Functional Singular Spectrum Analysis

Hossein Haghbin
Persian Gulf University

Seyed Morteza Najibi
Shiraz University

Rahim Mahmoudvand
Bu-Ali Sina University

Jordan Trinka
Marquette University

Mehdi Maadooliat
Marquette University, mehdi.maadooliat@marquette.edu

Follow this and additional works at: https://epublications.marquette.edu/math_fac

Part of the Mathematics Commons

Recommended Citation
Haghbin, Hossein; Najibi, Seyed Morteza; Mahmoudvand, Rahim; Trinka, Jordan; and Maadooliat, Mehdi, "Functional Singular Spectrum Analysis" (2021). Mathematical and Statistical Science Faculty Research and Publications. 111.
https://epublications.marquette.edu/math_fac/111
Functional Singular Spectrum Analysis

Hossein Haghbin  
Persian Gulf University  
Seyed Morteza Najibi  
Lund University  
Rahim Mahmoudvand  
Bu-Ali Sina University  
Jordan Tinka  
Marquette University  
Mehdi Maadooliat  
Marquette University

Abstract

In this paper, we develop a new extension of the singular spectrum analysis (SSA) called functional SSA to analyze functional time series. The new methodology is constructed by integrating ideas from functional data analysis and univariate SSA. Specifically, we introduce a trajectory operator in the functional world, which is equivalent to the trajectory matrix in the regular SSA. In the regular SSA, one needs to obtain the singular value decomposition (SVD) of the trajectory matrix to decompose a given time series. Since there is no procedure to extract the functional SVD (fSVD) of the trajectory operator, we introduce a computationally tractable algorithm to obtain the fSVD components. The effectiveness of the proposed approach is illustrated by an interesting
example of remote sensing data. Also, we develop an efficient and user-friendly R package and a shiny web application to allow interactive exploration of the results.

Abbreviations

dFPCA: dynamic FPCA
FDA: functional data analysis
FPCA: functional PCA
FSSA: functional SSA
fSVD: functional SVD
FTS: functional time series
MSSA: multivariate SSA
NDVI: normalized difference vegetation index
PCA: principal component analysis
SSA: singular spectrum analysis
SVD: singular value decomposition.

1 INTRODUCTION

One of the popular approaches in the decomposition of time series is accomplished using the rates of change. In this approach, the observed time series is partitioned (decomposed) into informative trends plus potential seasonal (cyclical) and noise (irregular) components. Aligned with this principle, singular spectrum analysis (SSA) is a model-free procedure that is commonly used as a non-parametric technique in analyzing the time series. SSA is intrinsically motivated as an exploratory and model building tool rather than a confirmatory procedure (Golyandina, Nekrutkin, & Zhigljavsky, 2001). SSA does not require restrictive assumptions such as stationarity, linearity, and normality. It can be used for a wide range of purposes such as trend and periodic component detection and extraction, smoothing, forecasting, change-point detection, gap filling, and causality (see, e.g., Golyandina & Osipov, 2007; Golyandina et al. 2001; Kondrashov, Shprits, & Ghil, 2010; Mahmoudvand & Rodrigues, 2016; Mohammad & Nishida, 2011; Moskvina & Zhigljavsky, 2003; Rodrigues & Mahmoudvand, 2016).

The implementation of SSA over time series is similar to that of principal component analysis (PCA) of multivariate data. In contrast to PCA, which is applied to a data matrix with independent rows, SSA is applied to a time series. It provides a representation of the given time series in terms of eigentriples of a so-called trajectory matrix (Alexandrov, 2009).

Up to this day, many studies have been published with extensions and applications of SSA. Extension to multivariate SSA (MSSA) is straightforward (Golyandina & Zhigljavsky, 2013; Golyandina, Korobeynikov, & Zhigljavsky, 2018; Golyandina, Korobeynikov, Shlemov, & Usevich, 2015; Hassani & Mahmoudvand, 2013, 2018), and an extension of SSA to a two-dimensional setting can be found in Golyandina and Usevich (2010), Golyandina et al. (2015), and references therein. In the regular SSA, we assume that the observation at each time point is scalar, vector, or array. As a matter of interest, one may consider a series of curves observed over time and use the basics of Hilbert space in the functional data analysis (FDA) framework to introduce the concept of functional SSA (FSSA). Specifically, in this work, we introduce the FSSA for a time series of functions, where each function can be represented using basis expansion models (e.g., B-spline or Fourier).

While the research in FDA has grown extensively in recent years, there have been relatively few contributions dealing with functional time series (FTS) (see, e.g., Bosq, 2000; Hörmann & Kokoszka, 2012). Although most of the current FTS approaches focus on a parametric fit for inferences and forecasting, there exist some other approaches in the literature that extend functional PCA (FPCA) to incorporate the temporal correlation of FTS.
For instance, Hörmann, Kidziński, and Hallin (2015) introduced dynamic FPCA (dFPCA) to analyze FTS. This approach assumes the strong assumption of stationarity, which is not generally held in practice. It would be of interest to non-parametrically decompose a nonstationary FTS to reveal the respective trends plus seasonal and irregular components in an appropriate manner. Consistent with this approach, and as a first step, Fraiman, Justel, Liu, and Llop (2014) introduced a new concept of trends for the FTS. Furthermore, Hörmann, Kokoszka, and Nisol (2018) considered the periodic components for the FTS and derived several procedures to test the periodicity using frequency domain analysis. To the best of our knowledge, existing studies mainly focus on detecting rather than extracting interpretable components.

Since one of the primary missions of SSA is to extract trends and periodic components of a regular (non-functional) time series, it would be rational to establish a similar elegant non-parametric procedure to extract such components for FTS. In this paper, we use the basics of univariate SSA and FDA to develop FSSA. In a nutshell, we develop a matrix operator with functional entities and provide a procedure to obtain the functional singular value decomposition (fSVD) of this operator, which can be seen as a generalization of the regular SVD used in the SSA literature. The new methodology, FSSA, not only can serve as a non-parametric dimension reduction tool to decompose the FTS, it can also be used as a visualization tool to illustrate the concept of seasonality and periodicity in the functional space over time.

