Chiang Mai J. Sci. 2023; 50(3): 1-14
https://doi.org/10.12982/CMJS.2023.026
Journal homepage : http://epg.science.cmu.ac.th/ejournal/

Research Article

Strategies for Producing Reliable Trends Forecasting of COVID-19 Pandemic in Malaysia using Dynamic Mode Decomposition

Noor Atinah Ahmad* and Nurul Ashikin Othman

School of Mathematical Sciences, Universiti Sains Malaysia, USM Penang 11800, Malaysia

*Author for correspondence; e-mail: nooratinah@usm.my

Received: 1 November 2022
Revised: 17 January 2023
Accepted: 18 February 2023

ABSTRACT

Dynamic Mode Decomposition (DMD) with time delay embedding is used to predict dynamic patterns in univariate time series. An important pattern that can be extracted using DMD is the trend or global change in a time series which is useful for producing reliable forecast. DMD utilizes the computationally efficient singular value decomposition (SVD) to produce a low rank approximation of the linear operator that brings about the dynamic patterns in the time series. Trend in the time series is translated as dynamic modes of the operator with low frequencies. The time evolution of this low frequency pattern produces forecast of the time series. In this paper, we outline the strategies for extracting trend component from COVID-19 time series of Malaysia. It is discovered that, other than identifying modes with slow varying frequencies, we need to also resolve the time stamp delay, so that mean-square error of the reconstructed time series is minimal. Information of the magnitude and phase of DMD modes are useful to identify persistent patterns and remove nonstationary ones. We compare the performance of DMD with another SVD-based method which is the singular spectrum analysis (SSA) and our results highlight certain fundamental difference between these two methods. The forecasts from SSA tend to lean towards the direction of maximum variance, producing low reconstruction error but slow to detect sudden changes in the time series. On the other hand, forecasts from DMD captures the phases of dominant modes that dictates the overall global pattern, hence providing a better prediction of future dynamics of the time series.

Keywords: dynamic mode decomposition, singular value decomposition, univariate time series

1. INTRODUCTION

The study of time series analysis includes extracting trend or information about the time series global change to produce a reliable forecast and it is widely applied in various field such as finance [1], biomedical engineering [2-4], geophysical study [5], and epidemiological study [6, 7]. Many methods for time series analysis rely on decomposition techniques to separate different variations in a series which are classified as trends, seasonality and/or cyclic changes and 'irregular' fluctuations often termed as 'noise'. Examples of the decomposition methods are Singular Spectrum Analysis (SSA) [2, 6, 8], Time Series Decomposition [9], and Auto Regressive Integrated Moving Average (ARIMA) [10].

Another type of decomposition technique that shares similar features with SSA is the Dynamic Mode Decomposition (DMD). It was first introduced in the fluid dynamics community [11, 12] to uncover spatio-temporal coherent structures of complex flows but has since evolved into other areas of applications [13-22]. It was shown in [12] that DMD approximates the spectra of the Koopman operator [23] which is an infinite-dimensional linear operator that represents nonlinear and finite-dimensional dynamics without linearization. The modes and eigenvalues of the Koopman operator capture future evolution of observables of a dynamical system. The use of its data-driven spectral decomposition and model reduction was first proposed in [24]. This approach captures intrinsic dynamics that is observed over time thereby allowing reduced order DMD modes to be computed without relying on any underlying assumptions on physical laws. Like SSA, it has the ability to decompose signals into interpretable components such as trends, oscillation, and noise. In terms of formulation, DMD has the ability to capture frequency components which allows the extraction of information regarding the growth/decay in the data as well as computing the modes that are associated with complex eigenvalues.

Although DMD is more popularly used for multi-dimensional, it can also be used on univariate time series by the use of time delay embedding. Delay embedding is a key methodology for predicting and forecasting of complex dynamical behavior [25,26]. A complete treatment of DMD for univariate time series can be found in [27] and various other applications [28, 29]. In this paper, we investigate the applicability of DMD with time delay embedding on the COVID-19 time series of Malaysia. The objective of the investigation is to determine the strategies for extracting global trends in the time series to produce reliable forecasts. In our previous work [6, 7], we investigated the capabilities of SSA in the analysis and forecasting of COVID-19 Pandemic in Malaysia. Since both SSA and DMD utilize the singular value decomposition (SVD) to separate components in a time series, we aim to highlight the fundamental difference in the attributes of these methods. The remainder of the paper is structured as follows: First we describe the methodology behind the DMD algorithm. Then we explain the experimental design, strategies and results before we conclude the paper with a discussion.

