clustering. They later introduced three robust variants of this approach [42], and extended one of them into a spatial weighted clustering framework [43].

FCPCA can be regarded as a membership-informed subspace method tailored to MTS: fuzzy memberships weight the estimation of cluster-specific common principal components (CPCs), and these CPCs define the subspaces themselves. Subspace clustering, long studied in areas such as face recognition [68] and motion segmentation [78], is commonly grouped into four families [1]: spectral [58,87], iterative [83], algebraic [70], and statistical [62,11] approaches. In contrast to these traditional feature-space techniques, FCPCA works directly in lag-covariance space, enabling it to capture the richer temporal dynamics inherent to MTS. Despite the maturity of subspace clustering in general, methods explicitly designed for MTS remain scarce. The soft-subspace ensemble of [29], which assigns variable-weight vectors in the raw feature domain and ignores cross-lag dependence, and the recent fuzzy-granule soft-subspace work [72] also weights variables rather than modeling lag-covariance structure. This persistent gap motivates FCPCA: it jointly learns cluster-specific lag-covariance subspaces and fuzzy memberships, capturing the temporal dynamics that other subspace methods overlook.

Collectively, most existing MTS fuzzy clustering algorithms stumble on at least one key challenge: they are either (i) computationally heavy or infeasible once the number of channels exceeds a modest level (20–30), (ii) based on the assumption of equal sequence lengths, (iii) unable to capture cross-variable lagged correlations, or (iv) reliant on fixed, predefined values for the fuzziness parameter and the number of clusters. To address these limitations, we introduce a novel fuzzy clustering method tailored for moderately high-dimensional MTS that accommodates both fixed and varying sequence lengths. The proposed method, fuzzy clustering based on common principal component analysis (FCPCA), represents a new fuzzy clustering framework for multivariate time series. It extends the classical common principal component analysis (CPCA) approach by incorporating fuzzy membership degrees into the projection construction process. The method begins by estimating cross-covariance matrices at different lags and constructing block matrices that capture both zero-lag and lagged dependencies. These matrices are then aggregated into a single weighted covariance matrix (weighted according to current membership degrees), on which singular value decomposition (SVD) is performed to extract principal components. A subset of these components, chosen to capture a high proportion of the total variance, is then used to project the original MTS onto a lower-dimensional common subspace. Cluster membership degrees are iteratively updated by minimizing the total weighted reconstruction error, with a fuzziness parameter controlling the balance between crisp and soft assignments. Although methodologically sophisticated, the algorithm is computationally efficient in practice.

The methodology is validated on both simulated and real-world datasets. EEG data provide an ideal proving ground: they embody all the aforementioned challenges, high dimensionality, potentially unequal lengths, rich lagged cross-channel interactions, and continuous transitions between brain states (e.g., alert and drowsy). The clinical relevance of capturing subtle, mixed brain states further amplifies the value of fuzzy assignments. Accordingly, the main application focuses on an EEG driving dataset. In this case study, 3-second EEG segments from multiple channels are clustered to distinguish between alert and drowsy states, revealing clear groupings as well as transitional states that traditional hard clustering methods would likely overlook. Additional experiments on well-known datasets, such as Japanese Vowels, Basic Motions, and other benchmark collections, demonstrate that the proposed fuzzy

clustering approach effectively captures nuanced and overlapping dynamics. This enables the algorithm to deliver richer information and supports its potential integration into systems such as adaptive driver-monitoring platforms.

The main contributions of this work can be summarized as follows: (1) Fuzzy clustering via lagged cross-covariances. We extend CPCA-based clustering into a fuzzy framework, explicitly incorporating lagged cross-covariance blocks to better capture the temporal dependencies intrinsic to MTS. (2) Data-driven tuning of the fuzziness parameter and number of clusters. We introduce an FCPCA-based validity index that simultaneously selects the fuzziness parameter and the number of clusters based on the data. (3) Membership-weighted projection axes. Projection directions are obtained from membership-weighted covariances at each iteration, yielding subspaces that are more representative of the underlying group structure and naturally down-weighting outlying observations. (4) Improved performance and interpretability. FCPCA provides richer clustering insights than traditional hard clustering and often outperforms, or achieves results comparable to, existing fuzzy clustering benchmarks in terms of accuracy.

The remainder of the paper is organized as follows. Section 2 reviews related work. Section 3 presents the FCPCA methodology, including the dimensionality reduction strategy, dissimilarity measure, and fuzzy clustering framework. Section 4 reports numerical results. In Section 5, the proposed approach is applied to the EEG driver drowsiness dataset. Finally, Section 6 concludes with a discussion of the findings and potential directions for future research.

2. Hard clustering of MTS based on CPCA

This section briefly reviews a clustering method for MTS based on CPCA. Although the original method, referred to as Mc2PCA (MTS clustering based on CPCA), was proposed by [39], we modified it by incorporating the ROBCPCA algorithm (a robust MTS clustering method based on CPCA) introduced by [45]. This robust extension accounts for both temporal (serial) dependencies and cross-variable relations within MTS. These modifications allow the hard clustering method to better capture temporal patterns in multivariate time series. As suggested by the authors, we consider lags up to 2.

Consider a dataset $\boldsymbol{X} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N\}$ of N MTS, where $\boldsymbol{X}_i \in \mathbb{R}^{T_i \times p}$ represents the ith time series with T_i time points and p observed dimensions. Each $\boldsymbol{X}_i$ is assumed to be a realization of a p-dimensional second-order stationary process $\{\boldsymbol{\mathcal{X}}_t^i, t \in \mathbb{Z}\} = \{(\mathcal{X}_t^{i,1}, \ldots, \mathcal{X}_t^{i,p})^\top, t \in \mathbb{Z}\}$, where $\mathcal{X}_t^{i,j}$ is the value of the jth dimension at time t for the ith series. Second-order stationarity indicates that the mean vector $\boldsymbol{\mu}_i = E[\boldsymbol{\mathcal{X}}_t^i]$ is constant over time, and the cross-covariance matrix $\Gamma_i(l) = E[(\boldsymbol{\mathcal{X}}_t^i - \boldsymbol{\mu}_i)(\boldsymbol{\mathcal{X}}_{t-l}^i - \boldsymbol{\mu}_i)^\top]$ depends only on the lag $l \in \mathbb{Z}$ and not on t. The main task is to divide the N MTS into S non-overlapping groups.

For a given lag l, we estimate $\Gamma_i(l)$ using the natural sample estimate:

$$\hat{\Gamma}_i(l) = \frac{1}{T_i - l} \sum_{t=l+1}^{T_i} (\boldsymbol{X}_{i,t} - \overline{\boldsymbol{X}}_i)(\boldsymbol{X}_{i,t-l} - \overline{\boldsymbol{X}}_i)^\top,$$

where $\overline{\boldsymbol{X}}_i$ is the column-wise sample mean of $\boldsymbol{X}_i$.

For each lag $l \in \{1, \ldots, L\}$, a block matrix is constructed by combining lag-0 and lag-l cross-covariance estimates:

$$\hat{\boldsymbol{\Gamma}}_i(l) = \begin{pmatrix} \hat{\Gamma}_i(0) & \hat{\Gamma}_i(l) \\ \hat{\Gamma}_i(l)^\top & \hat{\Gamma}_i(0) \end{pmatrix}, \tag{1}$$

where $\hat{\boldsymbol{\Gamma}}_i(l) \in \mathbb{R}^{2p \times 2p}$. The collection of all block matrices across N samples for lag l is denoted as $\hat{\boldsymbol{\Gamma}}(l) = \{\hat{\boldsymbol{\Gamma}}_1(l), \ldots, \hat{\boldsymbol{\Gamma}}_N(l)\}$.

Next, the common matrix $\overline{\boldsymbol{\Sigma}}(l)$ is obtained by averaging over all N samples:

$$\overline{\boldsymbol{\Sigma}}(l) = \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\Gamma}}_i(l). \tag{2}$$

Applying SVD to $\overline{\boldsymbol{\Sigma}}(l)$ yields eigenvalues $\boldsymbol{\Lambda}^l = (\lambda_1^l, \ldots, \lambda_{2p}^l)$ and eigenvectors $\mathbf{V}^l = \{\mathbf{V}_1^l, \ldots, \mathbf{V}_{2p}^l\}$. A common projection space $\mathbf{C}(l)$ is constructed by selecting the first $k(l)$ components of $\mathbf{V}^l$, where $k(l)$ is chosen to capture at least 95% of the total variance:

$$k(l) = \min\left\{ r \in \{1, \ldots, 2p\} : \frac{\sum_{j=1}^{r} \lambda_j^l}{\sum_{j=1}^{2p} \lambda_j^l} \geq 0.95 \right\}. \tag{3}$$

The clustering procedure divides $\boldsymbol{X}$ into S groups. For each cluster, the common spaces $\boldsymbol{C}(l) = \{\boldsymbol{C}_1(l), \ldots, \boldsymbol{C}_S(l)\}$ are constructed. To avoid of having non-conformable arguments, $\hat{\mathbf{X}}_i(l)$ is introduced as

$$\hat{\mathbf{X}}_i(l) = (\mathbf{X}_{i,t-l}^*, \mathbf{X}_{i,t}^*), \tag{4}$$

where

$$\mathbf{X}_{i,t-l}^* = (\mathbf{X}_{i,1}, \ldots, \mathbf{X}_{i,T_i-l})^\top,$$

and

$$\mathbf{X}_{i,t}^* = (\mathbf{X}_{i,1+l}, \ldots, \mathbf{X}_{i,T_i})^\top.$$

Later, each $\hat{\mathbf{X}}_i$ is projected onto the common spaces as

$$\boldsymbol{Y}_i^s(l) = \hat{\boldsymbol{X}}_i(l)\boldsymbol{C}_s(l)\boldsymbol{C}_s(l)^\top, \quad \text{where } s \in \{1, \cdots, S\}.$$

After projection, the reconstruction error $E_{is}(l)$ of projecting i-th MTS object $\hat{\mathbf{X}}_i(l)$ onto the s-th cluster at lag l is computed as:

$$E_{is}(l) = \|\hat{\boldsymbol{X}}_i(l) - \boldsymbol{Y}_i^s(l)\|_2. \tag{5}$$

The object $\boldsymbol{X}_i$ will be grouped into the cluster where its total reconstruction error across all lags

$$E_i = \min_{s\in\{1,\ldots,S\}} \sum_{l=1}^{L} E_{is}(l),$$

is the minimum.
As a result, the overall error is:

$$E = \sum_{i=1}^{N} E_i. \tag{6}$$

The procedure iteratively updates the cluster partition and projection spaces until the convergence of Equation (6) or the maximum number of iterations.

Although CPCA-based hard clustering approaches provide valuable insight into the structure of multivariate time series data, they suffer from inherent limitations. They assign each data point to a single cluster, which can be overly restrictive, especially when the boundaries of the cluster are ambiguous or overlapped. Consequently, the fuzzy version is preferred because it allows data points to belong to multiple clusters with varying degrees of membership, providing a more flexible approach to modeling complex, overlapping structures in MTS data.

3. FCPCA: fuzzy clustering of MTS based on CPCA

In this section, we extend the previously described hard clustering algorithm to a fuzzy clustering framework. We call the corresponding procedure FCPCA. Clustering is performed based on the reconstruction error of each MTS in each common space, which differs from distance-based clustering in classical fuzzy C-means [8] or fuzzy C-medoids [37].

3.1. The FCPCA clustering problem

In this framework, we propose to perform fuzzy clustering on $\boldsymbol{X}$ by using the FCPCA algorithm. The main goal is to find the $N \times S$ matrix of fuzzy membership, $\mathbf{U} = (u_{is})$, $i = 1, \ldots, N$, $s = 1, \ldots, S$, and sets of common projection axes, $\boldsymbol{C}(l) = \{\mathbf{C}_1(l), \ldots, \mathbf{C}_S(l)\}$, to minimize the total weighted reconstruction error. Formally, the clustering problem is stated as

$$\begin{cases} \min\limits_{\mathbf{U},\boldsymbol{C}(l)} \quad \sum\limits_{i=1}^{N}\sum\limits_{s=1}^{S} u_{is}^m \sum\limits_{l=1}^{L} \|\hat{\boldsymbol{X}}_i(l) - \hat{\boldsymbol{X}}_i(l)\mathbf{C}_s(l)\mathbf{C}_s(l)^\top\|^2 \\ \text{subject to} \quad \sum\limits_{s=1}^{S} u_{is} = 1 \quad \text{and} \quad u_{is} \geq 0 \end{cases}, \tag{7}$$

where $u_{is} \in [0, 1]$ represents the membership degree of the ith MTS in the sth cluster, and $m > 1$ is a real number, usually referred to as the fuzziness parameter, regulating the fuzziness of the partition. For $m = 1$, the crisp version of the algorithm is obtained, so the solution takes the form $u_{is} = 1$ if the ith series pertains to cluster c and $u_{is} = 0$ otherwise. As the value of m increases, the boundaries between clusters get softer and the resulting partition is fuzzier.

To integrate fuzzy membership degrees into the projection space, we define a common weighted covariance matrix for cluster s:

$$\tilde{\boldsymbol{\Sigma}}_s(l) = \frac{\sum_{i=1}^{N} u_{is}^m \hat{\boldsymbol{\Gamma}}_i(l)}{\sum_{i=1}^{N} u_{is}^m}, \quad s \in \{1, \cdots, S\}, \tag{8}$$

where $\hat{\boldsymbol{\Gamma}}_i(l)$ is the estimated block covariance matrix for the ith MTS object at lag l given by Equation (1). The set of projection axes $\boldsymbol{C}(l)$ is then obtained by selecting principal components from each $\tilde{\boldsymbol{\Sigma}}_s(l)$ (e.g., retaining enough components to capture 95% of the variance or a fixed number of components).