In order to depict the idea of our approach and to show its utility, consider the following motivating example involving a real dataset, which is described in detail in the Supporting Information. These data provide the intraday number of calls to a call center, during different times of the days for 1 year. The associated 365 curves are represented in an overlapping pattern in Figure 1 (left). In Figure 1 (right), we investigate the pattern among weekdays and weekend days. As we can see, the intraday patterns of weekends (Friday and Saturday) are significantly different from workdays, while workdays seem to have similar patterns. Investigators used variants of FPCA to analyze the call center data in literature (Huang, Shen, & Buja, 2008; Maadooliat, Huang, & Hu, 2015; Shen & Huang, 2005). For illustration purpose, we compare the results of the proposed method (FSSA) and dFPCA on this dataset. Figure 2 (top) presents the projection of the data into the first four FPCs of the dFPCA obtained from the freqdom.fda package in R (Hörmann et al. 2015). We used seven different colors to differentiate between different days of a week. As one may observe, visually, there is no clear separation in either one of the FPC graphs in the top row. This may not be surprising, as the purpose of PCA of any type is to reduce the dimensionality and not necessarily to decompose the data into regular trends and periodic and irregular components. In contrast, the grouping results that we obtained using the FSSA on the call center data are given in Figure 2 (middle). We also present the result of MSSA obtained from the Rssa package (Golyandina et al. 2015) in Figure 2 (middle). It can be seen that the functional behavior of 7 days of a week can be well distinguished, visually, using either one of the last two groups (Groups 3 and 4) of the FSSA.

FIGURE 1 The number of calls to a call center between January 1 to December 31 in the year 1999
The rest of the paper is organized as follows. Section 2 reviews the core of SSA for completeness. Section 3 presents the theoretical foundations and some properties of the proposed method (FSSA), and Section 4 provides implementation details. Section 5.1 reports simulation results to illustrate the use of the proposed approach in analyzing FTS and to compare it with MSSA and dFPCA. Application to a real data example on remote sensing is given in Section 5.2. Section 6 provides some discussions and concluding remarks.

2 GENERAL SCHEME OF SSA

As we mentioned in Section 1, SSA can be used for many purposes. However, as we intend to introduce the functional version of SSA for decomposing FTS, we review a general scheme of SSA to perform time series decomposition in this section.

2.1 Univariate SSA

Throughout this section, we consider that $y_i$,s are elements of Euclidean space $\mathbb{R}$. Suppose that $y_N = (y_1, y_2, ..., y_N)^T$ is a realization of size $N$ from a time series. The basic SSA algorithm consists of four steps: embedding, decomposition, grouping, and reconstruction.

2.1.1 Step 1. Embedding

This step generates a multivariate object by tracking a moving window of size $L$ over the original time series, where $L$ is called window length parameter and $1 < L < N/2$. Embedding can be regarded as a mapping operator $T$ that transfers the series $y_N$ into a so-called trajectory matrix $X$ of dimension $L \times K$, defined by

$$X = T(y_N) = [x_1, ..., x_K],$$

(1)

where $K = N - L + 1$ and $x_j = (y_j, y_{j+1}, ..., y_{j+L-1})^T$, for $j = 1, ..., K$, are called lagged vectors. Note that the trajectory matrix $X$ is a Hankel matrix, which means that all the elements along the antidiagonals are equal. Indeed, the embedding operator $T$ is a one-to-one mapping from $\mathbb{R}^N$ into $\mathbb{R}^{L \times K} \subseteq \mathbb{R}^{L \times K}$, where $\mathbb{R}^{L \times K}$ is the set of all $L \times K$ Hankel matrices.

2.1.2 Step 2. Decomposition

In this step, the singular value decomposition (SVD) for the trajectory matrix is computed as
\[
X = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i v_i^T = \sum_{i=1}^{r} X_i,
\]

(2)

where \( r \) is the rank of the matrix \( X \), \( \sqrt{\lambda_i} \) is the \( i \)th singular value of \( X \), \( u_i \) and \( v_i \) are the associated (orthonormal) left and right singular vectors, and \( X_i = \sqrt{\lambda_i} u_i v_i^T \) is called the respective elementary matrix.

2.1.3 Step 3. Grouping
Consider a partition of the set of indices \( \{1, 2, \ldots, r\} \) into \( m \) disjoint subsets \( \{I_1, I_2, \ldots, I_m\} \). For any positive integer \( q \), that is, \( 1 \leq q \leq m \), the matrix \( X_{I_q} \) is defined as \( X_{I_q} = \sum_{i \in I_q} X_i \). Thus, by the expansion (2), we have the grouped matrix decomposition

\[
X = X_{I_1} + X_{I_2} + \cdots + X_{I_m}.
\]

(3)

Each group in (3) should correspond to a component in a time series decomposition. These components can be considered as trend, cycle, seasonal, noise, and so forth.

2.1.4 Step 4. Reconstruction
Finally, the resulting matrices \( X_{I_q} \) in (3) are transformed back into the form of the original series \( y_N \) by an inverse operator \( J^{-1} \). In order to do this, first, it is necessary that each matrix \( X_{I_q} \) be approximated by a matrix in \( \mathbb{R}^{H \times K} \). This approximation is performed optimally in the sense of orthogonal projection of \( X_{I_q} \) on \( \mathbb{R}^{H \times K} \) with respect to the Frobenius norm. Denote this projection by \( \Pi: \mathbb{R}^{L \times K} \to \mathbb{R}^{H \times K} \). It is shown that the projection \( \Pi \) is the averaging of the matrix elements over the antidiagonals (where \( i + j = \text{const} \)). By combining the results of this step and (3), we obtain the final decomposition of the series in the form of

\[
y_N = \bar{y}_1 + \bar{y}_2 + \cdots + \bar{y}_m,
\]

(4)

where \( \bar{y}_q = J^{-1} \Pi(X_{I_q}) \), for \( q = 1, \ldots, m \).

The above algorithm can be extended to perform MSSA for analyzing multivariate time series. The only difference is in defining the trajectory matrix, which can be defined by stacking univariate trajectory matrices horizontally or vertically (Hassani & Mahmoudvand, 2018).

It is well known that SSA does not require restrictive assumptions; however, it is ideal to have a time series with separable components. Therefore, we present tools to measure the separability of components in the next subsection.

2.2 Separability
Let \( y_N^{(i)} = (y_1^{(i)}, \ldots, y_N^{(i)})^T \) for \( i = 1, 2 \) be two time series and consider an additive model as \( y_N = y_N^{(1)} + y_N^{(2)} \). The series \( y_N^{(1)} \) and \( y_N^{(2)} \) are called separable when each lagged vector of \( y_N^{(1)} \) is orthogonal to the lagged vectors of \( y_N^{(2)} \). To measure the degree of separability between two time series \( y_N^{(1)} \) and \( y_N^{(2)} \), Golyandina et al. (2001) introduced the so-called \( w \)-correlation
\[
\rho^{(w)}(y_N^{(1)}, y_N^{(2)}) = \frac{\left\langle y_N^{(1)}, y_N^{(2)} \right\rangle_w}{\sqrt{\left\langle y_N^{(1)}, y_N^{(1)} \right\rangle_w \left\langle y_N^{(2)}, y_N^{(2)} \right\rangle_w}}
\]

(5)

where \( \left\langle y_N^{(t)}, y_N^{(s)} \right\rangle_w = \sum_{i=1}^{N} w_i y_i^{(t)} y_i^{(s)} \) for \( t, s = 1,2 \) and \( w_i = \min\{i, N - i + 1\} \). We call two series \( y_N^{(1)} \) and \( y_N^{(2)} \) \( w \)-orthogonal if \( \rho^{(w)}(y_N^{(1)}, y_N^{(2)}) = 0 \) for appropriate values of \( L \) (see the next subsection for more details).