2. METHODOLOGY

2.1 Infectious Disease Data and Nonlinear Dynamical System

Mathematical modeling of the transmission of infectious disease has a long and varied history, as indicated by the great number of epidemiological models developed to explain the spread of a wide range of diseases [30]. These models are in the form of dynamical system models and frequently include nonlinear components to describe the observed occurrences in a nonlinear system. To draw the connection between infectious disease data and nonlinear dynamical system, we begin by considering a nonlinear discrete dynamical system in its general form as follows:

$$\mathbf{x}_{n+1} = \mathcal{F}(\mathbf{x}_n, \boldsymbol{\theta}), \tag{1}$$

where n denotes the discrete time step, $\mathcal{F}(x_n, \theta)$ is a mapping that maps the current state vector x_n to its future state $\mathbf{x}_{n+1}$ and $\boldsymbol{\theta}$ is a vector of system parameters. Most compartmental epidemic models can be represented in the form as in Eqn. (1), where a typical state variable is $\mathbf{x}_n = (S_n, I_n, R_n, \dots)^T$ where S_n is the number of susceptible cases, I_n is the number of infected cases, R_n is the number of recovered cases, etc. In an epidemic model, $\mathcal{F}$ represents the physical rules that govern how disease propagates and these rules depend on certain epidemic parameters $\boldsymbol{\theta} = (\beta, \gamma, \dots)^T$ where β and γ could be the transmission and removal rate of the disease respectively.

A more practical view of Eqn. (1) is to describe a dynamical system from the point of view of observables, or measurements of $\mathbf{x}_n$.

Let g maps $\mathbf{x}_n$ to a scalar valued measurement $g(\mathbf{x}_n)$. Then, instead of Eqn. (1), we consider

$$g(\mathbf{x}_{n+1}) = g\big(\mathcal{F}(\mathbf{x}_n)\big). \tag{2}$$

To predict future values of the measurements, linear operator known as Koopman operator is used. A Koopman operator defines a rule for mapping the current measurement $g(\mathbf{x}_n)$ to a future measurement $g(\mathbf{x}_{n+1})$. Let A be a Koopman operator for Eqn. (2), then

$$Ag(\mathbf{x}_n) = g(\mathbf{x}_{n+1}) = g\big(\mathcal{F}(\mathbf{x}_n)\big). \tag{3}$$

Now, let's consider a vector-valued measurement $\mathbf{y}_n = \mathbf{g}(\mathbf{x}_n) = \big(g_1(\mathbf{x}_n), g_2(\mathbf{x}_n), \dots, g_m(\mathbf{x}_n)\big)^T$, where m is the number of measurements. In the case of epidemic data $\mathbf{y}_n$, can be a set of measurements for the number of daily cases from the start of the epidemic up to day n. The linear nature of the Koopman operator allows us to represent the future measurement $\mathbf{y}_{n+1}$ as an expansion of the form

$$\mathbf{y}_{n+1} = A\mathbf{y}_n = \sum_{i=1}^{\infty} \lambda_i^n \varphi_i(\mathbf{x}_n)\, \mathbf{v}_i, \tag{4}$$

where $\big(\lambda_i, \varphi_i(\mathbf{x}_n)\big)$ is the eigenvalue-eigenfunction pair of A associated with measurement $\mathbf{y}_n$, and $\mathbf{v}_i$ are the Koopman modes of the map $\mathcal{F}$ [12].

The relationship in Eqn. (4) provides a direct way of representing measurements of infectious disease data in terms of the Koopman decomposition which consists of the eigenvalues λ_i, eigenvectors $\varphi_i(x_n)$, and $\mathbf{v}_i$. But it is only feasible if a finite-dimensional approximation of the Koopman operator is available. Central to the dynamic mode decomposition [12, 13] is the low rank approximation of the Koopman operator using measurement data and computing an approximation to Eqn. (4) as a finite expansion.