To solve the minimization problem above, an iterative algorithm that alternately optimizes the membership degrees and the projection axes is considered. For a given membership matrix $\mathbf{U}$, the algorithm first computes all $\tilde{\boldsymbol{\Sigma}}_s(l)$ and their corresponding $\boldsymbol{C}(l)$. It then evaluates the reconstruction error for each MTS $\boldsymbol{X}_i$ under the projection space of each cluster. The membership degrees u_{is} are updated by comparing reconstruction errors across clusters as follows:

$$u_{is} = \left[\sum_{s^*=1}^{S} \left(\frac{\sum_{l=1}^{L} \| \hat{\boldsymbol{X}}_i(l) - \hat{\boldsymbol{X}}_i(l) \mathbf{C}_s(l) \mathbf{C}_s(l)^\top \|^2}{\sum_{l=1}^{L} \| \hat{\boldsymbol{X}}_i(l) - \hat{\boldsymbol{X}}_i(l) \mathbf{C}_{s^*}(l) \mathbf{C}_{s^*}(l)^\top \|^2} \right)^{\frac{1}{m-1}} \right]^{-1}. \tag{9}$$

The total weighted reconstruction error is recalculated after updating $\mathbf{U}$. The number of principal components is set during the first iteration and remains fixed thereafter. The procedure continues until convergence (i.e., until the total reconstruction error ceases to decrease) or until a maximum number of iterations is reached.

3.2. The cluster validity indices

Choosing the fuzziness parameter m and the number of clusters S is crucial: an unsuitable pair (S, m) can either under–partition the data or over–smooth the memberships, dramatically degrading the quality of the final solution. Following the classical idea of the cluster validity indices (CVI) [84,76,36,24], we extend two widely–used fuzzy indices so that they exploit quantities inherent to the FCPCA algorithm. As a result, each index can be computed directly from a single run of the algorithm, making a systematic grid search over (S, m) computationally inexpensive, as the indices are obtained at no additional computational cost.

(i) **FCPCA-based Xie–Beni (XB) index.**

We adapt the compactness/separation idea of [77] as follows:

$$\mathrm{XB}_{S,m} = \frac{\sum_{i=1}^{N} \sum_{s=1}^{S} u_{is}^m \sum_{l=1}^{L} \left\| \hat{\boldsymbol{X}}_i(l) - \hat{\boldsymbol{X}}_i(l)\, \mathbf{C}_s(l)\, \mathbf{C}_s(l)^\top \right\|^2}{N \min\limits_{r \neq t} \sum_{l=1}^{L} \left\| P_r(l) - P_t(l) \right\|^2}, \tag{10}$$

where

$$P_s(l) = \mathbf{C}_s(l)\, \mathbf{C}_s(l)^\top, \quad s \in \{1, \ldots, S\}.$$

Each $P_s(l)$ is a $2p \times 2p$ matrix. The set

$$\mathcal{P} = \left\{ P_s(l) \right\}, \quad s \in \{1, \ldots, S\},\ l \in \{1, \ldots, L\},$$

is stored after convergence and includes the prototypes of the final solution. Note that each cluster is represented by a collection of L prototypes.

Notice that the numerator is exactly the FCPCA objective, while the denominator includes the smallest Frobenius distance between any two cluster projection subspaces, thus measuring separation. Smaller values of $\mathrm{XB}_{S,m}$ are preferred.

(ii) **FCPCA-based fuzzy silhouette index.**

Here, we consider an FCPCA-based variant of the fuzzy silhouette (FS) proposed by [10]. We denote

$$z_i = \arg\min_{s} \sum_{l=1}^{L} E_{is}(l)$$

as the index of the cluster for which the projection of the ith time series gives the lowest total reconstruction error,

$$a_i = \sum_{l=1}^{L} E_{i\, z_i}(l)$$

as the intra-cluster reconstruction error of the ith time series, i.e., how well it fits its assigned cluster, and

$$b_i = \min_{s \neq z_i} \sum_{l=1}^{L} E_{is}(l),$$

the smallest reconstruction error for the ith time series among all other clusters excluding z_i.

Let $u_{i(1)}$ and $u_{i(2)}$ denote the largest and second-largest membership degrees of the ith time series across all clusters. We then define the FS as

$$\mathrm{FS}_{S,m} = \frac{\sum_{i=1}^{N} \left(u_{i(1)} - u_{i(2)} \right) \frac{b_i - a_i}{\max\{a_i, b_i\}}}{\sum_{i=1}^{N} \left(u_{i(1)} - u_{i(2)} \right)}, \tag{11}$$

Here, the numerator computes a weighted average of individual silhouette scores, where each score $(b_i - a_i)/\max\{a_i, b_i\}$ quantifies how the ith time series is well-separated and well-clustered, and the weight $u_{i(1)} - u_{i(2)} \in [0, 1]$ reflects the confidence of cluster

assignment. Objects with ambiguous memberships (i.e., $u_{i(1)} \approx u_{i(2)}$) contribute less to the final score. A value of $\mathrm{FS}_{S,m}$ close to 1 indicates strong clustering structure, while values near 0 suggest poor separation or misassignments [69].

In practice, we compute both indices over a modest grid and select the parameter (either S, m, or both) that simultaneously minimizes $\mathrm{XB}_{S,m}$ and maximizes $\mathrm{FS}_{S,m}$. When the two criteria disagree, we prioritize the lower XB, as it is directly aligned with the objective optimized by FCPCA.

3.3. The fuzziness parameter m

The FCPCA procedure requires the user to specify the fuzziness parameter m a priori. The fuzzy clustering literature offers several empirical guidelines for choosing this parameter. When $m \to 1^+$, the memberships approach the binary set $\{0, 1\}$, and the procedure degenerates into its hard clustering counterpart. Conversely, if m is set too high, each MTS tends to receive a membership value of approximately $1/S$ in every cluster, resulting in an almost uniform, and thus uninformative, partition. Consequently, both extremes are typically avoided in practice [6,74].

The value $m = 2$ is widely regarded as a standard choice in the fuzzy clustering literature [33,57,52,12]. Furthermore, [80] analyzed the effect of the fuzziness parameter and also recommended $m = 2$ as the preferred value. Nonetheless, $m = 2$ may not always be optimal [86]. Many studies thus suggest selecting m from a plausible range, typically $[1.5, 2.5]$ [25,55]. However, as emphasized by [46], "there seems to exist no theoretically justifiable manner of selecting m", and, as noted by [19], "there is a lack of sound theoretical basis to justify the selection of the fuzziness parameter". In fact, selection of the optimal value of m is still an open problem.

While $[1.5, 2.5]$ is a commonly recommended range for the fuzziness parameter, our empirical results suggest that values of m in the range $[1.1, 2.2]$ perform best for the proposed FCPCA method. In particular, we observe that $m > 2.2$ often yields overly fuzzy partitions with diminished cluster structure, which negatively impacts both clustering accuracy and internal validity indices. Therefore, if the user does not specify m in advance, our implementation automatically selects it from the interval $[1.1, 2.2]$ by running the algorithm across a proper grid and choosing the solution that minimizes the CVI. The considered range can still be adjusted by the user if prior knowledge or application-specific considerations warrant a broader or narrower search.

If one considers only the associated hard clustering result, such as assigning each object to the cluster with the highest membership, the resulting partitions may appear identical across different values of m. However, this hard assignment masks the variation in fuzzy memberships that can occur under different m values. In such cases, external metrics based solely on single group labels may remain unchanged, despite meaningful differences in the underlying uncertainty or overlap between clusters. The proposed XB and FS indices are more sensitive to these differences and provide additional guidance in selecting the most informative fuzzy partition.

3.4. The FCPCA clustering procedure summary

Algorithm 1 outlines the FCPCA clustering procedure. Here, we make a short summary of the whole algorithm. In Step 1, each $\boldsymbol{X}_i$ is mean-centered column-wise, removing potential biases in each dimension. Steps 2–3 compute the block matrices (lagged covariance structures) and the number of principal components to retain. Steps 6–15 provide the core iterative process, updating membership degrees and re-estimating projection axes until convergence. Finally, Steps 16–18 select the best replication among R runs, providing the final membership matrix $\boldsymbol{U}^*$, the projection axes $\boldsymbol{C}^*(l)$, the set of class prototypes $\mathcal{P}^*$, and the total weighted reconstruction error E^*. The algorithm also returns the hard clustering labels (obtained by taking the largest membership for each $\boldsymbol{X}_i$) and the corresponding CVI. Before each run, the membership matrix $\mathbf{U}$ is randomly initialized (normalizing rows to ensure $\sum_s u_{is} = 1$). Since different random starts can lead to different results, we suggest repeating the entire clustering process (Steps 6–15) multiple times (3 in practice) and selecting the final solution that yields the best CVI.

4. Numerical studies

This section shows the performance of the proposed method through several simulated scenarios and some well-known MTS datasets.

4.1. MTS generated by VARMA process

In this section, we focus on MTS data generated from Vector Autoregressive Moving Average (VARMA) processes.

4.1.1. The VARMA model

This study is designed to address the limited research on fuzzy clustering for MTS with both high dimensions and varying lengths. VARMA models are widely used in time series analysis to capture lagged dependencies and moving average effects across multiple variables [16]. The general form of the VARMA process of order (a, b) is expressed as [67]

$$\mathbf{X}_t = \sum_{i=1}^{a} \mathbf{\Phi}_i \mathbf{X}_{t-i} + \sum_{j=1}^{b} \mathbf{\Theta}_j \varepsilon_{t-j} + \varepsilon_t, \tag{12}$$

where

Algorithm 1 Fuzzy clustering of MTS based on CPCA: FCPCA.

Input: MTS dataset $\boldsymbol{X} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N\}$, number of clusters S, fuzziness parameter m, maximum number of iterations $max\_iter = 1000$, number of replicates $R = 3$, convergence tolerance $\epsilon = 10^{-3}$.
Output: Optimal membership matrix $\boldsymbol{U}^*$, projection axes $\boldsymbol{C}^*(l)$, class prototypes $\mathcal{P}^*$, final total weighted reconstruction error E^*, the hard clustering labels vector, and the CVI value.

Normalize the MTS data: mean-center each column of $\boldsymbol{X}_i \leftarrow$ for every unit i.
Compute the block covariance matrices for lag 1 and lag 2 using Equation (1).
Compute $k(l)$ using Equation (3) for each lag l and retain this $k(l)$ for all iterations.
Initialize $E^* \leftarrow \infty$.
for $r = 1, \ldots, R$ **do** ▷ Replication loop
  Initialize membership matrix $\boldsymbol{U}$ randomly.
  Initialize $E \leftarrow \infty$, iteration counter $t \leftarrow 0$.
  repeat:
    Construct common weighted covariance matrix for each cluster using Equation (8).
    Perform SVD on the matrices in Step 9 to obtain the common projection axes $\boldsymbol{C}(l)$.
    Compute reconstruction error for each MTS to each cluster using Equation (5).
    Compute total weighted reconstruction error (across all clusters) E using Equation (7).
    Update membership matrix $\boldsymbol{U}$ using Equation (9).
    Increment iteration counter $iter \leftarrow iter + 1$.
  until $|E_{\text{prev}} - E| < \epsilon$ or $iter \geq max\_iter$.
  if $r = 1$ **or** $E < E^*$ **then** ▷ Select the best replication
    $E^* \leftarrow E$, $\boldsymbol{U}^* \leftarrow \boldsymbol{U}$, $\boldsymbol{C}^*(l) \leftarrow \boldsymbol{C}(l)$.
end for

- $\mathbf{X}_t \in \mathbb{R}^p$ is the p-dimensional vector of observed time series at time t.
- $\boldsymbol{\Phi}_i \in \mathbb{R}^{p\times p}$ are the coefficient matrices for the autoregressive (AR) terms.
- $\boldsymbol{\Theta}_j \in \mathbb{R}^{p\times p}$ are the coefficient matrices for the moving average (MA) terms.
- $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ is the p-dimensional vector of white noise with zero mean and covariance matrix $\boldsymbol{\Sigma}$. Hereafter, we take $\boldsymbol{\Sigma}$ to be the identity matrix.

4.1.2. Simulation setup

In each simulation replicate, we fix the series dimension and generate three groups of MTS: 10 VAR(1), 10 VMA(1), and 2 VARMA(1,1) time series. For the VAR and VMA groups, we randomly draw square coefficient matrices with entries in [–1,1]. We then evaluate their overall feedback strength (i.e., the largest one-step impact), and if it approaches unity, we lightly rescale the matrix to ensure stationarity (for VAR) or invertibility (for VMA), while preserving the random interaction pattern. The two VARMA series reuse these same VAR and VMA matrices, one slightly damped, the other amplified, so they naturally alternate between slow, VAR-like drifts and sharp, VMA-style reversals, as illustrated in Fig. 1. This design yields two well-separated clusters and a small hybrid group, providing a clear yet challenging testbed for evaluating the proposed fuzzy clustering method.

We consider two scenarios with different time series lengths:

1. All MTS have length 200.
2. All MTS have lengths randomly sampled from the range 200–600.

In each scenario, we also vary the dimensionalities of the time series (20, 60, or 100). Thus, this setup helps to evaluate the performance of our clustering method under various lengths and dimensionalities. In all cases, the true partition is given by two groups containing the 10 elements and 1 group containing 2 elements, in accordance with the generating mechanism.

4.1.3. Clustering and membership thresholds

The algorithm is executed with two clusters ($S = 2$). Since our method produces fuzzy memberships, We apply a membership threshold of 0.7 to assign a final label to each time series. A detailed justification for this threshold is provided by [46], and it is also commonly adopted in the literature; see, for example, [22,44]. Specifically, a given time series is assigned to the cluster in which it has a membership degree above the threshold [47]. We expect this to occur for the well-separated VAR(1) and VMA(1) time series. In contrast, if a given time series has both membership degrees below the threshold, we label such a series as “mixed”, creating a third group. We expect this to happen for the VARMA(1,1) time series, indicating that these series do not strongly favor either cluster, reflecting the combined characteristics of VAR(1) and VMA(1) [46]. Later, we also consider a threshold of 0.6 to assess the sensitivity of our fuzzy clustering results to the choice of the membership cutoff.