Note that \( \tilde{y}_q, q = 1, \ldots, m, \) is the reconstructed component produced by the group \( l_q \), and the matrix of \( \rho^{(w)} = \left\{ \rho^{(w)}(\tilde{y}_i, \tilde{y}_j) \right\}_{i,j=1}^{m} \) is called a \( w \)-correlation matrix.

2.3 Parameter selection

There are two basic parameters in the SSA procedure, window length \( L \) and grouping parameters. Choosing improper values for these parameters yields an incomplete reconstruction and misleading results in subsequent analysis. In spite of the importance of choosing \( L \) and grouping parameters for SSA, no ideal solution has been yet proposed. A thorough review of the problem shows that the optimal choice of the parameters depends on the intrinsic structure of the data and the purposes of the study (Golyandina & Zhigljavsky, 2013; Golyandina et al. 2001). However, there are several recommendations and rules that work well for a wide range of scenarios. It is recommended to select the window length parameter, \( L \), to be a large integer that is a multiple of the periodicities of the time series, but not larger than \( \frac{N}{2} \).

In addition, there are several utilities for effective grouping. These tools include analyzing the periodogram, paired plot of the singular vectors, scree plot of the singular values, and \( w \)-correlation plot (see Golyandina et al. 2001 for more details).

3 THEORETICAL FOUNDATIONS OF FSSA

We start this section with the mathematical foundations that are used to develop the fSSA procedure. Hereafter, we consider \( y_N = (y_1, \ldots, y_N)^T \) as an FTS of length \( N \). This means that each element \( y_i: [0,1] \to \mathbb{R} \) belongs to \( \mathbb{H}: = L^2([0,1]) \), the space of square integrable real functions defined on the interval \([0,1]\). Here, the space \( \mathbb{H} \) is a Hilbert space, equipped with inner product \( \langle x, y \rangle = \int_0^1 x(s) y(s) \, ds \). For a given positive integer \( k \), the space \( \mathbb{H}^k \) denotes the Cartesian product of \( k \) copies of \( \mathbb{H} \); that is, for an element \( x \in \mathbb{H}^k \), it has the form \( x(s) = (x_1(s_1), x_2(s_2), \ldots, x_k(s_k))^T \), where \( x_i \in \mathbb{H} \), and \( s = (s_1, s_2, \ldots, s_k) \in [0,1]^k \). Then \( \mathbb{H}^k \) is a Hilbert space equipped with the inner product \( \langle x, y \rangle_{\mathbb{H}^k} = \sum_{i=1}^{k} \langle x_i, y_i \rangle \). The norms will be denoted by \( \| \cdot \| \) and \( \| \cdot \|_{\mathbb{H}^k} \) in the spaces \( \mathbb{H} \) and \( \mathbb{H}^k \), respectively. For \( x \in \mathbb{H}_1 \), and \( y \in \mathbb{H}_2 \), where \( \mathbb{H}_1 \) and \( \mathbb{H}_2 \) are two Hilbert spaces, we define the tensor (outer) product corresponding to the operator \( x \otimes y: \cdot \to \mathbb{H}_1 \to \mathbb{H}_2 \) as \( (x \otimes y)h = (x,h)y \), where \( h \in \mathbb{H}_1 \).

For positive integers \( L \) and \( K \), we denote \( \mathbb{H}^{L \times K} \) as the linear space spanned by operators \( Z: \cdot \to \mathbb{H}^K \to \mathbb{H}^L \), specified by \( [Z_{i,j}]_{i=1,...,L,j=1,...,K} \), where
\[
Z \mathbf{a} = \begin{pmatrix}
\sum_{j=1}^{K} a_j z_{1,j} \\
\vdots \\
\sum_{j=1}^{K} a_j z_{L,j}
\end{pmatrix}, z_{i,j} \in \mathbb{H}, \text{and } \mathbf{a} = (a_1, \ldots, a_K) \in \mathbb{R}^K.
\]
(6)
We call an operator \( \hat{\mathcal{Z}} = [\hat{z}_{i,j}] \in \mathbb{H}^{L \times K} \) Hankel if \( \| \hat{z}_{i,j} - g_s \| = 0 \) for some \( g_s \in \mathbb{H} \) where \( s = i + j \). The space of such Hankel operators will be denoted by \( \mathbb{H}^{L \times K}_H \). For two given operators \( \mathcal{Z}_1 = [z_{i,j}^{(1)}] \) and \( \mathcal{Z}_2 = [z_{i,j}^{(2)}] \) in \( \mathbb{H}^{L \times K}_H \), define
\[
\langle \mathcal{Z}_1, \mathcal{Z}_2 \rangle_F := \sum_{i=1}^{L} \sum_{j=1}^{K} \langle z_{i,j}^{(1)}, z_{i,j}^{(2)} \rangle.
\]
(7)
It follows immediately that \( \langle \cdot, \cdot \rangle_F \) defines an inner product on \( \mathbb{H}^{L \times K}_H \). We will call it the Frobenius inner product of two operators in \( \mathbb{H}^{L \times K}_H \). The associated Frobenius norm is \( \| Z \|_F = \sqrt{\langle Z, Z \rangle_F} \). Before discussing the FSSA algorithm, here we present a lemma that will be used in the last step of the proposed algorithm. We offer proofs of all lemmas, propositions, and theorems in the supporting information.

**Lemma 1.** Let \( x_i, i = 1, \ldots, N \) be elements of the Hilbert space \( \mathbb{H} \). If \( \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \), then
\[
\sum_{i=1}^{N} \| x_i - \bar{x} \|^2 \leq \sum_{i=1}^{N} \| x_i - y \|^2,
\]
for all \( y \in \mathbb{H} \).

3.1 FSSA algorithm
For an integer \( 1 < L < N/2 \), let \( K = N - L + 1 \) and define a set of multivariate functional vectors in \( \mathbb{H}^{L} \) by
\[
\mathbf{x}_j := (y_j, y_{j+1}, \ldots, y_{j+L-1})^T, j = 1, \ldots, K,
\]
(7)
where the \( \mathbf{x}_j \)s denote the functional \( L \)-lagged vectors. The following algorithm provides the FSSA results in four steps.