2.2 Dynamic Mode Decomposition

Starting with a collection of L consecutive snapshots of measurements $\mathbf{y}_n, n = 1,2, \dots, L$, two different matrices can be constructed as follows:

$$\begin{aligned} Y_1 &= [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{L-1}], \\ Y_2 &= [\mathbf{y}_2, \mathbf{y}_3, \dots, \mathbf{y}_L]. \end{aligned} \tag{5}$$

The Koopman matrix A advances each snapshot $\mathbf{y}_n$ by a single time step into the corresponding future measurement $\mathbf{y}_n$ so that the evolution of all L snapshots is given as

$$Y_2 = AY_1. \tag{6}$$

A best approximation of A in the least-squares sense is obtained by minimizing the Frobenius norm

$$\|Y_2 - AY_1\|_F \tag{7}$$

To use DMD as a prediction tool, delay embedding is used [30]. In particular, for univariate time series, the snapshots $\mathbf{y}_n$ take the form of time delayed samples $\mathbf{y}_n = (y_n, y_{n+1}, \dots, y_{n+K-1})^T$ [27] where, for a time series of length l, $K = l - L + 1$. The optimal construction of matrix A for the linear approximation in Eqn. (6), is given by

$$A = Y_2 Y_1^{\dagger}, \tag{8}$$

where $(.)^{\dagger}$ denotes the "Moore-Penrose" pseudoinverse. Let $Y_1 = U\Sigma V^*$ be the Singular Value Decomposition (SVD) of Y_1, with $(^*)$ denotes the conjugate transpose, $U \in \mathbb{C}^{K\times(L-1)}$, $\Sigma \in \mathbb{C}^{(L-1)\times(L-1)}$ and $V \in \mathbb{C}^{(L-1)\times(L-1)}$. Then $Y_1^{\dagger} = V\Sigma^{-1}U^*$, which gives

$$A = Y_2 V\Sigma^{-1}U^*. \tag{9}$$

Eqn. (9) provides the full-rank matrix A, however, the computation is expensive as the size of the matrix is big which is $K \times K$. A low rank approximation of A is obtained on the subspace spanned by the orthogonal singular vectors U of Y_1, which is [31]

$$\tilde{A} = U^* Y_2 V\Sigma^{-1}. \tag{10}$$

The dynamic modes of the time series reconstruction are made up of the eigenvalues and the eigenvectors of $\tilde{A}$. To achieve these, the eigendecomposition of $\tilde{A}$ is computed as follows:

$$\tilde{A}W = W\Lambda, \tag{11}$$

where the columns of W are the eigenvectors and Λ is a diagonal matrix containing the corresponding eigenvalues of $\tilde{A}$ matrix. The DMD modes are reconstructed using the eigenvectors W according to

$$\Psi = Y_2 V \Sigma^{-1} W. \tag{12}$$

It has been shown in [31] that the columns of Ψ (i.e. the DMD modes) are the eigenvectors of the high-dimensional matrix A that corresponds to the eigenvalues in Λ.

2.2.1 DMD reconstruction, forecasting and trend extraction

Given the DMD modes in Eqn. (12), a complete reconstruction of the time series is given by

$$\mathbf{y}_n = \sum_{i=1}^{L-1} \Psi_i \lambda_i^{n-1} v_i = \Psi \Lambda^{n-1} \mathbf{v}, \tag{13}$$

where Ψ_i is the ith column of Ψ, and $\mathbf{v}$ is the initial amplitude of the dynamic modes obtained from the multiplication of inverse DMD modes and the first lagged vector, i.e, $\mathbf{v} = \Psi^{-1}\mathbf{y}_1$. Based on Eqn. (13), a reconstruction of the matrix Y_2 in Eqn. (5) is given by,

$$Y_r = \Psi \begin{bmatrix} v_1 & \lambda_1^1 v_1 & \lambda_1^2 v_1 & \cdots & \lambda_1^{L-1} v_1 \\ v_2 & \lambda_2^1 v_2 & \lambda_2^2 v_2 & \cdots & \lambda_2^{L-1} v_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{L-1} & \lambda_{L-1}^1 v_{L-1} & \lambda_{L-1}^2 v_{L-1} & \cdots & \lambda_{L-1}^{L-1} v_{L-1} \end{bmatrix} = \Psi \mathrm{T}_r, \tag{14}$$

where the $(L-1) \times L$ matrix T describes the time dynamic of the DMD modes.