4.1.4. Evaluation criterion

To quantify the agreement between the true partition and the crisp version of the clustering solution obtained as described above, we employ the Rand index (RI) [61]. The RI measures the proportion of pairwise agreements between two partitions:

$$\text{RI} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{13}$$

where TP and TN are the numbers of correctly clustered and correctly separated pairs, respectively, while FP and FN represent misclustered pairs. RI has range $[0, 1]$. Higher RI values indicate greater similarity between the experimental and true partitions.

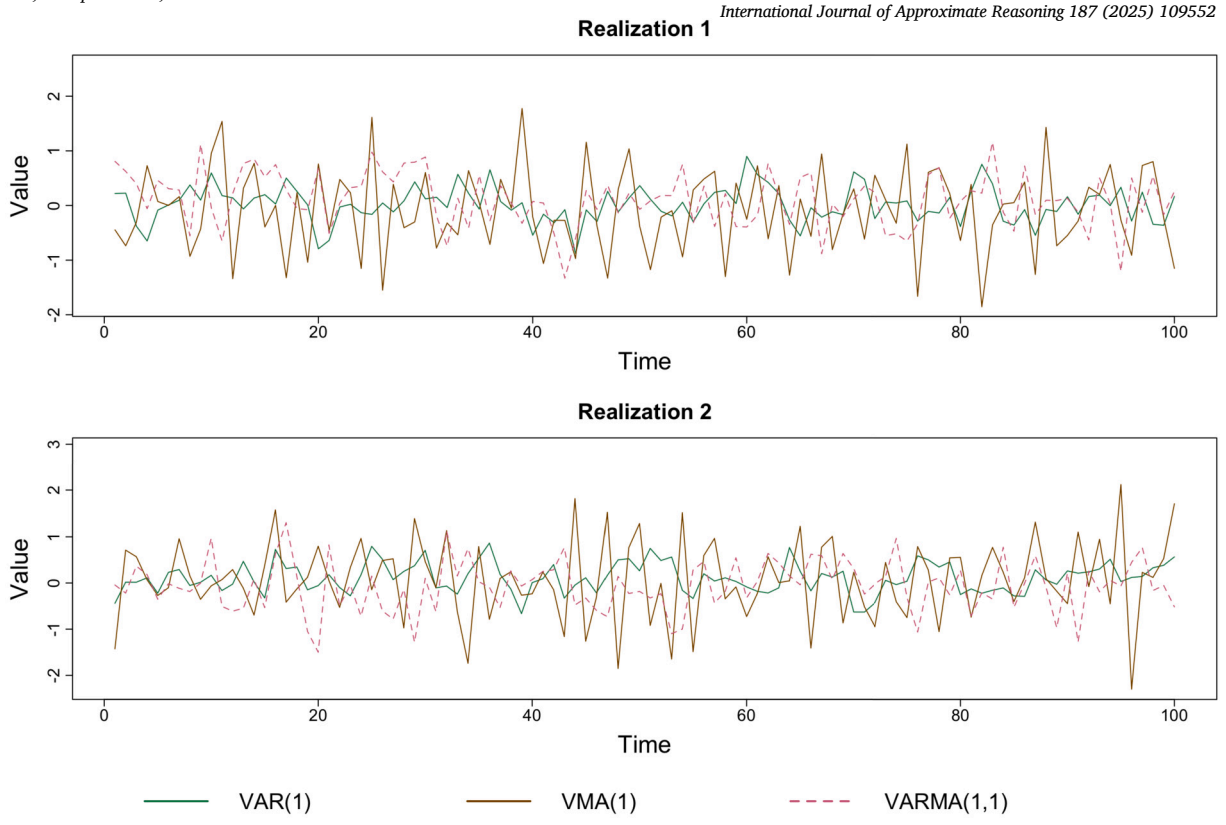

Fig. 1. Example of two realizations of VARMA models. In each case, realizations of the associated VAR(1) and VMA(1) processes are also shown.

This setup allows us to evaluate our proposed clustering method under both well-separated and partially overlapping time-series dynamics. This criterion was also suggested and used by [40] in their paper. Besides RI, we also consider the adjusted Rand index (ARI) [32] and the fuzzy version of the Rand index (RIF) [9].

4.1.5. Competing methods

To better understand the performance of FCPCA, we compare it with the following competing approaches:

- **Variable-based principal component analysis (VPCA) clustering.** This method combines VPCA with a spatial-weighted matrix distance fuzzy clustering approach, capturing both value and spatial differences among features [30]. We include this method as a natural benchmark because it also involves dimensionality reduction. We use the `vpca_clustering()` function from the R package `mlmts` [41].
- **Fuzzy C-medoids (FCMD).** This is a fuzzy variant of the well-known K-medoids algorithm, where each cluster is represented by a single medoid (i.e., an actual data point) rather than a centroid derived from averaging. Data points receive partial memberships in each cluster based on their distance to the medoids [13]. In our implementation, we use the function `tsclust()` from the R package `dtwclust` [66], and the distance is set as `dtw_basic`.

- **Wavelet-based clustering (MODWT).** This method first extracts features based on the maximum overlap discrete wavelet transform (MODWT) for each MTS using the function `dis_modwt()` from the `mlmts` package, following [18]. The resulting feature vectors are then clustered with fuzzy C-means.
- **Structure-based statistical-feature clustering (WWW).** For each MTS, we compute the structure-based statistical features proposed by [73] using the function `dis_www()` from `mlmts`. These feature vectors are subsequently clustered with fuzzy C-means.
- **Quantile cross-spectral density clustering (QCD).** Each MTS is first mapped to a vector of quantile cross-spectral density-based estimates obtained with the function `dis_qcd()` from `mlmts`. We then apply PCA to these vectors and cluster the resulting low-rank scores with fuzzy C-means, following [44].

Note that only one replicate is used for FCPCA to obtain the grouping results. This is also the case in the following simulation study.

Table 1
Mean RI (top), ARI (middle), and RIF (bottom) for the six methods, reported over varying time-series lengths, fuzziness exponents m, and dimensionalities (20 / 60 / 100); membership threshold fixed at 0.7.

Length	m	FCPCA	VPCA	FCMD	MODWT	WWW	QCD
200	1.2	**0.91/0.91/0.91**	0.49/0.52/0.46	0.35/0.28/0.28	**0.91/0.91/0.91**	0.80/0.85/**0.91**	**0.91/0.91/0.91**
		0.83/0.83/0.83	0.20/0.10/0.30	0.13/0.04/0.04	**0.83/0.83/0.83**	0.50/0.61/0.81	**0.83/0.83/0.83**
		0.91/0.91/0.91	0.50/0.51/0.49	0.40/0.36/0.37	**0.91/0.91/0.91**	0.75/0.79/0.87	**0.91**/0.90/0.88
	1.4	**0.91/0.91/0.91**	0.50/0.55/0.55	0.23/0.28/0.35	**0.91/0.91/0.91**	0.81/0.80/0.87	**0.91/0.91/0.91**
		0.83/0.83/0.83	0.21/0.30/0.30	0.11/0.20/0.20	**0.83/0.83/0.83**	0.50/0.59/0.81	**0.83/0.83/0.83**
		0.90/0.89/0.90	0.54/0.58/0.58	0.30/0.48/0.49	0.89/0.88/0.85	0.73/0.72/0.76	0.87/0.81/0.87
	1.6	**1.00/0.92/0.92**	0.55/0.55/0.55	0.22/0.43/0.28	0.92/**0.92/0.92**	0.82/0.74/0.71	0.98/0.86/0.91
		1.00/0.83/0.83	0.29/0.30/0.30	0.10/0.19/0.22	**0.83/0.83/0.83**	0.54/0.61/0.81	**0.83/0.83/0.83**
		0.99/0.83/0.86	0.60/0.62/0.62	0.40/0.43/0.52	0.89/0.81/0.76	0.71/0.69/0.68	0.80/0.72/0.81
	1.8	0.45/**1.00**/0.91	0.55/0.55/0.55	0.25/0.28/0.24	0.96/**1.00/1.00**	0.80/0.69/0.50	**1.00**/0.75/0.86
		0.61/**1.00/0.83**	0.29/0.30/0.30	0.10/0.11/0.10	**0.83/0.83/0.83**	0.61/0.62/0.81	**0.83/0.83/0.83**
		0.60/**0.76/0.80**	0.60/0.62/0.62	0.45/0.47/0.44	**0.86**/0.75/0.69	0.69/0.65/0.62	0.74/0.06/0.75
	2.0	0.09/0.40/**0.93**	0.55/0.48/0.55	0.36/0.20/0.21	**1.00/1.00**/0.09	0.76/0.56/0.23	0.92/0.32/0.10
		0.23/0.35/**0.83**	0.30/0.19/0.34	0.18/0.09/0.10	**0.83/0.83/0.83**	0.61/0.66/0.82	**0.83/0.83/0.83**
		0.51/0.65/**0.73**	0.60/0.58/0.62	0.57/0.47/0.47	**0.83/0.70**/0.64	0.67/0.62/0.58	0.70/0.62/0.50
	2.2	0.09/0.09/0.09	0.18/0.27/0.27	0.14/0.23/0.43	**1.00/0.78/0.09**	0.70/0.48/0.09	0.70/0.19/0.10
		0.23/0.24/0.23	0.04/0.10/0.11	0.03/0.09/0.28	**1.00/0.80**/0.23	0.61/0.72/**0.57**	0.73/0.31/0.23
		0.50/0.51/0.54	0.38/0.45/0.43	0.34/0.35/0.56	**0.80/0.67/0.59**	0.65/0.60/0.54	0.66/0.58/0.50
200-600	1.2	**0.91/0.91/0.91**	NA	0.35/0.25/0.25	**0.91/0.91/0.91**	0.80/0.85/**0.91**	NA
		0.83/0.83/0.83	NA	0.25/0.11/0.10	**0.83/0.83/0.83**	0.50/0.61/0.81	NA
		0.91/0.91/0.91	NA	0.61/0.43/0.42	**0.91/0.91/0.91**	0.75/0.79/0.87	NA
	1.4	0.91/**0.93/0.91**	NA	0.28/0.35/0.35	**0.99**/0.91/**0.91**	0.73/0.86/0.62	NA
		0.83/0.83/0.83	NA	0.18/0.23/0.24	**0.83/0.83/0.83**	0.35/0.65/0.35	NA
		0.89/**0.91/0.90**	NA	0.41/0.59/0.60	**0.91**/0.88/0.85	0.67/0.82/0.58	NA
	1.6	0.91/**1.00/0.91**	NA	0.20/0.17/0.17	**1.00**/0.91/**0.91**	0.72/0.87/0.61	NA
		0.83/0.83/0.83	NA	0.10/0.07/0.05	**0.83/0.83/0.83**	0.35/0.65/0.37	NA
		0.82/**0.90/0.86**	NA	0.45/0.36/0.37	**0.91**/0.88/0.85	0.67/0.70/0.54	NA
	1.8	0.97/**0.98/0.94**	NA	0.30/0.39/0.35	**1.00**/0.93/0.91	0.70/0.76/0.58	NA
		0.83/0.95/0.83	NA	0.20/0.24/0.22/	**0.83/0.83/0.83**	0.35/0.68/0.39	NA
		0.74/**0.83/0.86**	NA	0.51/0.54/0.52	**0.88**/0.76/0.71	0.67/0.65/0.52	NA
	2.0	0.63/0.09/0.63	NA	0.23/0.32/0.39	**1.00/0.92/0.70**	0.69/0.54/0.55	NA
		0.73/0.23/0.65	NA	0.11/0.21/0.25	**1.00/0.83/0.67**	0.74/0.68/0.39	NA
		0.67/0.50/**0.68**	NA	0.41/0.51/0.59	**0.85/0.70**/0.66	0.67/0.62/0.50	NA
	2.2	0.23/0.09/0.09	NA	0.14/0.23/0.43	**1.00/0.81**/0.35	0.68/0.28/**0.49**	NA
		0.30/0.23/0.23	NA	0.05/0.12/0.23	**0.83/0.83**/0.45	0.37/0.68/**0.62**	NA
		0.58/0.50/0.50	NA	0.30/0.45/0.60	**0.83/0.68/0.63**	0.66/0.59/0.49	NA

4.1.6. Results

The quality of the clustering outcomes of the proposed and alternative approaches is assessed as follows. For each simulation scenario and time series dimension, we randomly generate three distinct sets of coefficient matrices that satisfy the stationarity condition for the VAR(1) and VMA(1) processes. For each set, we simulate MTS processes 100 times, compute the Rand Index (RI) for each method at each simulation, and then take the per-method average across all 300 runs (3 sets $\times$ 100 runs). For the three methods, this is done for several values of the fuzziness parameter.

The clustering performance of all approaches is shown in Table 1 using a threshold of 0.7. The headers of each column refer to the number of dimensions of the time series. Since VPCA cannot handle MTS of varying lengths, and both VPCA and QCD are computationally expensive, especially as dimensionality and series length increase. We omit their results in these cases and mark the corresponding entries as NA. The best mean performance in each setting is highlighted in bold.

When the fuzziness parameter is set to a high value (e.g., $m = 2.2$), the accuracy of FCPCA decreases, as overly diffuse memberships weaken cluster separations. This finding also motivates our decision to restrict the implementation of FCPCA to the range $m \in [1.1, 2.2]$. However, at more moderate fuzziness levels (around 1.2–1.9), FCPCA clearly outperforms many of the competing methods, regardless of the lengths and dimensionalities of the MTS data. We can notice that MODWT is also a competitive method, and in many cases, it has comparable results to FCPCA. We observe that the RI value is frequently 0.91. A detailed examination of the obtained partitions reveals that, in such instances, the method fails to detect the two fuzzy series at the corresponding threshold, although it works perfectly for the well-separated groups. Therefore, an additional measure must be provided to quantify the success rate of detecting the fuzzy series.