3.1.1 Step 1. Embedding
Define the operator \( \mathcal{X}: \mathbb{R}^K \to \mathbb{H}^L \) by
\[
\mathcal{X} \mathbf{a} := \sum_{j=1}^{K} a_j \mathbf{x}_j, a = (a_1, \ldots, a_K)^T \in \mathbb{R}^K.
\]
(9)
Proposition 1. The operator $\mathcal{X}$ is a bounded linear operator. If we define $\mathcal{X}^*: \mathbb{H}_L \rightarrow \mathbb{R}_K$, given by

$$\mathcal{X}^* z = (\langle x_1, z \rangle_{\mathbb{H}_L}, \langle x_2, z \rangle_{\mathbb{H}_L}, ..., \langle x_K, z \rangle_{\mathbb{H}_L})^T, z \in \mathbb{H}_L,$$

(10)

then $\mathcal{X}^*$ is an adjoint operator for $\mathcal{X}$.

3.1.2 Step 2. Decomposition

In this step, we decompose the trajectory operator $\mathcal{X}$ into a set of rank one operators. To this end, we denote the range of $\mathcal{X}$ by $R(\mathcal{X})$. Clearly, $R(\mathcal{X}) = \text{span}\{x_i\}_{i=1}^K$ is $r$-dimensional, where $0 < r \leq K$. Therefore, $\mathcal{X}$ is a finite-rank ($r$-dimensional) operator.

Theorem 1. Consider the trajectory operator $\mathcal{X}$ in (8). There exist linearly independent elements $\psi_1, ..., \psi_r$ from $\mathbb{H}_L$ and $v_1, ..., v_r$ from $\mathbb{R}_K$ that are orthonormal and

$$\mathcal{X}a = \sum_{i=1}^r \sqrt{\lambda_i} (v_i, a)_{\mathbb{R}_K} \psi_i, \text{ for all } a \in \mathbb{R}_K,$$

(11)

where the $\lambda_i$s are non-ascending positive scalars. Then

$$\mathcal{X}^* z = \sum_{i=1}^r \sqrt{\lambda_i} \langle \psi_i, z \rangle_{\mathbb{H}_L} v_i, \text{ for all } z \in \mathbb{H}_L.$$

(12)

Proposition 2. In Theorem 1, the set $\{\psi_i\}_{i=1}^r$ forms a basis system for $R(\mathcal{X})$, and each $v_i$ can be written as

$$v_i = \frac{\mathcal{X}^* \psi_i}{\sqrt{\lambda_i}} = \left(\frac{\langle \psi_i, x_1 \rangle_{\mathbb{H}_L}}{\sqrt{\lambda_i}}, ..., \frac{\langle \psi_i, x_K \rangle_{\mathbb{H}_L}}{\sqrt{\lambda_i}}\right)^T, i = 1, \ldots, K.$$

(13)

One may consider Theorem 1 as the fSVD of $\mathcal{X}$. We shall call $\sqrt{\lambda_i}$ the singular value, $\psi_i$ the left singular function, and $v_i$ the right singular vector of the trajectory operator. Therefore, the collection $\left(\sqrt{\lambda_i}, \psi_i, v_i\right)$ can be seen as the $i$th eigentriple of $\mathcal{X}$, and the right singular vectors, the $v_i$s, can be used to produce paired plots similar to the ones in the SSA literature (Golyandina et al. 2001). Now define $\mathcal{X}_i = \sqrt{\lambda_i} v_i \otimes \psi_i$, for $i = 1, \ldots, r$. One can use (11) and decompose $\mathcal{X}$ as

$$\mathcal{X} = \sum_{i=1}^r \mathcal{X}_i.$$

(14)

Clearly, the $\mathcal{X}_i$s are rank one operators. We refer to the $\mathcal{X}_i$s as elementary operators.

3.1.3 Step 3. Grouping

The grouping step is the procedure of rearranging and partitioning the elementary operators $\mathcal{X}_i$ in (14). To do this, we mimic the approaches used in Step 3 of Section 2 for the univariate SSA and implement the equivalent functional version of those in Haghbin, Najibi, Trinka, and Maadooliat (2019). Note that, in practice, we use a
finite set of elementary operators, and one can consider a partition \( \{ I_1, I_2, \ldots, I_m \} \) of the set of indices such that we have the expansion

\[
X = X_{I_1} + X_{I_2} + \cdots + X_{I_m}.
\]  

(15)

3.1.4 Step 4. Reconstruction

At this step, for any given \( q (1 \leq q \leq m) \), we would like to use \( T^{-1} : \mathbb{H}_L^L \rightarrow \mathbb{H}^N \) to transform back each operator \( X_{I_q} \) in (15) to \( Y_q \) and hence construct a functional version of the decomposition given in (4). But since \( X_{I_q} \in \mathbb{H}_L \), first, it is necessary to project \( X_{I_q} \) to \( \mathbb{H}_L \). Note that \( \mathbb{H}_L \) is a closed subspace of \( \mathbb{H}^N \), therefore by the projection theorem, there exists a unique \( \hat{X}_{I_q} \in \mathbb{H}_L \) such that

\[
\| X_{I_q} - \hat{X}_{I_q} \|_F^2 \leq \| X_{I_q} - Z \|_F^2, \text{ for any } Z \in \mathbb{H}_L.
\]

To specify \( \hat{X}_{I_q} \), we denote the elements of \( X_{I_q} \) and \( \hat{X}_{I_q} \) by \( x_{i,j} \) and \( \hat{x}_{i,j} \), respectively. Using Lemma 1, it is easy to extend the diagonal averaging approach in Golyandina et al. (2001) to obtain \( \hat{x}_{i,j} \) as follows:

\[
\hat{x}_{i,j} = \frac{1}{n_s} \sum_{l+k=s} x_{i,j},
\]

(16)

where \( s = i + j \) and \( n_s \) stands for the number of \((l, k)\) pairs such that \( l + k = s \). Denote this projection by \( \Pi : \mathbb{H}_L \rightarrow \mathbb{H}_L \), and set \( \hat{X}_{I_q} = \Pi X_{I_q} \). Now we can define \( Y_q = T^{-1} \hat{X}_{I_q} \) and reconstruct the FTS.