It is now apparent that the evolution of the DMD modes for $n = L+1, L+2, \ldots, L+F$, produces a forecast of the time dynamic into F time steps ahead and this is

$$Y_f = \Psi \begin{bmatrix} \lambda_1^L v_1 & \lambda_1^{L+1} v_1 & \lambda_1^{L+2} v_1 & \cdots & \lambda_1^{L+F-1} v_1 \\ \lambda_2^L v_2 & \lambda_2^{L+1} v_2 & \lambda_2^{L+2} v_2 & \cdots & \lambda_2^{L+F-1} v_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_{L-1}^L v_{L-1} & \lambda_{L-1}^{L+1} v_{L-1} & \lambda_{L-1}^{L+2} v_{L-1} & \cdots & \lambda_{L-1}^{L+F-1} v_{L-1} \end{bmatrix} = \Psi \mathrm{T}_f. \tag{15}$$

Combining the reconstruction in Eqn. (14) and the forecast in Eqn. (15) gives

$$Y = [Y_r \quad Y_f] = \Psi [T_r \quad T_f]. \tag{16}$$

The reconstruction in Eqn. (14) can be grouped further into several sub-components such as trend, periodic component, and noise. To elucidate further, we write Eqn. (14) as a sum of three groups of matrices as follows:

$$\begin{aligned} Y_r &= \sum_{j \in G_1} \Psi_j \lambda_j^{n-1} v_j + \sum_{j \in G_2} \Psi_j \lambda_j^{n-1} v_j \\ &\quad + \sum_{j \in G_3} \Psi_j \lambda_j^{n-1} v_j \\ &= \hat{Y}_{trend} + \hat{Y}_{periodic} + \hat{Y}_{noise}, \end{aligned} \tag{17}$$

where G_1 is the group of indices corresponds to trend component, G_2 is the group of indices corresponds to periodic components, and G_3 is the group of indices corresponds to minor components and noise. When constructing the groups, it is useful to consider the frequencies of the DMD modes which are defined as

$$\omega_j = \frac{\ln(\lambda_j)}{\delta t}, \forall\, j \in \{1, 2, \ldots, L-1\}. \tag{18}$$

Here δt refers to the lag between samples of the time series. The real part of ω_i is responsible for the growth/decay of the DMD modes while the imaginary parts bring about the oscillatory behaviour. For trend extraction, we look for sub-component(s) with low oscillation, i.e. the j-th sub-component $\Psi_j \lambda_j^{n-1} v_j$ is a trend if $im(\omega_j) \approx 0$. Then, the evolution of the trend component for $n = L+1, L+2, \ldots, L+F$, produces a forecast of

the time dynamic of trend in the time series into F time steps ahead given by

$$\hat{Y}^{f}_{trend} = \Psi_{G_1} T^{f}_{G_1}, \tag{19}$$

where Ψ_{G_1} consists of the columns of Ψ associated with G_1, while $T^{f}_{G_1}$ consists of the rows of T_f associated with G_1.

2.2.2 Diagonal averaging, time series reconstruction and time stamp delay

Similar to the Singular Spectrum Analysis (SSA), a diagonal averaging of matrix Y in Eqn. (16) is needed to transform it into a time series of length $l + F$ to produce a reconstruction of the trend component of the time series and the F-steps ahead forecast of the trend. Let $\hat{y}_{ij}, 1 \le i \le K, 1 \le j < L + F - 1$ be the entries of $Y, L^* = min(L + F - 1, K)$, and $K^* = max(L + F - 1, K)$.. Also, let $\hat{y}^*_{ij} = \hat{y}_{ij}$ if $L + F - 1 < K$ and $\hat{y}^*_{ij} = \hat{y}_{ji}$ otherwise. Diagonal averaging of Y to produce the time series $\tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \dots, \tilde{y}_{l+F}$ is given by [8]

$$\tilde{y}_n = \begin{cases} \dfrac{1}{n+1}\displaystyle\sum_{i=1}^{n+1} Re\left(\hat{y}^*_{i,n-i+2}\right) & \text{for} \quad 1 \le n < L^* - 1, \\ \dfrac{1}{K^*}\displaystyle\sum_{i=1}^{L^*} Re\left(\hat{y}^*_{i,n-i+2}\right) & \text{for} \quad L^* - 1 \le n < K^*, \\ \dfrac{1}{l+F-K}\displaystyle\sum_{i=n-K^*+2}^{l+F-K^*+1} Re\left(\hat{y}_{i,n-i+1}\right) & \text{for} \quad K^* \le n < l + F. \end{cases} \tag{20}$$

It is important to note that, since $\tilde{A}$ is reconstructed from the matrix Y_2, and Y_2 does not include the first data point, a time stamp delay is introduced [27]. Let τ be this time stamp delay, then in general $\tilde{y}_{n+\tau} \approx y_n$.