Specifically, we consider the average number of VARMA(1, 1) series that are correctly recognized as fuzzy out of the two. Table 2 presents the results for FCPCA using the optimal value of m and also considers another threshold at 0.6. Results for MTS with a fixed length of 400 are also reported as an intermediate case. As a reference, the most selected m for this simulation across all scenarios is

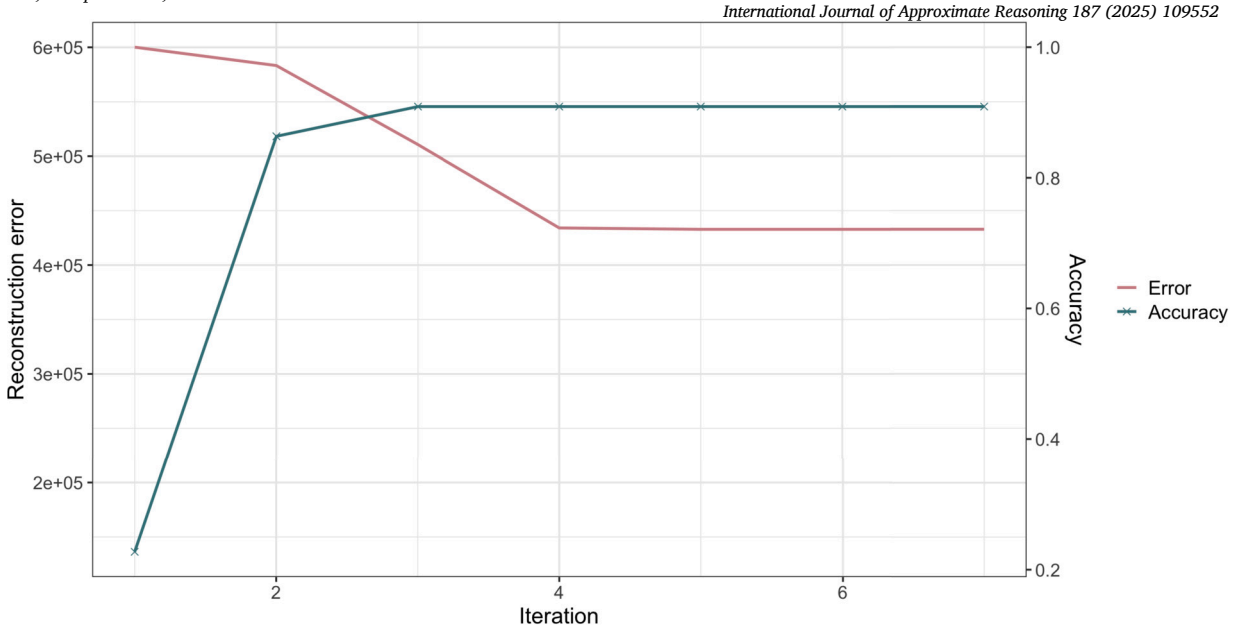

Fig. 2. Intermediate optimization path for FCPCA with $m = 1.2$ on MTS of length 200–600. The left axis tracks the total reconstruction error, while the right axis shows RI accuracy at each iteration.

Table 2
Average number of fuzzy series detected out of two using the optimal m for method FCPCA with thresholds 0.7 and 0.6.

		Dim		
Threshold	Length	20	60	100
0.7	200	0.24	2.00	2.00
	400	0.00	2.00	2.00
	200–600	2.00	1.78	2.00
0.6	200	0.00	1.55	0.00
	400	0.00	0.00	0.00
	200–600	2.00	0.00	1.77

mostly 1.7 or 1.8 using the XB index. We see that FCPCA can detect the fuzzy series in most cases, although at a threshold of 0.6, the results appear less ideal. This highlights the influence of the chosen threshold on the detection performance.

Fig. 2 illustrates the intermediate behavior of FCPCA. We plot the total reconstruction error alongside the RI accuracy over successive iterations for MTS generated under $m = 1.2$, with series lengths ranging from 200 to 600. The sharp drop in error during the first two iterations, followed by a gradual plateau, indicates that the algorithm quickly identifies the dominant common principal components. Simultaneously, clustering accuracy exceeds 0.90 after iteration 2 and stabilizes thereafter.

Fig. 3 explores the impact of the stopping tolerance ϵ on convergence, with the fuzziness parameter fixed at $m = 1.2$ and series lengths again between 200 and 600. We vary ϵ from 1×10^{-5} to 1×10^{-2}. As expected, looser tolerances lead to faster convergence (fewer iterations), yet final clustering accuracy remains effectively unchanged. This pattern holds consistently across both simulated and real-world datasets, suggesting that ϵ has minimal practical influence on clustering quality.

Fig. 4 reports the mean runtime of FCPCA under three series lengths and three different MTS dimensions. Each point represents the average over 10 random simulations, with vertical bars denoting ± 1 standard deviations. As anticipated, runtime increases with both dimension and length, yet remains under 8 seconds even for 100-dimensional series spanning 200–600 time points.

We further assessed the computational time of FCPCA in two large–scale regimes.

(a) High–dimensional, relatively smaller sample: $p = 200$ and $p = 400$ with 50 MTS per cluster $(N = 150)$.

(b) Larger amount of series, moderate–dimensional: $p = 20, 60$, and 100 with $N = 2000$ series in total.

In both settings the series length was fixed at $T = 200$. Fig. 5 shows the mean runtime across 20 replicates. The number of replicates is reduced here because each run is moderately time-consuming. Error bars represent ± 1 standard deviation.

Panel (a) indicates that runtime increases approximately quadratically with the dimension: doubling p from 200 to 400 leads to a four-fold rise in runtime. This behavior is consistent with the dominant $\mathcal{O}(Np^2)$ cost of computing and averaging the $p \times p$ lag-covariance matrices. The memory footprint likewise scales with p^2.

Panel (b) indicates a near–linear increase with p when N is large, confirming the theoretical $\mathcal{O}(Npk)$ time complexity: doubling p from 60 to 100 raises runtime by a factor of ≈ 1.8, consistent with the slope observed in the high-N experiment. These results suggest

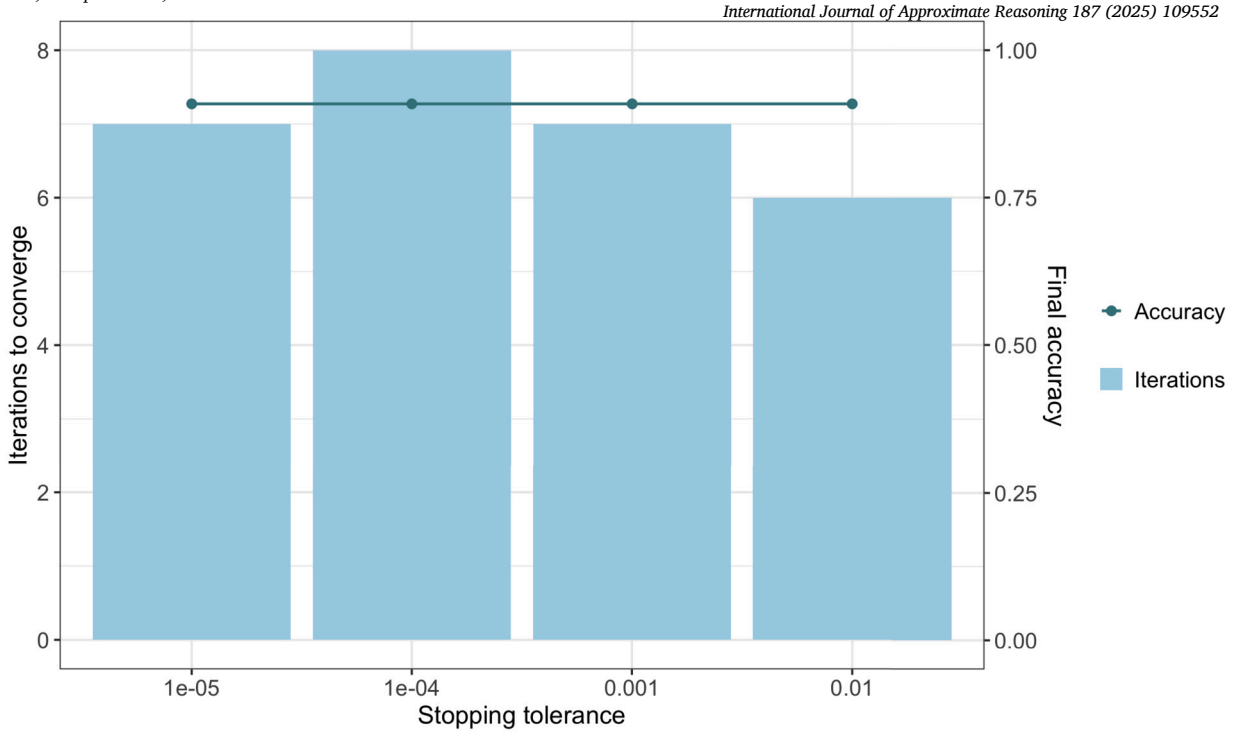

Fig. 3. Convergence behavior versus stopping tolerance – iterations (bars) and RI accuracy (line) for FCPCA with $m = 1.2$ on MTS of length 200–600.

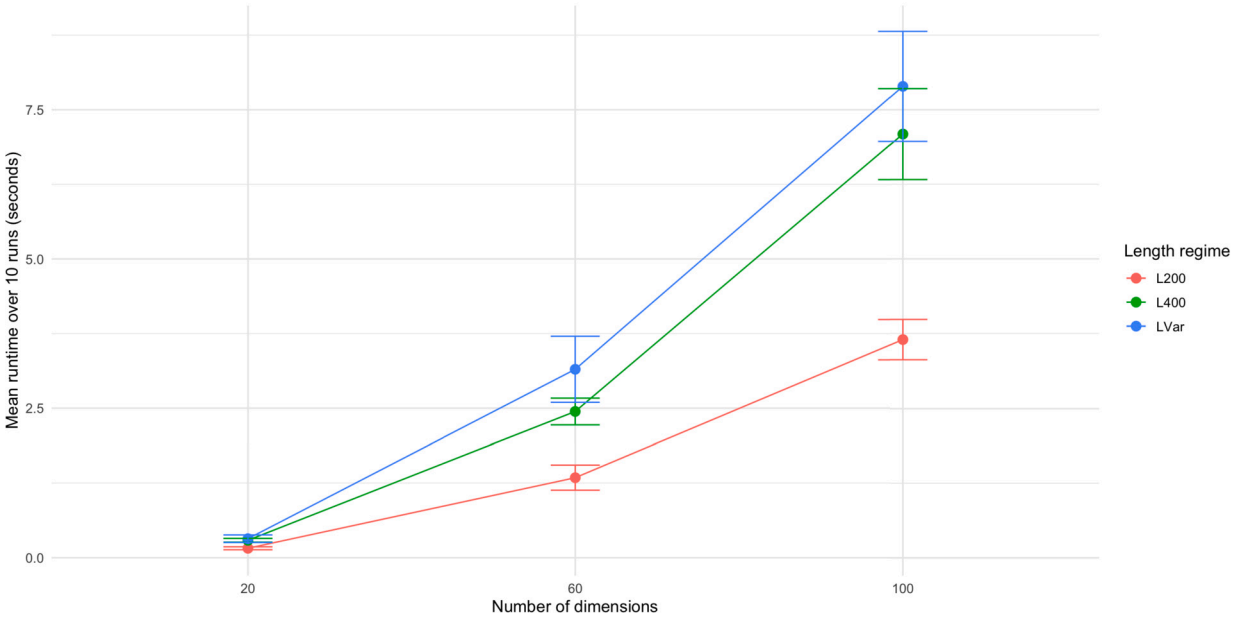

Fig. 4. Mean runtime of FCPCA under different series lengths and dimensions.

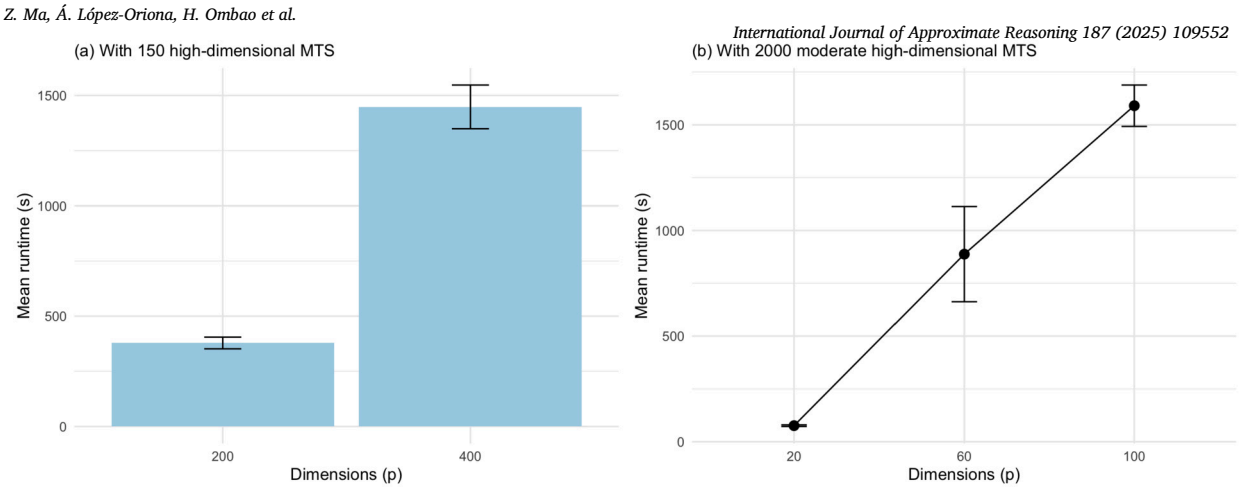

Fig. 5. FCPCA performance in larger scale regimes.

Table 3
The mean RI and number of fuzzy series successfully detected out of ten for FCPCA in the simulated EEG data with a threshold of 0.7.

Length	256			512		
Channels	32	64	128	32	64	128
RI	0.84	0.86	0.82	0.89	0.85	0.87
Number of fuzzy series detected	5.10	5.24	5.78	6.70	5.44	6.04

that CPU time remains manageable when either p or N is extreme in isolation. The practical bottleneck in the joint extreme regime is memory, not computation.