3.2 Separability

Let \( y = y^{(1)} + y^{(2)} \), where \( y^{(i)} = (y^{(i)}_1, \ldots, y^{(i)}_N) \), \( i = 1, 2 \), are FTS. Using a fixed window length \( L \) for each series \( y^{(i)} \), denote \( \{ x^{(i)}_k \}_{k=1}^K \) as a sequence of functional lagged vectors and \( L^{(i)} \) as the linear space spanned by \( \{ x^{(i)}_k \}_{k=1}^K \). Analogous to univariate SSA, separability of the series \( y^{(1)} \) and \( y^{(2)} \) is equivalent to \( L^{(1)} \perp L^{(2)} \), which is the same as \( \{ x^{(1)}_k, x^{(2)}_k \}_{k=1}^K \) of \( 0 \) for all \( k, k' = 1, \ldots, K \). Furthermore, a necessary condition for separability can be defined based on \( w \)-correlation measure. To do this, consider the weighted inner product of two series \( y^{(1)} \) and \( y^{(2)} \) as

\[
\langle y^{(1)}, y^{(2)} \rangle_w = \sum_{i=1}^N w_i \langle y^{(1)}_i, y^{(2)}_i \rangle,
\]

(17)

where \( w_i = \min\{i, L, N - i + 1\} \). We call the series \( y^{(1)} \) and \( y^{(2)} \) \( w \)-orthogonal if

\[
\langle y^{(1)}, y^{(2)} \rangle_w = 0.
\]

(18)
**Theorem 2.** If the series $y_N^{(1)}$ and $y_N^{(2)}$ are separable, then they are $w$-orthogonal.

Also, to quantify the degrees of separability of two FTS, the functional version of the $w$-correlation measure can be obtained by replacing the new definition of the weighted inner product (17) into (5).

4 IMPLEMENTATION STRATEGY

In practice, functional data are being recorded discretely and then converted to functional objects using proper smoothing techniques. We refer to Ramsay and Silverman (2007) for more details on preprocessing the raw data. Let $\{v_i\}_{i \in \mathbb{N}}$ be a known basis system (not necessarily orthogonal) of the space $\mathbb{H}$. Each functional observation in $\mathbb{H}$ can be projected into the subspace $\mathbb{H}_d := sp\{v_i\}_{i=1}^d$, where $d \leq N$ can be determined by a variety of techniques (e.g., cross-validation). Therefore, each $y_j \in \mathbb{H}_d$ is uniquely represented by

$$y_j = \sum_{i=1}^{d} c_{ij} v_i, j = 1, \ldots, N.$$  \hfill (19)

Let us define the quotient sequence, $q_k$, and remainder sequence, $r_k$, by

$$k = (q_k - 1)L + r_k, 1 \leq r_k \leq L, 1 \leq q_k \leq d.$$  \hfill (20)

Note that for any given $k \leq Ld$, one may use (20) to determine $q_k$ and $r_k$ uniquely, so these sequences are well defined. Now, consider $\phi_k$ as the functional vector of length $L$ with all zero functions except the $r_k$th element, which is $v_{q_k}$.

**Lemma 2.** The sequence $\{\phi_k\}_{k=1}^{Ld}$ is a basis system for $\mathbb{H}_d^L$, where $\mathbb{H}_d^L$ is the Cartesian product of $L$ copies of $\mathbb{H}_d$.

Now using Lemma 2 and (7), each lag vector $x_j$ admits a unique representation as

$$x_j = \sum_{i=1}^{Ld} b_{ij} \phi_i = \Phi b_j = 1, \ldots, K,$$  \hfill (21)

where $\Phi = [\phi_1, \ldots, \phi_{Ld}]$ and

$$b_j = (c_{1j}, \ldots, c_{1j+L-1}, c_{2j}, \ldots, c_{2j+L-1}, \ldots, c_{dj+L-1})^T.$$  \hfill (22)

Note that here, $r = di m(R(\mathcal{X})) \leq \min\{Ld, K\}$. Hence, for any $a \in \mathbb{R}^K$, we have

$$\mathcal{X}a = \sum_{j=1}^{K} a_j x_j = \sum_{j=1}^{K} a_j \Phi b_j = \Phi B a,$$  \hfill (23)

where $B = [b_{ij}]_{i=1..d, j=1..K} = [b_1 | b_2 | \ldots | b_K]_{Ld \times K}$. 

**Theorem 3.** Suppose $X = G^{1/2}B$, where $G = \left(\left(\Phi_i, \Phi_j\right)_{i,j=1}^{ld}\right)$ is the Gram matrix. Let us denote the collection $\left(\sqrt{\lambda_i}u_i, v_i\right)$ as the $i$th eigentriple of $X$. Now define $\psi_i = \Phi w_i$, where $w_i = G^{-1/2}u_i$. The following holds:

(i) $X^* \psi_i = \sqrt{\lambda_i} v_i$

(ii) $X v_i = \sqrt{\lambda_i} \psi_i$

(iii) $\left\{\psi_i\right\}_{i=1}^r$ is an orthonormal basis for $R(X)$.

**Corollary 1.** The collection of triples $\left(\sqrt{\lambda_i}, \psi_i, v_i\right)$, $i = 1, \ldots, r$ defined in Theorem 3 defines the fSVD of $X$.

Note that the Gram matrix $G$ used in Theorem 3 can be further simplified to $G = \left(\delta_{i,j} v_i v_j^T\right)_{i,j=1}^{ld}$, where $\delta_{i,j} = 1$ if $i = j$ and zero otherwise. Now we have the recipes to proceed with Algorithm 1, which we use to obtain the FSSA results.

**Algorithm 1 FSSA algorithm**

Initialize the finite-dimensional subspace spanned by $H = \ker G^T G$. Project $(x_i)_{i=1}^r$ into $H$ and extract the coefficients $(c_i)_{i=1}^{r,ld}$ as given in equation (2) of the manuscript.

Define the Gram matrix $G$, and the basis matrix $\Phi$ using the following loops

for $r_i = 1$ to d do

for $q_i = 1$ to d do

Set $\Phi_{i,j} = \delta_{i,j} + r_i$

end for

end for

Define the matrix B using the following loop

for $j = 1$ to K do

Set $H_j = (c_{j,i}c_{j,i})^T$

end for

Set $X = G^{1/2}B$

for $i = 1$ to d do

Extract the $i$th eigentriple of $X$: $\left(\lambda_i u_i, v_i\right)$

Set $w_i = G^{-1/2}u_i$

Set $v_i = \Phi w_i$

Set $\psi_i = \sqrt{\lambda_i} \Phi w_i$

end for

// Obtain the partition $\left\{\{r_1, r_2, \ldots, r_d\}\right\}$ of the set of indices in the grouping stage.

for $q = 1$ to m do

Set $X_q = \sum_{r_i \leq q} \psi_i$

Set $v_q = \Pi_{r_i \leq q} \psi_i$

end for

5 NUMERICAL STUDY

In this section, first, we present a simulation study to elaborate the use of the FSSA compared with dFPCA and MSSA under different scenarios. To do so, we utilize the implementation of the proposed model that is available as an R package named Rfssa in the CRAN repository (Haghbin et al. 2019). We also use the freqdom.fda (Hörmann et al. 2015) and Rssa (Golyandina et al. 2015) packages to obtain the dFPCA and MSSA results. In the second subsection, we analyze a remote sensing dataset using Rfssa and provide some visualization tools that come in handy in the grouping step.
We developed a shiny app, included in the Rfssa package, also available at [https://fssa.shinyapps.io/fssa/](https://fssa.shinyapps.io/fssa/), to demonstrate and reproduce different aspects of the simulation setup. Furthermore, it can be used to compare the results of dFPCA, MSSA and FSSA on the remote sensing data, call center dataset, or any other FTS provided by the end user.