4.2. EEG data simulated by AR(2) mixture model

In this section, we simulate EEG data using AR(2) mixing processes, chosen for their ability to generate oscillatory signals with well-defined peaks and spectral bandwidths. EEG, a widely used non-invasive measure of brain activity, captures essential neural dynamics [65]. By replicating these rhythms, AR(2) mixing processes provide realistic models for studying cognitive mechanisms and neurological disorders. Specifically, MTS are generated by AR(2) mixing processes that represent five distinct EEG frequency bands (delta, theta, alpha, beta, and gamma). The AR(2) coefficients for each band are designed to capture its characteristic oscillatory properties. A more detailed explanation of the simulation mechanism for EEG data using the AR(2) mixture model can be found in the paper by [54].

In our simulations, we generate MTS with two lengths (256 and 512), each with three different channel counts (32, 64, and 128). These lengths and channel configurations align with typical EEG sampling rates (e.g., 256 or 512 Hz) and commonly used EEG channel setups. We consider two well-separated groups. Group 1 is dominated by delta and gamma frequencies, and group 2 is dominated by theta, alpha, and beta frequencies. We then form the third group by combining the first half of the MTS from group 1 with the second half of the MTS from group 2. An illustrative example is provided in Fig. 6. As shown, the time series in group 3 clearly exhibits characteristics from both group 1 and group 2, with a visible transition at the midpoint. This mixed structure reflects the fuzzy nature of the group. Each group consists of 10 MTS objects.

We consider a threshold of 0.7 and conduct 100 simulation replicates. The results are presented in Table 3, where we report both the mean RI and the mean number of fuzzy series successfully detected (out of ten) using the optimal fuzziness parameter using the XB index. Only the FCPCA results are presented, as the primary objective of this paper is to demonstrate the practical utility of FCPCA, with EEG serving as the main application. Accordingly, we aim to keep this section concise and focused solely on FCPCA. The most selected m here across all scenarios is 1.2. We can see that the clustering performance of FCPCA is quite consistent, considering the different lengths and dimensions of the MTS. Our experimental results indicate that selecting a higher threshold (e.g., 0.8) can further improve clustering accuracy in this scenario.

4.3. Evaluating FCPCA in a hard/crisp clustering context

Since FCPCA is a fuzzy generalization of the CPCA-based method, we aim to assess the benefits of introducing fuzziness. Mc2PCA and ROBCPCA employ a hard clustering approach, whereas FCPCA assigns degrees of membership to each cluster. To facilitate a direct comparison with the hard clustering baseline, we use a relatively low fuzziness parameter, namely $m = 1.1$, so that the memberships remain nearly crisp. In addition, we convert the resulting fuzzy partition into a hard one using the row-wise maximum rule, assigning each series to the cluster with the highest membership. Table 4 provides a brief overview of the real datasets used in this study, which

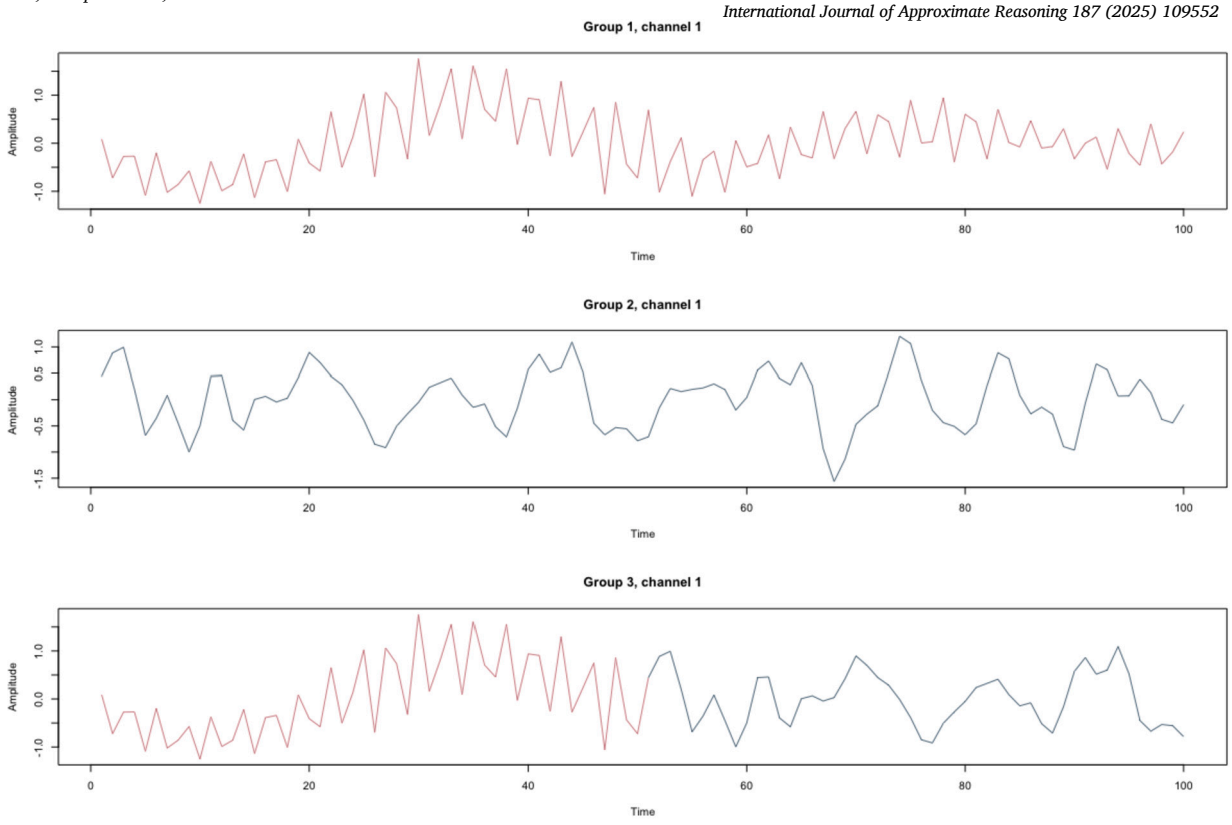

Fig. 6. The simulated EEG data example.

Table 4
Datasets brief description.

Dataset	Size	Length	Dimension	Number of classes
Articulary word recognition	300	144	9	25
Atrial fibrillation	15	600	2	3
Basic motions	40	100	6	4
Japanese vowels	370	29	12	9
NATOPS	180	51	24	6
Racket sports	152	30	6	4

are commonly considered to evaluate the performance of various clustering algorithms (all of them have associated true partitions). The RI and ARI are used for evaluation. The RIF is not used since we have hard clustering methods. The results presented in Table 5 are obtained by averaging the performance of each method over 100 replications for each dataset. Standard deviations are shown in parentheses. Note that some real datasets do not meet the conditions required to compute certain features of WWW, therefore, these values are reported as NA.

Interestingly, Table 5 shows that FCPCA achieves a higher clustering accuracy than the original hard clustering methods in several real-world datasets when a low value for the fuzziness parameter is considered. This improvement could suggest that even a small degree of fuzziness can help capture the underlying uncertainties in the data, leading to more robust and accurate cluster assignments. Notably, FCPCA never ranks lower than third place on any dataset.

5. Application of FCPCA to EEG drowsiness data

Driver drowsiness is one of the leading causes of road fatalities [14], and EEG analysis provides a direct window into the neurophysiological activities of the brain that signal the onset of fatigue. This analysis is critical because, despite the susceptibility of EEG to inter-subject variability (including mental and physical drifts), it remains one of the most informative signals for detecting drowsiness. Ultimately, analyzing EEG drowsiness data can deepen our understanding of the neurocognitive transitions from alertness to drowsiness, which is essential for designing interventions that can enhance driver safety and prevent transportation hazards.

The dataset we use for the study is available on the website https://figshare.com/articles/dataset/EEG_driver_drowsiness_dataset/14273687?file=30707285. It contains 2022 EEG samples of size 384×30 from 11 subjects with labels of alert and drowsy. Each

Table 5
Mean RI (upper) and ARI (lower) on six datasets. Standard deviations are reported in parentheses. The best result in each row is shown in bold.

Dataset	Mc2PCA	ROBCPCA	FCPCA	VPCA	FCMD	MODWT	WWW	QCD
Articulary word recognition	0.17 (0.02)	0.21 (0.02)	0.50 (0.01)	**0.77** (0.04)	0.66 (0.05)	0.62 (0.02)	NA (NA)	0.68 (0.03)
	0.00 (0.00)	0.02 (0.00)	0.34 (0.02)	**0.75** (0.02)	0.50 (0.04)	0.46 (0.02)	NA (NA)	0.54 (0.03)
AtrialFibrillation	0.38 (0.11)	0.49 (0.06)	0.47 (0.01)	0.43 (0.04)	0.45 (0.04)	0.46 (0.03)	**0.53** (0.00)	0.38 (0.05)
	0.18 (0.17)	0.36 (0.04)	0.32 (0.00)	0.30 (0.03)	0.34 (0.11)	0.34 (0.01)	**0.42** (0.01)	0.17 (0.06)
Basic motions	0.38 (0.04)	0.46 (0.07)	0.79 (0.04)	0.46 (0.05)	0.50 (0.14)	0.69 (0.03)	NA (NA)	**0.90** (0.11)
	0.23 (0.02)	0.28 (0.03)	0.65 (0.04)	0.26 (0.02)	0.33 (0.14)	0.65 (0.02)	NA (NA)	**0.81** (0.14)
Japanese Vowels	0.36 (0.07)	0.57 (0.06)	**0.73** (0.04)	**0.73** (0.04)	0.28 (0.21)	0.56 (0.03)	NA (NA)	0.43 (0.02)
	0.21 (0.11)	0.42 (0.06)	**0.62** (0.05)	0.58 (0.05)	0.42 (0.08)	0.38 (0.02)	NA (NA)	0.22 (0.01)
NATOPS	0.38 (0.11)	0.47 (0.08)	0.53 (0.01)	**0.56** (0.04)	0.43 (0.15)	0.50 (0.03)	NA (NA)	0.48 (0.06)
	0.28 (0.17)	0.32 (0.04)	0.35 (0.00)	**0.40** (0.03)	0.34 (0.11)	0.31 (0.00)	NA (NA)	0.38 (0.07)
Racket sports	0.31 (0.06)	0.33 (0.03)	0.47 (0.03)	0.44 (0.09)	0.26 (0.01)	0.40 (0.00)	NA (NA)	**0.64** (0.03)
	0.01 (0.02)	0.02 (0.01)	0.22 (0.03)	0.20 (0.04)	0.00 (0.00)	0.04 (0.01)	NA (NA)	**0.42** (0.04)

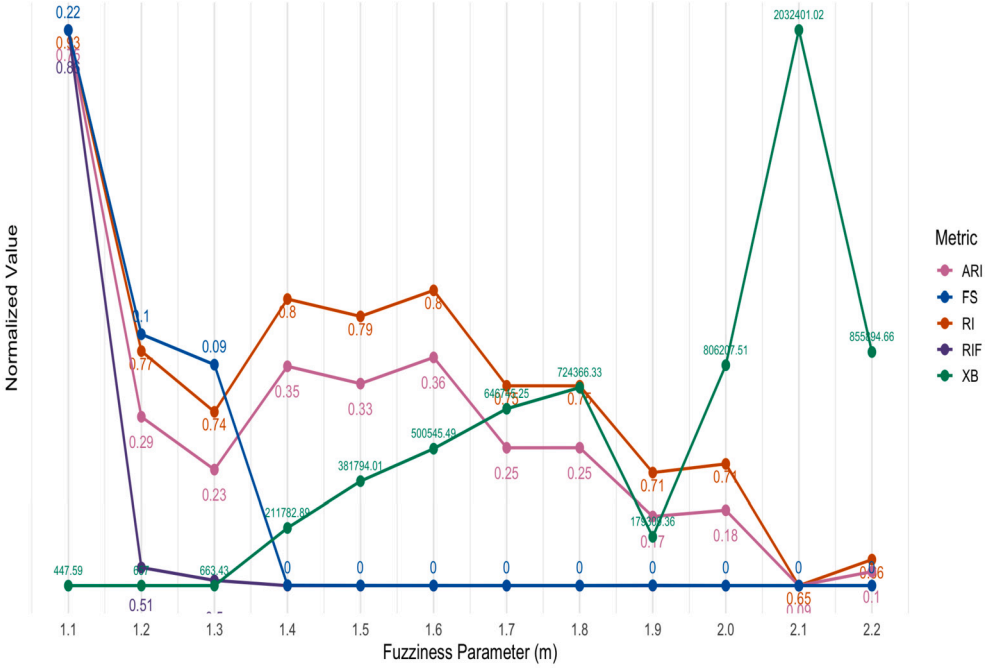

Fig. 7. The selection of m and comparison to the CVIs and clustering performances of subject 11.

sample is a 3-second EEG data with 128 Hz from 30 EEG channels. In this dataset, participants operated a simulated vehicle, attempting to maintain their position in the center of a lane. Lane-departure events were randomly introduced to mimic real-world conditions, such as slight road curvatures or small obstacles, causing the car to drift left or right. Each lane-departure event (a "trial") consisted of a baseline period, the moment the car started drifting (deviation onset), the steering response of the participant (response onset), and the time when the car returned to the center (response offset). Because the task was monotonous, participants often became drowsy. For each trial, a 3-second EEG segment was extracted immediately before the deviation onset, capturing the brain state leading up to the drift. A detailed description of this dataset is provided in [14].

As mentioned by [82], since EEG signals have high variability and instability, data can differ substantially across subjects. Thus, we decided to split the whole dataset into 11 groups, each containing only the EEG signals from one subject. Later, our FCPCA algorithm is applied to each individual set of signals, considering two groups, $S = 2$. All results are obtained with $R = 3$ replicates.

Here, we first present the results for subject number 11 from the original dataset. In particular, 226 samples were collected for this subject.

Fig. 7 shows the normalized value of XB, FS, and the corresponding RI, RIF and ARI values. Both XB and FS suggest the selection of $m = 1.1$. Notably, the best clustering performance is also obtained with this m. The RI and ARI exhibit fluctuations as m increases, while the membership matrix $\mathbf{U}$ remains approximately balanced, with membership degrees near 0.5 for both clusters. In certain instances, the RI and ARI appear to be influenced by random assignment due to identical membership degrees. When we move from $m = 1.2$ to $m = 1.4$, the FS drops to 0 and the XB explodes, indicating blurred, poorly separated clusters, while the RIF settles at its baseline level of 0.50 (expected when two clusters share little meaningful overlap with the reference). These internal, fuzzy-aware

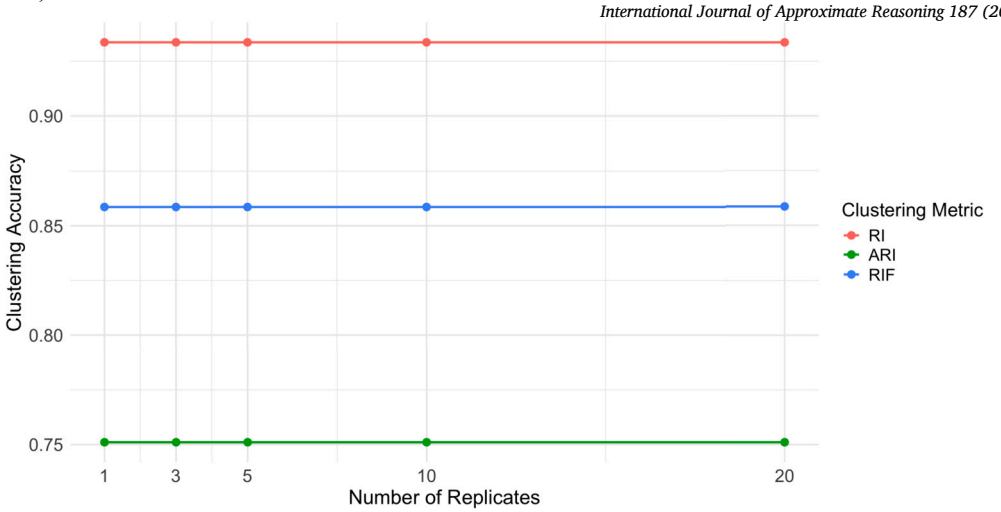

Fig. 8. The influence of the number of replicates in FCPCA.

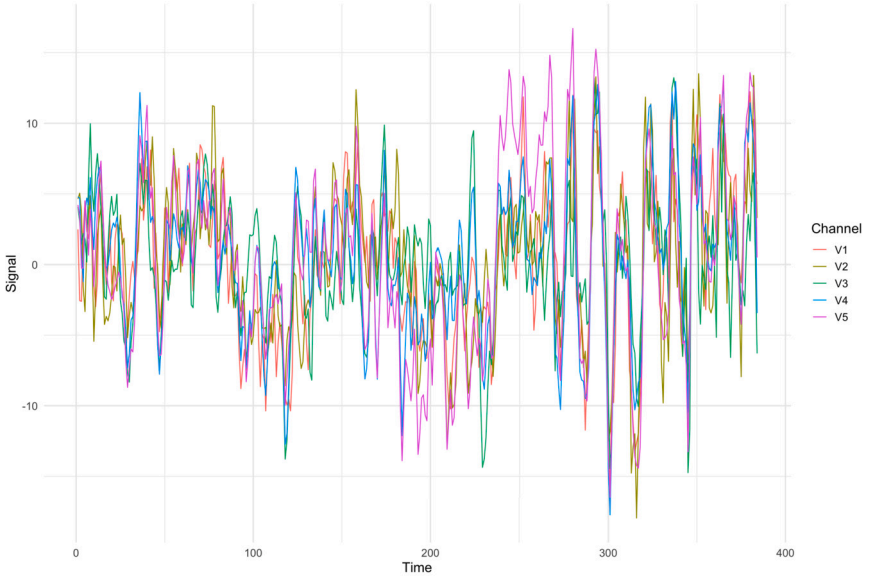

Fig. 9. The example EEG signals from the first five channels of subject 11 with high membership in the alert state.

measures speak directly to cluster compactness and separation, and therefore carry more weight than the modest improvements observed in ARI/RI. In practice, we favor the parameter setting that yields better validity scores, even if the corresponding external accuracy is slightly lower (note that the true labels are usually unknown), making $m = 1.2$ preferable to $m = 1.4$. That said, for subject 11, the optimal choice of m is 1.1.

To investigate the sensitivity of our fuzzy clustering algorithm to the number of replicates R used during initialization, we conducted a controlled experiment on subject 11 with $m = 1.1$. We varied $R \in \{1, 3, 5, 10, 20\}$ and evaluated the clustering performance.

As shown in Fig. 8, the clustering results remain relatively stable across the considered range of R. However, increasing R can result in improved clustering accuracy compared to using $R = 1$ in some contexts, where a higher number of replicates can help avoid poor local optima and enhance performance.

Based on these experiments, we recommend $R = 3$ as a reasonable and computationally efficient default that balances robustness and speed. That said, for simpler datasets or when computational resources permit, users may choose larger values of R to further improve stability.

The resulting membership degree structure is shown in Fig. 12. As one can see, some samples exhibit nearly crisp membership in either the drowsy or the alert cluster. For instance, a high drowsy state membership might occur when a driver travels along a monotonous route with little traffic, where boredom can set in and induce drowsiness. Conversely, a high alert state membership may appear when the driver encounters heavier traffic, needs to stay vigilant to avoid collisions, or notices engaging stimuli on road signs or in music. Fig. 9 shows an EEG sample from subject 11 with high membership in the alert state. Fig. 10 presents an

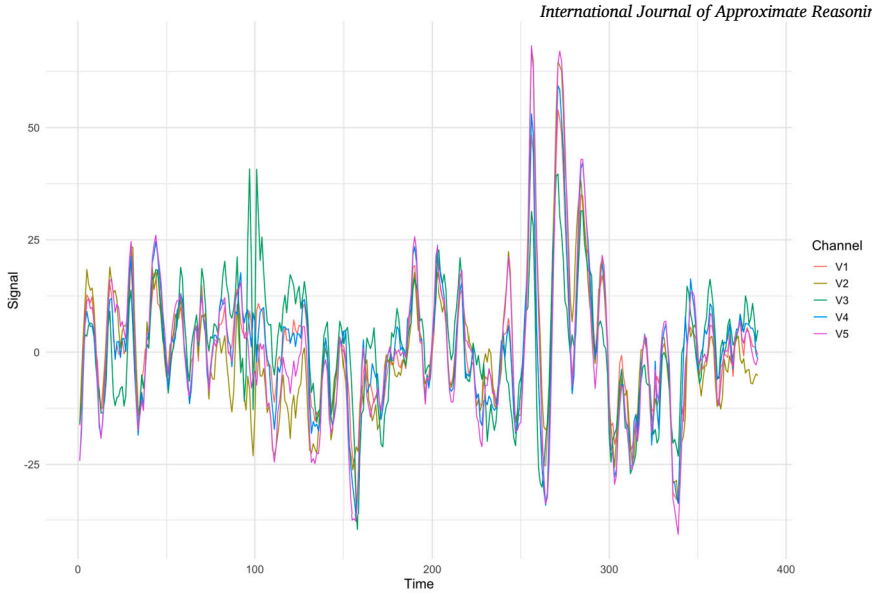

Fig. 10. The example EEG signals from the first five channels of subject 11 with high membership in the drowsy state.

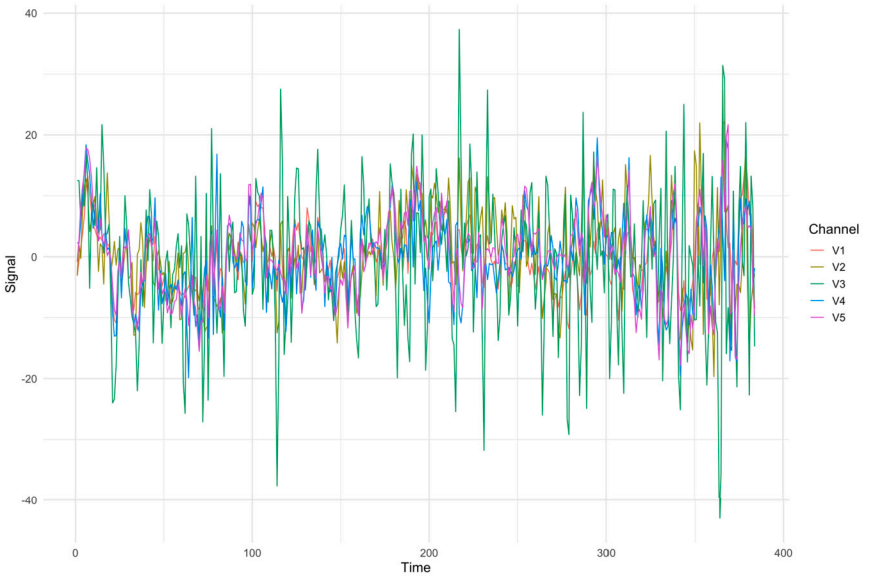

Fig. 11. The example EEG signals from the first five channels of subject 11 with substantial memberships in both alert and drowsy states.

example of the drowsy state. Fig. 11 illustrates a sample exhibiting substantial memberships in both brain states. The drowsy state is primarily characterized by slow oscillations in the delta and theta bands, reflected in the EEG signals as larger amplitude fluctuations and lower frequency components across all channels. This contrasts with the alert state, which exhibits higher-frequency, lower-amplitude patterns, consistent with increased beta and gamma band activity commonly associated with cognitive engagement and alertness. The mixed state displays a combination of these features: the amplitude is more moderate, and the frequency content appears blended, with traces of both lower and higher-frequency components. This variability across channels in the mixed state explains its substantial fuzzy membership in both clusters, as the signal exhibits partial characteristics of both the alert and drowsy conditions [49,27].

More importantly, our fuzzy clustering algorithm reveals that certain samples exhibit substantial memberships in both clusters, indicating a transitional or mixed state between drowsiness and alertness. This phenomenon cannot be captured by hard clustering, which forces each sample into a single cluster. However, it is realistic to assume that drivers experience a continuum of alertness, transitioning gradually from an alert state to a drowsy state (and vice versa) over the course of a long trip. For instance, a driver who has been on the road for an extended period may start feeling fatigued but then becomes partially re-alerted upon noticing a lane drift or other stimuli, requiring additional time to regain full attentiveness.

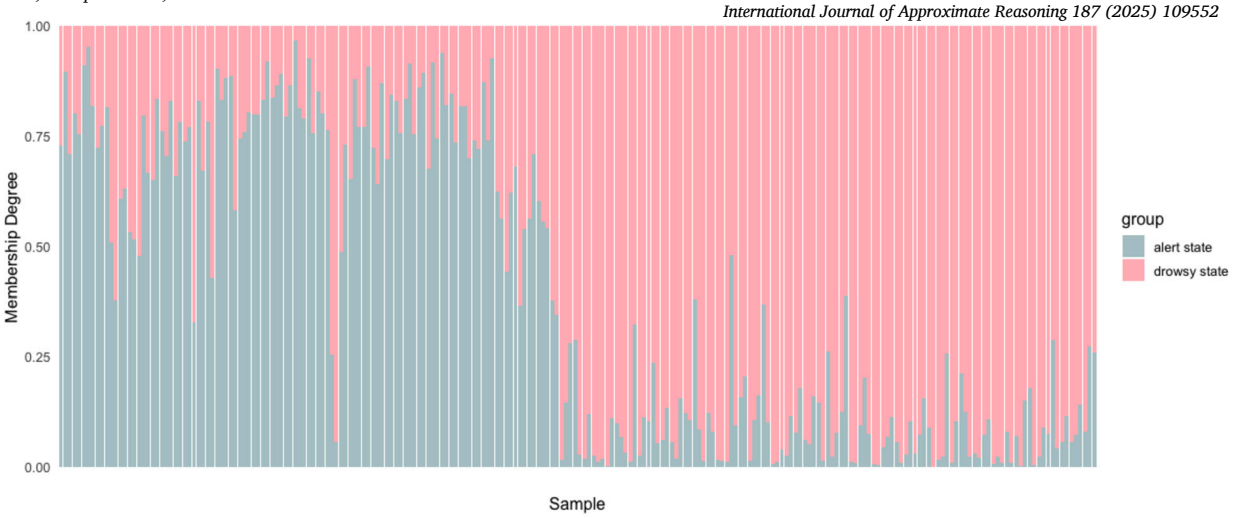

Fig. 12. Membership matrix of the samples of subject 11.

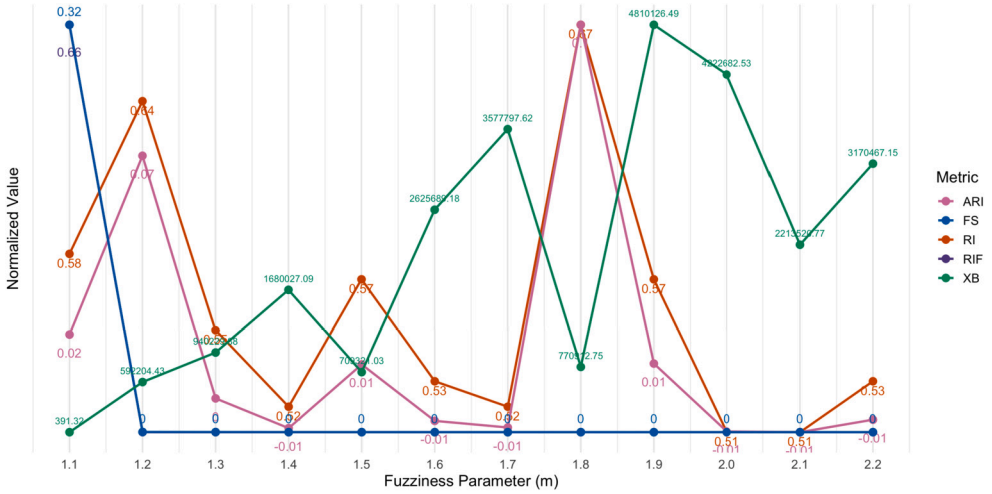

Fig. 13. The selection of m and comparison to the CVIs and clustering performances of subject 7.