5.1 Simulation study

For the simulation setup, consider the FTS of lengths $N = 50, 100, 150,$ and $200$, which are observed in $n = 100$ fixed equidistant discrete points on $[0, 1]$ from the following model:

$$Y_t(s_i) = m_t(s_i) + X_t(s_i), s_i \in [0,1], i = 1, \ldots, n, \text{and } t = 1, \ldots, N.$$  

(24)

A cubic B-spline basis functions with 15 degrees of freedom is used to convert the $\{Y_t(s_i)\}$s into smooth (continuous) functional curves. In this model, $m_t(s)$ is considered to be a periodic component defined as

$$m_t(s) = e^{s^2} \cos(2\pi\omega t) + \cos(4\pi s) \sin(2\pi\omega t),$$

(25)

where $\omega$ is the model frequency with three different values ($\omega = 0$, 0.1, and 0.25). Figure S1 depicts a perspective and a heatmap view of an FTS given in (25) for $N = 50$ and $\omega = 0.1$.

The $X_t(s)$ in (24) is a stochastic term that is generated under four different settings with an increasing trend in complexity. In the first setting, we consider that $\{X_t(s_i), t = 1, \ldots, N \text{ and } i = 1, \ldots, n\}$ are drawn from an independent Gaussian white noise (GWN) process with zero mean and standard deviation equal to 0.1. It is expected to obtain acceptable performance from FPCA for reconstructing the FTS in the first setting as intuitively FPCA performs well under this ideal framework (see Maadooliat et al. 2015, for more details).

In the remaining three settings, the $\{X_t(s)\}$ processes are simulated from a functional autoregressive model of Order 1, $\text{FAR}(1)$, defined by

$$X_t(s) = \Psi X_{t-1}(s) + \varepsilon_t(s),$$

(26)

where $\Psi$ is an integral operator with a parabolic kernel as follows:

$$\psi(s, u) = \gamma_0(2 - (2s - 1)^2 - (2u - 1)^2).$$

The constant $\gamma_0$ is chosen such that the Hilbert–Schmidt norm defined by

$$\|\Psi\|_\mathcal{S}^2 = \int_0^1 \int_0^1 |\psi(s, u)|^2 dsdu$$

acquires the values $\|\Psi\|_\mathcal{S}^2 = 0, 0.5, \text{and } 0.9$ for the remaining three settings, respectively. In these settings, the white noise terms $\varepsilon_t(s)$ are considered as independent trajectories of the standard Brownian motion over the interval $[0, 1]$. It is worth noting that as we increase the Hilbert–Schmidt norm, $\|\Psi\|_\mathcal{S}^2$, in the FAR(1) models, the dependency structure of consecutive FTS gets more twisted, and we expect it would be more challenging to reconstruct the true structures, $\{X_t(s)\}$.

To compare the performance of FSSA and MSSA, we further consider three window length parameters ($L = 20, 30, \text{and } 40$) in our simulation setup. For the sake of consistency in all of the reconstruction procedures
(dFPCA, MSSA and FSSA), we use the first two leading eigencomponents. The ratio of amplitude of the signal to error is used as the signal-to-noise ratio \( SNR = \frac{\sum_{i=1}^{n} m_t(s_i)^2}{\sum_{t=1}^{N} \sum_{i=1}^{n} x_t(s_i)^2} \). As a measure of goodness of fit, we use the root-mean-square error (RMSE) defined as
\[
RMSE = \sqrt{\frac{1}{N \times n} \sum_{t=1}^{N} \sum_{i=1}^{n} (m_t(s_i) - \hat{y}_t(s_i))^2},
\]
where \( \hat{y}_t(s_i) \) is the FTS reconstructed by each method. We repeat each setting 1000 times and report the mean of the RMSEs in Table 1.