Conversely, for subject 7, the proposed method exhibits a low clustering accuracy (0.58), with most samples classified as "drowsy." In total, 102 samples are collected for this subject. Figs. 13 and 14 show the selection of m and the corresponding membership matrix, respectively. Several factors may contribute to this outcome. First, EEG signals of this participant may lack clear-cut distinctions between drowsy and alert states, causing the algorithm to group most observations into a single cluster. Second, the data could be noisy or ambiguous, especially if the subject frequently occupies borderline states that do not neatly fit a strict binary label. Finally, the ground truth labels in this dataset are assigned using strict threshold-based rules on reaction times [75], which can mislabel borderline or transitional states by forcing them into one of two categories. Such rules may also misalign physiological responses with the assigned labels, particularly if reaction times do not accurately reflect true drowsiness. Consequently, the so-called "true" cluster may not fully capture the actual cognitive state.

Table 6 and Table 7 display the clustering accuracy (RI, ARI, and RIF) for each subject using the optimal fuzziness parameter m selected according to the two cluster validity indices, XB and FS, respectively. These tables allow a comparison of the performance resulting from different validity-based choices of m. In most cases, XB and FS lead to the same optimal value, suggesting consistency between the two internal indices. For subject 10, our FCPCA achieves 100% accuracy and delivers richer information than hard clustering methods, where one can observe the 'transition' state on some samples. Figures 15 to 32 in the Supplementary material show the results for the remaining subjects.

In short, by accommodating partial memberships, our fuzzy clustering approach provides a more nuanced view of driver states, capturing these transitional phases that hard clustering would overlook. This capability is especially valuable for designing adaptive driver-monitoring systems, as it can help identify early warning signs of drowsiness or distraction and allow for timely interventions.

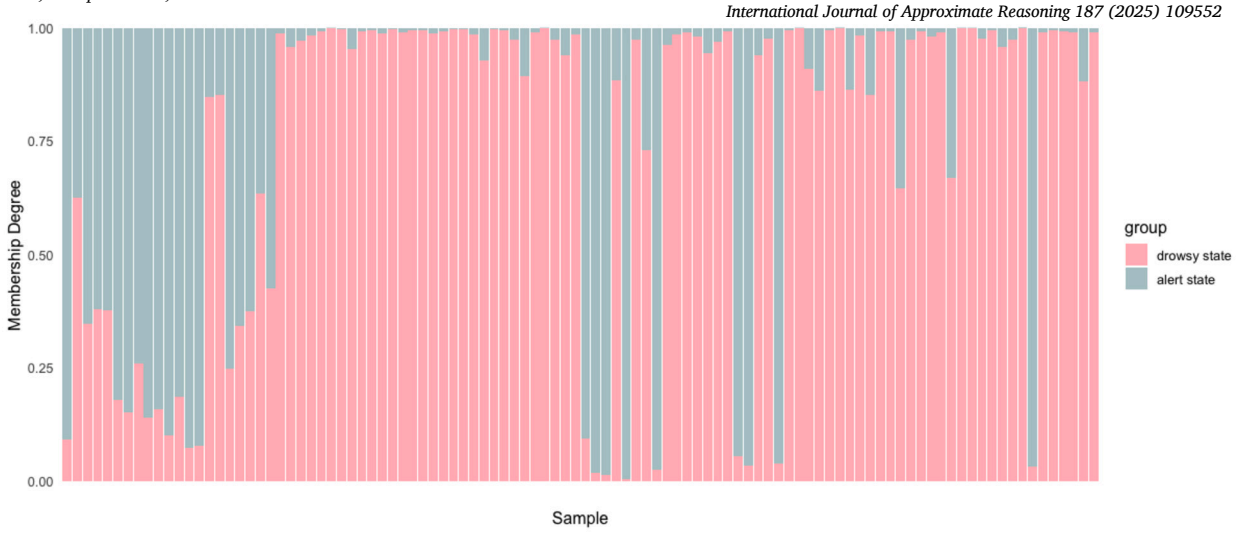

Fig. 14. Membership matrix of the samples of subject 7.

Table 6
Clustering accuracy using FCPCA with the optimal m selected by XB.

Subject #	Sample size	Optimal m	RI	ARI	RIF
1	188	1.5	0.74	0.24	0.51
2	132	1.1	0.89	0.62	0.82
3	150	1.1	0.55	0.01	0.89
4	148	1.1	0.57	0.02	0.58
5	224	1.1	0.88	0.56	0.74
6	166	1.2	0.51	0.00	0.52
7	102	1.1	0.58	0.02	0.66
8	264	1.1	0.66	0.09	0.61
9	314	1.1	0.75	0.26	0.84
10	108	1.1	1.00	1.00	0.92
11	226	1.1	0.93	0.75	0.86

Table 7
Clustering accuracy using FCPCA with the optimal m selected by FS.

Subject #	Sample size	Optimal m	RI	ARI	RIF
1	188	1.1	0.53	0.00	0.52
2	132	1.1	0.89	0.62	0.82
3	150	1.1	0.55	0.01	0.89
4	148	1.1	0.57	0.02	0.58
5	224	1.1	0.88	0.56	0.74
6	166	1.2	0.51	0.00	0.52
7	102	1.1	0.58	0.02	0.66
8	264	1.1	0.66	0.09	0.61
9	314	1.1	0.75	0.26	0.84
10	108	1.1	1.00	1.00	0.92
11	226	1.1	0.93	0.75	0.86

6. Discussion and future research

This work introduced FCPCA, a novel fuzzy clustering framework specifically designed for high-dimensional, variable-length MTS. By extending CPCA into a fuzzy clustering context, FCPCA directly addresses several long-standing challenges in time series clustering. Unlike traditional distance-based fuzzy clustering methods that often rely on dissimilarity measures (e.g., DTW distances) and struggle with the curse of dimensionality [23,34], FCPCA employs a reconstruction error criterion within a CPCA-based subspace. This shift from distance minimization to reconstruction minimization allows the algorithm to capture cross-channel lagged correlations while projecting data onto a low-dimensional common subspace, thereby maintaining efficiency even as the number of channels grows. As a result, the proposed method can handle moderately high-dimensional MTS (tens to hundreds of variables) and does not require all sequences to have the same length, a notable improvement over many previous approaches that assume common length or become computationally infeasible beyond even 10 dimensions. In addition, FCPCA introduces a data-driven procedure to tune the

fuzziness parameter m and the number of clusters, using suitable internal validity indices to select these parameters from the data. This automated selection addresses the common need to predefine fuzzy clustering parameters, thus enhancing objectivity and adaptability of the clustering process.

Empirically, FCPCA demonstrated competitive results in both synthetic simulations and real-world case studies. By incorporating fuzzy membership degrees, the method allows each time series to belong partially to multiple clusters, effectively capturing transitional or mixed states that hard clustering would overlook. This fuzzy assignment yields a richer, more nuanced representation of complex temporal phenomena. Our results showed that even a small amount of fuzziness can lead to more robust and accurate clustering outcomes: for example, using a low fuzziness value ($m = 1.1$) improved the clustering accuracy in several benchmark datasets compared to purely hard clustering. The iterative calculation of cluster-specific projection subspaces, which are weighted by current memberships, enables FCPCA to reveal subtle temporal relationships without forcing every series into a single cluster. Consequently, FCPCA not only achieves competitive or superior clustering accuracy relative to existing CPCA-based hard clustering, but also provides interpretable membership patterns that reflect underlying uncertainties and overlaps in the data.

A primary application highlighted in this study is the clustering of EEG data related to driver drowsiness. FCPCA proved capable of distinguishing alert versus drowsy brain states from 30-channel EEG segments while also identifying gradual transitional states between alertness and drowsiness. Such transitional or fuzzy states were clearly revealed by FCPCA but would have been lost under hard clustering. This finding is scientifically important. In neuroscience, cognitive states like alertness are not binary but change continuously, so capturing these in-between states provides a more comprehensive understanding of brain dynamics. Indeed, whereas most prior EEG clustering or classification efforts focus on improving accuracy for well-defined states (alert versus drowsy), our fuzzy clustering approach illuminates the continuum of vigilance levels. The ability to detect subtle, mixed brain states has practical implications: for instance, an adaptive driver monitoring system could use fuzzy membership outputs to warn of incipient drowsiness before a full transition to the drowsy state occurs, enabling earlier interventions. Beyond the EEG example, additional experiments on benchmark datasets indicate that FCPCA effectively captures nuanced overlapping dynamics in diverse settings. This suggests the broad applicability of the method.

Despite its contributions, FCPCA has certain limitations that must be acknowledged. First, the method relies on estimating covariance matrices (and lagged covariance blocks) to build the CPCA projection. This makes FCPCA somewhat sensitive to noise and outliers: if the time series data contain heavily corrupted, non-stationary, or noisy segments, the covariance estimates may be distorted, potentially degrading clustering performance. Even when using a reconstruction error criterion and weighted projection axes, extreme anomalies can still influence the common subspace derived for each cluster. Second, the theoretical convergence properties of the FCPCA algorithm are not yet fully characterized. While in practice the objective function (total weighted reconstruction error) tends to decrease monotonically with each iteration, formal guarantees, such as convergence to a stationary point, remain to be established. Third, the current implementation of FCPCA assumes that the lag structure (i.e., the set of time lags used for cross-covariance estimation) is predetermined. If the chosen lag order is suboptimal, the clustering may fail to capture important temporal dependencies or may include unnecessary information. Automatically selecting an optimal lag window remains an open issue in this study. Finally, computational complexity may become a bottleneck as the data scale increases. While we demonstrated feasibility on moderately high-dimensional datasets with reasonable runtime, very large-scale applications, such as those involving hundreds of thousands of variables or extremely long time series, may face challenges. In such scenarios, computing and decomposing large covariance matrices via SVD for each cluster can be computationally intensive. Addressing this would require further optimization or approximate techniques to maintain tractability.

Building on these insights, several directions for future research can be pursued to enhance and extend the FCPCA framework. (1) Robust extensions of FCPCA: A key priority is to improve the robustness of FCPCA against outliers. We aim to develop a robust variant of FCPCA that can effectively down-weight or trim anomalous series, leading to more reliable covariance estimates even when data are contaminated. (2) Integration with forecasting: Another potential direction is the integration of FCPCA into predictive modeling frameworks. The fuzzy clusters identified by FCPCA can serve as high-level features or states that enhance forecasting models. For instance, a clustering-based forecasting approach could first use FCPCA to recognize latent “regimes” in a MTS collection and then fit specialized forecasting models within each regime. Since FCPCA yields soft membership degrees, a time series could partially belong to multiple regimes, allowing a forecasting system to weigh predictions from different regime-specific models according to memberships (reflecting uncertainty during regime transitions). This could improve forecast accuracy, especially in systems that undergo gradual changes over time.

CRediT authorship contribution statement

Ziling Ma: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Ángel López-Oriona:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Methodology, Conceptualization. **Hernando Ombao:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Ying Sun:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Code availability

The FCPCA code and associated simulation studies are implemented in R and are freely available on the following GitHub repository: https://github.com/arbitraryma/FCPCA.git. For any questions or inquiries, please contact the corresponding author.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We sincerely thank the three referees for their careful review and valuable comments, which have greatly helped us improve the quality of our manuscript. This research was supported by King Abdullah University of Science and Technology (KAUST).

Appendix A. Supplementary material: Fuzziness parameter selection for each subject

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijar.2025.109552.

Data availability

The dataset used in this study is publicly available.

References

[1] M. Abavisani, V. Patel, Domain adaptive subspace clustering, in: BMVC, 2016.
[2] U. Ahmed, M. Nazir, A. Sarwar, T. Ali, E.-H.M. Aggoune, T. Shahzad, M.A. Khan, Signature-based intrusion detection using machine learning and deep learning approaches empowered with fuzzy clustering, Sci. Rep. 15 (2025) 1726.
[3] A. Alqahtani, M. Ali, X. Xie, M.W. Jones, Deep time-series clustering: a review, Electronics 10 (2021) 3001.
[4] M. Alruwaili, R. Alruwaili, U.A. Kumar, A.M. Albarrak, N.H. Ali, R. Basri, Human emotion recognition based on brain signal analysis using fuzzy neural network, Soft Comput. (2023) 1–15.
[5] A. Anuragi, D.S. Sisodia, R.B. Pachori, Mitigating the curse of dimensionality using feature projection techniques on electroencephalography datasets: an empirical review, Artif. Intell. Rev. 57 (2024) 75.
[6] P. Arabie, J.D. Carroll, W. DeSarbo, J. Wind, Overlapping clustering: a new method for product positioning, J. Mark. Res. 18 (1981) 310–317.
[7] A. Babii, E. Ghysels, J. Striaukas, Machine learning time series regressions with an application to nowcasting, J. Bus. Econ. Stat. 40 (2022) 1094–1106.
[8] J.C. Bezdek, R. Ehrlich, W. Full, Fcm: the fuzzy C-means clustering algorithm, Comput. Geosci. 10 (1984) 191–203.
[9] R.J. Campello, A fuzzy extension of the rand index and other related indexes for clustering and classification assessment, Pattern Recognit. Lett. 28 (2007) 833–841.
[10] R.J. Campello, E.R. Hruschka, A fuzzy extension of the silhouette width criterion for cluster analysis, Fuzzy Sets Syst. 157 (2006) 2858–2875.
[11] Y. Chen, Z. Wang, X. Bai, Fuzzy sparse subspace clustering for infrared image segmentation, IEEE Trans. Image Process. 32 (2023) 2132–2146.
[12] Y. Chen, S. Zhou, X. Zhang, D. Li, C. Fu, Improved fuzzy C-means clustering by varying the fuzziness parameter, Pattern Recognit. Lett. 157 (2022) 60–66.
[13] R. Coppi, P. D'Urso, P. Giordani, Fuzzy C-medoids clustering models for time-varying data, in: Modern Information Processing, Elsevier, 2006, pp. 195–206.
[14] J. Cui, Z. Lan, Y. Liu, R. Li, F. Li, O. Sourina, W. Müller-Wittig, A compact and interpretable convolutional neural network for cross-subject driver drowsiness detection from single-channel eeg, Methods 202 (2022) 173–184.
[15] A. Deng, B. Hooi, Graph neural network-based anomaly detection in multivariate time series, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, 2021, pp. 4027–4035.
[16] M.-C. Düker, D.S. Matteson, R.S. Tsay, I. Wilms, Vector autoregressive moving average models: a review, Wiley Interdiscip. Rev. Comput. Stat. 17 (2025) e70009.
[17] P. D'Urso, Fuzzy C-means clustering models for multivariate time-varying data: different approaches, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 12 (2004) 287–326.
[18] P. D'Urso, E.A. Maharaj, Wavelets-based clustering of multivariate time series, Fuzzy Sets Syst. 193 (2012) 33–61.
[19] P. D'Urso, L. De Giovanni, R. Massari, Trimmed fuzzy clustering of financial time series based on dynamic time warping, Ann. Oper. Res. 299 (2021) 1379–1395.
[20] P. D'Urso, L. De Giovanni, R. Massari, D. Di Lallo, Noise fuzzy clustering of time series by autoregressive metric, Metron 71 (2013) 217–243.
[21] P. D'Urso, L. De Giovanni, V. Vitale, Robust dtw-based entropy fuzzy clustering of time series, Ann. Oper. Res. (2023) 1–35.
[22] P. D'Urso, E.A. Maharaj, Autocorrelation-based fuzzy clustering of time series, Fuzzy Sets Syst. 160 (2009) 3565–3589.
[23] E. Egrioglu, C.H. Aladag, U. Yolcu, Fuzzy time series forecasting with a novel hybrid approach combining fuzzy C-means and neural networks, Expert Syst. Appl. 40 (2013) 854–857.
[24] A.E. Ezugwu, A.M. Ikotun, O.O. Oyelade, L. Abualigah, J.O. Agushaka, C.I. Eke, A.A. Akinyelu, A comprehensive survey of clustering algorithms: state-of-the-art machine learning applications, taxonomy, challenges, and future research prospects, Eng. Appl. Artif. Intell. 110 (2022) 104743.
[25] M.-J. Fadili, S. Ruan, D. Bloyet, B. Mazoyer, On the number of clusters and the fuzziness index for unsupervised fca application to bold fmri time series, Med. Image Anal. 5 (2001) 55–67.
[26] J.-Y. Franceschi, A. Dieuleveut, M. Jaggi, Unsupervised scalable representation learning for multivariate time series, Adv. Neural Inf. Process. Syst. 32 (2019).
[27] J. Frohlich, D. Toker, M.M. Monti, Consciousness among delta waves: a paradox?, Brain 144 (2021) 2257–2277.
[28] A.B. Geva, D.H. Kerem, Forecasting generalized epileptic seizures from the EEG signal by wavelet analysis and dynamic unsupervised fuzzy clustering, IEEE Trans. Biomed. Eng. 45 (1998) 1205–1216, PMID: 9775534.
[29] G. He, W. Jiang, R. Peng, M. Yin, M. Han, Soft subspace based ensemble clustering for multivariate time series data, IEEE Trans. Neural Netw. Learn. Syst. 34 (2022) 7761–7774.
[30] H. He, Y. Tan, Unsupervised classification of multivariate time series using vpca and fuzzy clustering with spatial weighted matrix distance, IEEE Trans. Cybern. 50 (2018) 1096–1105.
[31] Q. Huang, L. Shen, R. Zhang, S. Ding, B. Wang, Z. Zhou, Y. Wang, Crossgnn: confronting noisy multivariate time series via cross interaction refinement, Adv. Neural Inf. Process. Syst. 36 (2023) 46885–46902.
[32] L. Hubert, P. Arabie, Comparing partitions, J. Classif. 2 (1985) 193–218.
[33] H. Hwang, W.S. DeSarbo, Y. Takane, Fuzzy clusterwise generalized structured component analysis, Psychometrika 72 (2007) 181–198.
[34] H. Izakian, W. Pedrycz, I. Jamal, Fuzzy clustering of time series data using dynamic time warping distance, Eng. Appl. Artif. Intell. 39 (2015) 235–244.
[35] A. Javed, B.S. Lee, D.M. Rizzo, A benchmark study on time series clustering, Mach. Learn. Appl. 1 (2020) 100001.

[36] I. Khan, Z. Luo, J.Z. Huang, W. Shahzad, Variable weighting in fuzzy k-means clustering to determine the number of clusters, IEEE Trans. Knowl. Data Eng. 32 (2019) 1838–1853.
[37] R. Krishnapuram, A. Joshi, L. Yi, A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering, in: FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems, in: Conference Proceedings (Cat. No. 99CH36315), vol. 3, IEEE, 1999, pp. 1281–1286.
[38] B. Lafuente-Rego, P. D'Urso, J.A. Vilar, Robust fuzzy clustering based on quantile autocovariances, Stat. Pap. 61 (2020) 2393–2448.
[39] H. Li, Multivariate time series clustering based on common principal component analysis, Neurocomputing 349 (2019) 239–247.
[40] H. Li, M. Wei, Fuzzy clustering based on feature weights for multivariate time series, Knowl.-Based Syst. 197 (2020) 105907.
[41] A. Lopez-Oriona, A. Vilar J, mlmts: Machine Learning Algorithms for Multivariate Time Series, 2023, R package version 1.1.1.
[42] Á. López-Oriona, P. D'Urso, J.A. Vilar, B. Lafuente-Rego, Quantile-based fuzzy C-means clustering of multivariate time series: robust techniques, Int. J. Approx. Reason. 150 (2022) 55–82.
[43] Á. López-Oriona, P. D'Urso, J.A. Vilar, B. Lafuente-Rego, Spatial weighted robust clustering of multivariate time series based on quantile dependence with an application to mobility during covid-19 pandemic, IEEE Trans. Fuzzy Syst. 30 (2022) 3990–4004.
[44] Á. López-Oriona, J.A. Vilar, P. D'Urso, Quantile-based fuzzy clustering of multivariate time series in the frequency domain, Fuzzy Sets Syst. 443 (2022) 115–154.
[45] Z. Ma, A. Lopez Oriona, H. Ombao, Y. Sun, ROBCPCA: a robust multivariate time series clustering method based on common principal component analysis, 2024.
[46] E.A. Maharaj, P. D'Urso, Fuzzy clustering of time series in the frequency domain, Inf. Sci. 181 (2011) 1187–1211.
[47] E.A. Maharaj, P. D'Urso, D.U. Galagedera, Wavelet-based fuzzy clustering of time series, J. Classif. 27 (2010) 231–275.
[48] M.R. Mahmoudi, D. Baleanu, Z. Mansor, B.A. Tuan, K.-H. Pho, Fuzzy clustering method to compare the spread rate of covid-19 in the high risks countries, Chaos Solitons Fractals 140 (2020) 110230.
[49] A.S. Malik, H.U. Amin, Designing EEG Experiments for Studying the Brain: Design Code and Example Datasets, Academic Press, 2017.
[50] G. Marti, F. Nielsen, M. Bińkowski, P. Donnat, A review of two decades of correlations, hierarchies, networks and clustering in financial markets, Prog. Inf. Geom. Theory Appl. (2021) 245–274.
[51] P. Masulli, F. Masulli, S. Rovetta, A. Lintas, A.E. Villa, Fuzzy clustering for exploratory analysis of eeg event-related potentials, IEEE Trans. Fuzzy Syst. 28 (2019) 28–38.
[52] P.K. Mishro, S. Agrawal, R. Panda, A. Abraham, A novel type-2 fuzzy C-means clustering for brain mr image segmentation, IEEE Trans. Cybern. 51 (2020) 3901–3912.
[53] C.S. Möller-Levet, F. Klawonn, K.-H. Cho, H. Yin, O. Wolkenhauer, Clustering of unevenly sampled gene expression time-series data, Fuzzy Sets Syst. 152 (2005) 49–66.
[54] H. Ombao, M. Pinto, Spectral dependence, Econom. Stat. 32 (2024) 122–159.
[55] N.R. Pal, J.C. Bezdek, On cluster validity for the fuzzy C-means model, IEEE Trans. Fuzzy Syst. 3 (2002) 370–379.
[56] N.U. Patel, Comparative study between fuzzy clustering and hard clustering, Int. J. Futur. Trends Eng. Tech. 1 (2014) 16–19.
[57] W. Pedrycz, H. Izakian, Cluster-centric fuzzy modeling, IEEE Trans. Fuzzy Syst. 22 (2014) 1585–1597.
[58] C. Peng, Q. Zhang, Z. Kang, C. Chen, Q. Cheng, Kernel two-dimensional ridge regression for subspace clustering, Pattern Recognit. 113 (2021) 107749.
[59] J. Rabcan, V. Levashenko, E. Zaitseva, M. Kvassay, Eeg signal classification based on fuzzy classifiers, IEEE Trans. Ind. Inform. 18 (2021) 757–766.
[60] A. Rafiei, R. Zahedifar, C. Sitaula, F. Marzbanrad, Automated detection of major depressive disorder with eeg signals: a time series classification using deep learning, IEEE Access 10 (2022) 73804–73817.
[61] W.M. Rand, Objective criteria for the evaluation of clustering methods, J. Am. Stat. Assoc. 66 (1971) 846–850.
[62] S. Rao, R. Tron, R. Vidal, Y. Ma, Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2009) 1832–1845.
[63] A.P. Ruiz, M. Flynn, J. Large, M. Middlehurst, A. Bagnall, The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances, Data Min. Knowl. Discov. 35 (2021) 401–449.
[64] E.H. Ruspini, J.C. Bezdek, J.M. Keller, Fuzzy clustering: a historical perspective, IEEE Comput. Intell. Mag. 14 (2019) 45–55.
[65] S. Sanei, J.A. Chambers, EEG Signal Processing, John Wiley & Sons, 2013.
[66] A. Sarda-Espinosa, dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance, 2024, R package version 6.0.0.
[67] R.H. Shumway, D.S. Stoffer, D.S. Stoffer, Time Series Analysis and Its Applications, vol. 3, Springer, 2000.
[68] X. Si, Q. Yin, X. Zhao, L. Yao, Consistent and diverse multi-view subspace clustering with structure constraint, Pattern Recognit. 121 (2022) 108196.
[69] C. Subbalakshmi, G.R. Krishna, S.K.M. Rao, P.V. Rao, A method to find optimum number of clusters based on fuzzy silhouette on dynamic data set, Proc. Comput. Sci. 46 (2015) 346–353.
[70] R. Vidal, Y. Ma, S. Sastry, Generalized principal component analysis (gpca), IEEE Trans. Pattern Anal. Mach. Intell. 27 (2005) 1945–1959.
[71] J.A. Vilar, B. Lafuente-Rego, P. D'Urso, Quantile autocovariances: a powerful tool for hard and soft partitional clustering of time series, Fuzzy Sets Syst. 340 (2018) 38–72.
[72] H. Wang, F. Yu, Y. Tang, C. Wang, Linear fuzzy information granule and soft subspace based multivariate time series clustering, in: 2024 20th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), IEEE, 2024, pp. 1–7.
[73] X. Wang, A. Wirth, L. Wang, Structure-based statistical features and multivariate time series clustering, in: Seventh IEEE International Conference on Data Mining (ICDM 2007), IEEE, 2007, pp. 351–360.
[74] M. Wedel, J.-B.E. Steenkamp, A fuzzy clusterwise regression approach to benefit segmentation, Int. J. Res. Mark. 6 (1989) 241–258.
[75] C.-S. Wei, Y.-T. Wang, C.-T. Lin, T.-P. Jung, Toward drowsiness detection using non-hair-bearing EEG-based brain-computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (2018) 400–406.
[76] K.-L. Wu, Analysis of parameter selections for fuzzy C-means, Pattern Recognit. 45 (2012) 407–415.
[77] X.L. Xie, G. Beni, A validity measure for fuzzy clustering, IEEE Trans. Pattern Anal. Mach. Intell. 13 (1991) 841–847.
[78] X. Xu, L. Zhang, L.-F. Cheong, Z. Li, C. Zhu, Learning clustering for motion segmentation, IEEE Trans. Circuits Syst. Video Technol. 32 (2021) 908–919.
[79] R. Yan, J. Liao, J. Yang, W. Sun, M. Nong, F. Li, Multi-hour and multi-site air quality index forecasting in Beijing using cnn, lstm, cnn-lstm, and spatiotemporal clustering, Expert Syst. Appl. 169 (2021) 114513.
[80] M.-S. Yang, K.-L. Wu, J.-N. Hsieh, J. Yu, Alpha-cut implemented fuzzy clustering algorithms and switching regressions, IEEE Trans. Syst. Man Cybern., Part B, Cybern. 38 (2008) 588–603.
[81] H. Yu, X. Lei, Z. Song, C. Liu, J. Wang, Supervised network-based fuzzy learning of eeg signals for Alzheimer's disease identification, IEEE Trans. Fuzzy Syst. 28 (2019) 60–71.
[82] H. Zhang, H. Ji, J. Yu, J. Li, L. Jin, L. Liu, Z. Bai, C. Ye, Subject-independent EEG classification based on a hybrid neural network, Front. Neurosci. 17 (2023) 1124089.
[83] T. Zhang, A. Szlam, G. Lerman, Median k-flats for hybrid linear modeling with many outliers, in: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, IEEE, 2009, pp. 234–241.
[84] Y. Zhang, W. Wang, X. Zhang, Y. Li, A cluster validity index for fuzzy clustering, Inf. Sci. 178 (2008) 1205–1218.
[85] Z. Zhang, L. Meng, Y. Gu, Sageformer: series-aware framework for long-term multivariate time-series forecasting, IEEE Internet Things J. 11 (2024) 18435–18448.
[86] L. Zhu, F.-L. Chung, S. Wang, Generalized fuzzy C-means clustering algorithm with improved fuzzy partitions, IEEE Trans. Syst. Man Cybern., Part B, Cybern. 39 (2009) 578–591.
[87] J. Zhuang, L. Cui, T. Qu, C. Ren, J. Xu, T. Li, G. Tian, J. Yang, A streamlined scrna-seq data analysis framework based on improved sparse subspace clustering, IEEE Access 9 (2021) 9719–9727.