**TABLE 1.** The mean of RMSE for 1000 generations of the simulated model by dFPCA, MSSA, and FSSA approaches

<table>
<thead>
<tr>
<th>Model</th>
<th>( \omega )</th>
<th>( N )</th>
<th>dFPCA</th>
<th>MSSA ( L = 20 )</th>
<th>MSSA ( L = 30 )</th>
<th>MSSA ( L = 40 )</th>
<th>FSSA ( L = 20 )</th>
<th>FSSA ( L = 30 )</th>
<th>FSSA ( L = 40 )</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setting 1</td>
<td>0.00</td>
<td>50</td>
<td>0.092</td>
<td>0.026</td>
<td>0.022</td>
<td>0.019</td>
<td>0.010</td>
<td>0.011</td>
<td>0.014</td>
<td>241.91</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.088</td>
<td>0.024</td>
<td>0.020</td>
<td>0.018</td>
<td>0.008</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>241.57</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.074</td>
<td>0.024</td>
<td>0.020</td>
<td>0.017</td>
<td>0.007</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>240.37</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.071</td>
<td>0.024</td>
<td>0.019</td>
<td>0.017</td>
<td>0.007</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>241.08</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>50</td>
<td>0.062</td>
<td>0.028</td>
<td>0.024</td>
<td>0.022</td>
<td>0.009</td>
<td>0.011</td>
<td>0.014</td>
<td>146.95</td>
</tr>
<tr>
<td>GWN</td>
<td>100</td>
<td>0.045</td>
<td>0.027</td>
<td>0.023</td>
<td>0.020</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
<td>0.007</td>
<td>147.20</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.037</td>
<td>0.027</td>
<td>0.022</td>
<td>0.019</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>146.23</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.033</td>
<td>0.026</td>
<td>0.022</td>
<td>0.019</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>146.60</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>50</td>
<td>0.046</td>
<td>0.028</td>
<td>0.024</td>
<td>0.022</td>
<td>0.009</td>
<td>0.011</td>
<td>0.014</td>
<td>147.34</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.038</td>
<td>0.027</td>
<td>0.022</td>
<td>0.020</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
<td>0.007</td>
<td>145.87</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.032</td>
<td>0.027</td>
<td>0.022</td>
<td>0.019</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>147.35</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.030</td>
<td>0.026</td>
<td>0.022</td>
<td>0.019</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>147.07</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>50</td>
<td>1.008</td>
<td>0.251</td>
<td>0.246</td>
<td>0.270</td>
<td>0.244</td>
<td>0.248</td>
<td>0.284</td>
<td>7.62</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.993</td>
<td>0.205</td>
<td>0.185</td>
<td>0.179</td>
<td>0.190</td>
<td>0.176</td>
<td>0.174</td>
<td>5.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.974</td>
<td>0.192</td>
<td>0.166</td>
<td>0.153</td>
<td>0.175</td>
<td>0.155</td>
<td>0.145</td>
<td>5.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.960</td>
<td>0.186</td>
<td>0.158</td>
<td>0.143</td>
<td>0.168</td>
<td>0.145</td>
<td>0.134</td>
<td>5.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>50</td>
<td>0.636</td>
<td>0.220</td>
<td>0.211</td>
<td>0.235</td>
<td>0.204</td>
<td>0.216</td>
<td>0.264</td>
<td>3.34</td>
</tr>
<tr>
<td>Setting 2</td>
<td>100</td>
<td>0.633</td>
<td>0.193</td>
<td>0.165</td>
<td>0.154</td>
<td>0.158</td>
<td>0.144</td>
<td>0.143</td>
<td>4.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.633</td>
<td>0.188</td>
<td>0.158</td>
<td>0.141</td>
<td>0.149</td>
<td>0.130</td>
<td>0.120</td>
<td>4.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.632</td>
<td>0.186</td>
<td>0.154</td>
<td>0.136</td>
<td>0.144</td>
<td>0.123</td>
<td>0.112</td>
<td>3.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>50</td>
<td>0.636</td>
<td>0.223</td>
<td>0.214</td>
<td>0.235</td>
<td>0.206</td>
<td>0.218</td>
<td>0.263</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.633</td>
<td>0.194</td>
<td>0.167</td>
<td>0.156</td>
<td>0.160</td>
<td>0.146</td>
<td>0.144</td>
<td>4.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.633</td>
<td>0.188</td>
<td>0.157</td>
<td>0.140</td>
<td>0.148</td>
<td>0.129</td>
<td>0.120</td>
<td>3.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.632</td>
<td>0.185</td>
<td>0.153</td>
<td>0.136</td>
<td>0.143</td>
<td>0.122</td>
<td>0.111</td>
<td>4.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>50</td>
<td>0.841</td>
<td>0.316</td>
<td>0.310</td>
<td>0.350</td>
<td>0.305</td>
<td>0.311</td>
<td>0.364</td>
<td>7.20</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.818</td>
<td>0.271</td>
<td>0.240</td>
<td>0.231</td>
<td>0.250</td>
<td>0.228</td>
<td>0.224</td>
<td>5.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.806</td>
<td>0.262</td>
<td>0.223</td>
<td>0.203</td>
<td>0.238</td>
<td>0.207</td>
<td>0.192</td>
<td>5.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.801</td>
<td>0.258</td>
<td>0.216</td>
<td>0.194</td>
<td>0.233</td>
<td>0.198</td>
<td>0.180</td>
<td>5.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>50</td>
<td>0.696</td>
<td>0.277</td>
<td>0.267</td>
<td>0.300</td>
<td>0.255</td>
<td>0.270</td>
<td>0.330</td>
<td>3.27</td>
</tr>
<tr>
<td>Setting 3</td>
<td>100</td>
<td>0.709</td>
<td>0.241</td>
<td>0.207</td>
<td>0.193</td>
<td>0.198</td>
<td>0.180</td>
<td>0.178</td>
<td>4.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.710</td>
<td>0.235</td>
<td>0.197</td>
<td>0.176</td>
<td>0.187</td>
<td>0.162</td>
<td>0.151</td>
<td>4.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.710</td>
<td>0.232</td>
<td>0.192</td>
<td>0.170</td>
<td>0.180</td>
<td>0.153</td>
<td>0.139</td>
<td>3.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>50</td>
<td>0.707</td>
<td>0.214</td>
<td>0.205</td>
<td>0.225</td>
<td>0.198</td>
<td>0.209</td>
<td>0.254</td>
<td>2.93</td>
</tr>
</tbody>
</table>
By comparing the results in Table 1, it can be seen that FSSA outperforms dFPCA in different scenarios. This may not be surprising, as the main task of dFPCA is dimension reduction. Except for the first setting, MSSA also outperforms dFPCA significantly. Furthermore, FSSA performs better than MSSA in most of the cases except in the case where the length of the FTS is small ($L = 50$) and the window size, $L$, is getting closer to $N$. However, it is clear that FSSA is the optimal method for reconstructing the longer FTS ($N \geq 100$).

For all methods, RMSE decreases as the length of the series increases. For two smaller frequencies ($\omega = 0$ and $0.1$), the average of RMSE increases as the noise structure becomes more complex in Settings 1 through 4, while it decreases for $\omega = 0.25$. This might be happening due to the unpredicted cross-correlation of the functional noise structures and the periodic form of FTS.

The efficiency (ratio of RMSE) of MSSA to FSSA for different window lengths ($L$), time series lengths ($N$), and frequencies ($\omega$) is examined in Figure 3. The overall pattern confirms the improvement in RMSE for FSSA as the time series gets longer (larger $N$). Overall, as $L$ is increasing, the pattern of the ratio of RMSEs remains unchanged. Although as the window length becomes larger, the improvement either diminishes for longer FTS or disappears (or reverses) for smaller $N$. It is also worth noting that in Setting 1 (GWN), based on the right panel of Figure 3 and Table 1, the FSSA dominates the other two methods in all combinations of parameters with a better efficiency scale. The visualization of Figure 4 shows a sample of simulation setups where subplot (a) shows simulated functional data, subplot (b) shows the true signal, and subplots (c) and (d) show the results of FSSA and MSSA, respectively for varying the noise model and the frequency parameter ($\omega$). We see from Figure 4 that FSSA graphically appears to perform better in extracting the true signal as compared to MSSA.
5.2 Application to remote sensing data

Tracking changes in vegetation over time has become of interest to researchers who want to preserve wildlife. One technique used to track how much vegetation is present in a region is through field surveys. The problem with this technique is that it is especially difficult to implement in low population areas (Panuju & Trisasongko, 2012). The use of remote sensing data in this context is preferred in order to detect man-made or natural changes in vegetation. One source of remote sensing data is from NASA’s MODerate-resolution Imaging Spectroradiometer (MODIS) satellite, which provides images, twice daily, of regions around the globe at varying spatial resolutions (Tuck et al. 2014). Normalized difference vegetation index (NDVI) is a commonly used pixel-wise index in MODIS satellite images. The NDVI values are bounded between 0 and 1, where index values that are closer to 1 indicate that more vegetation is present and smaller values indicate the absence of vegetation (Tuck et al. 2014).

Many studies have used the spectral NDVI measure and its variants in order to remotely track changes in vegetation over time (Lambin, 1999). The temporal average and temporal variability of NDVI images have been used as explanatory variables for the number of different types of vegetation present in many regions (Tuck et al. 2014). Also, Panuju and Trisasongko (2012) used the maximum value of NDVI images taken of the Jambi province in Indonesia to form a time series. The resulting time series was then analyzed using an X12-ARIMA model in order to identify trend and seasonal changes in woody vegetation (Panuju & Trisasongko, 2012). While these statistics (e.g., maximum and average) have seen some success in tracking the changes of the NDVI.
images, they may fail to capture the distribution of the vegetation. Therefore, one may seek a more comprehensive measure (e.g., an FTS) to describe the distribution of the NDVI images.

Here we use 448 NDVI images taken in 16-day increments between February 18, 2002, and July 28, 2019. The images have a spatial resolution of 250 m and are of a square region just outside of the city of Jambi, Indonesia between 103.61°–103.68°E and 1.60°–1.67°S. Figure 5 shows the respective NDVI images taken on June 10, 2002, and June 10, 2019. It is interesting to note that, although the respective NDVI images are not similar, the sample means of the NDVI values are very close (they differ by only about 0.0032 units). As shown in Figure 5, the kernel density estimates (KDEs) are much more informative in representing the distribution of the NDVI images.

We follow Silverman's rule of thumb (Silverman, 1986) for bandwidth selection and obtain the KDEs. Then we project the results onto a functional space spanned by a cubic B-spline basis, selected via the GCV criterion. Figure 6a shows the projected KDEs onto the B-spline space. We pass the results (448 FTS) as input to the FSSA algorithm with the lag $L = 45$, and study the behavior of the NDVI images over almost two decades, using the Rfssa package.

According to the subplots in Figure 6 (b and c: the singular values and w-correlation plots), a suitable partition would be grouping the first and fourth components separately, plus the second and third components jointly ($G = \{1,2–3,4\}$). The remaining subplots in Figure 6 (d–f: the right singular vectors and left singular functions) indicate that the first component captures mean behavior, the second and third capture annual behavior, and the fourth captures trend.

We follow Silverman's rule of thumb (Silverman, 1986) for bandwidth selection and obtain the KDEs. Then we project the results onto a functional space spanned by a cubic B-spline basis, selected via the GCV criterion. Figure 6a shows the projected KDEs onto the B-spline space. We pass the results (448 FTS) as input to the FSSA algorithm with the lag $L = 45$, and study the behavior of the NDVI images over almost two decades, using the Rfssa package.

According to the subplots in Figure 6 (b and c: the singular values and w-correlation plots), a suitable partition would be grouping the first and fourth components separately, plus the second and third components jointly ($G = \{1,2–3,4\}$). The remaining subplots in Figure 6 (d–f: the right singular vectors and left singular functions) indicate that the first component captures mean behavior, the second and third capture annual behavior, and the fourth captures trend.
An important goal of analyzing the NDVI data is to investigate the existence of a temporal trend in the NDVI images over time. It is interesting to see that FSSA can distinguish and separate the overall mean structure (top-left subfigures in Figure 6d–f) and the trend pattern (bottom-right subfigures in Figure 6d–f) in two different components (the first and fourth components). The trend component is causing an interesting change-point behavior after almost a decade (bottom-right subfigures in Figures 6d–f and 7c). We would like to report that SSA algorithm with the lag $L = 45$ cannot separate this trend structure and the overall mean pattern. Therefore, SSA combines these two (the overall mean and the trend components) into one component (the SSA results are omitted due to space constraints).

FIGURE 7 NDVI image reconstructions using (a) the overall mean component (first group), (b) annual harmonic components (second group), (c) trend component (third group) and (d) sum of the trend and overall mean components (first and third groups)

Next, we build the reconstruction of the NDVI images using the grouping suggested by the FSSA exploratory plots, $G$, as shown in Figure 7. It is clear that Figure 7a shows the reconstructed overall mean, where it does not change over time. Figure 7b provides the annual harmonic behavior, and one may observe the change-point behavior after a decade in Figure 7c. The last subfigure, Figure 7d, presents the sum of the trend and overall mean components.

It would be of interest to confirm the properties of the reconstructed groups using some rigorous statistical procedures. In order to do this, we provide the multivariate trace periodicity test of Hörmann et al. (2018) to test for the annual seasonal effect of FTS and a bootstrap procedure to test existence of a trend (Grosjean & Ibanez, 2018) over the time series of the mean of the coefficients associated to the B-spline basis. We obtain the results of these two tests (periodicity and trend) on four sets of FTS (original signal $y_t(s)$, $R_1$, $R_2$, and $R_3$), where $R_i$ represents residual curves obtained via removing the reconstructed FTS by the first $i$ groups from the original signal $y_t(s)$. Table 2 provides the $p$ values of the tests for the periodicity and the trend. It is clear that the periodicity test captures the annual pattern for $y_t(s)$ and $R_1$ that contains the seasonal components ($p$ value = 0). Also the trend test is significant for all of the FTS except $R_3$, which does not contain the third group (fourth component).

TABLE 2. $P$ values of the multivariate trace periodicity test of Hörmann et al. (2018) and the bootstrap trend test (Grosjean & Ibanez, 2018) on four FTS: ($y_t(s)$, $R_1$, $R_2$, and $R_3$)

<table>
<thead>
<tr>
<th>FTS</th>
<th>Periodicity test</th>
<th>Trend test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_t(s)$</td>
<td>0.00</td>
<td>0.02</td>
</tr>
</tbody>
</table>
While the FSSA procedure was able to separate out the year-long seasonal pattern, it was also able to detect the less obvious trend component present in the NDVI data, which indicates a loss of vegetation over the previous 20 years. It was determined that between 2001 and 2015, grassland and shrubs in the Jambi region accounted for a large amount of the vegetation lost due to controlled burning for human use (Prasetyo, Dharmawan, Nasdian, & Ramdhoni, 2016). The region specified in images that we study here is primarily upland agriculture and grass, and it appears to have been a hotspot for controlled fires (Prasetyo et al. 2016).

6 DISCUSSION
In this paper, we constructed the FSSA procedure by incorporating the FDA techniques in basic SSA via MFPCA. The contribution of the proposed model is to provide practitioners with some tools to utilize the advantages of SSA in FTS. Accordingly, the researchers can analyze functional sequences (e.g., time series, longitudinal or spatial data) via FSSA. Alternatively one may approach the problem using MSSA, given that the data points are measured in fixed, regular grid points over time.

As for the ease of use, an efficient and user-friendly R implementation of FSSA is developed in the Rfssa package. Furthermore, a shiny web application is also included in the package, and it is available at https://fssa.shinyapps.io/fssa/ for reproducing the results of this paper or analyzing any other FTS.

ACKNOWLEDGEMENTS
This work used the high-performance computing cluster Perè at Marquette University funded by NSF awards OCI-0923037 and CBET-0521602. We thank an associate editor and an anonymous referee for their constructive and thoughtful comments that helped us to improve the paper.

References