2.3 DMD Modes and Epidemiological Interpretation

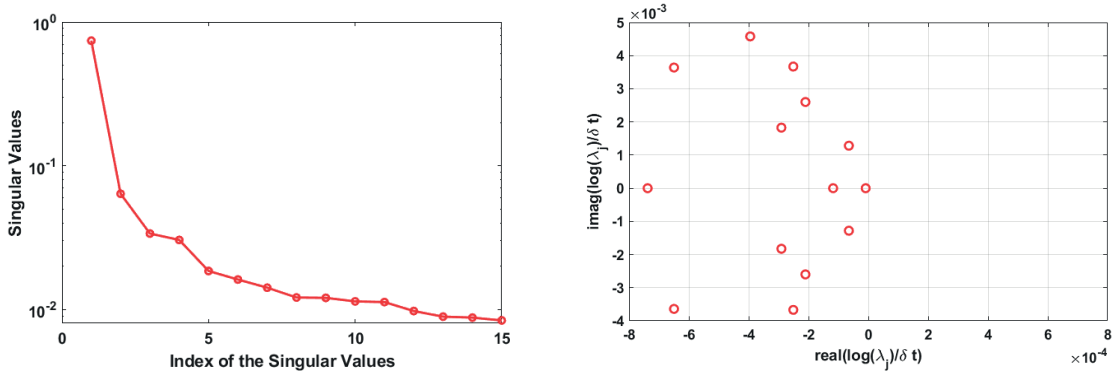
In epidemiological study, it is important to identify and analyze the dynamic patterns. The dynamic modes provided by DMD can help in extracting important information regarding the epidemiological system such that the ability to capture specific dynamic pattern of disease transmission as well as the relative phase of 'peaks' of epidemic. The DMD mode Ψ_j is a discrete approximation of the eigenfunction $\varphi_j(\mathbf{x}_n)$ (a complex-valued function) and it contains two important information pertaining to the dynamic pattern, i) the magnitude $|\Psi_j|$ which is a vector of absolute values of the elements in Ψ_j, and ii) the phase $\angle\Psi_j$ i.e. angle formed by the elements in Ψ_j in the complex plane. Trend forecasting of epidemic time series search for persistent dynamic patterns in the time series and these are characterized by persistent phase and magnitude.

3. Results and Analysis

The experiments in this section are conducted on COVID-19 time series involving the number of daily cases in Malaysia from 25 January 2020 until 31 December 2021 and the dataset is taken from https://www.worldometers.info/coronavirus/. The total length of the time series is $T = 858$. The choice of window size L should be between 2 and $T/2$ [27, 33] and to be able to extract seasonal component, it is necessary to use relatively large L, however, if L is too large, high frequency components will tend to creep in. Since we are only interested in slow varying components, we choose to use $L = 15$ in all the experiments.

3.1 Trend Extraction and Resolving the Time Stamp Delay

In order to extract trend component, we analyse both the singular spectrum and the DMD frequencies (eigenvalues). The log of the singular value spectrum is shown in Figure 1(a). Here we can see the 3rd and the 4th singular values having almost the same value. This indicates that the pair of singular values belong to a periodic component. Singular values ≥ 5 mostly belong to minor high frequency components that captures local properties of the time series. The observation strongly suggests that global trend can be found in only the first two principal singular values.



(a) (b)

Figure 1. Singular Value Spectrum (a) and DMD eigenvalues (b).

By analysing the DMD frequencies in Figure 1(b), we observe three components ($j = 1,2,3$) that have almost zero imaginary components. Imaginary component of the DMD eigenvalues captures oscillatory and high frequency part of the time series whereas the real part captures the growth/decay. The three components near the horizontal axis are therefore the components that represent trends. The indices of the DMD modes associated with the trend component ($j = 1,2,3$) are grouped together and then used to build a dynamical model for the time series which we later use to produce a reconstruction of the time series as well as for forecasting.

In Figure 2, the magnitudes $|\Psi_j|, , j = 1,2,3,4$ are plotted and compared with the original time series. The eigenvalues associated with Ψ_1, Ψ_2 , and Ψ_3 are is $\lambda_1 = 0.9939$, $\lambda_2 = 0.9224$ and $\lambda_3 = 0.6028$ respectively. It is observed that all three DMD modes capture distinctive patterns of the original time series. In particular, the location of the peaks are detected rather accurately especially in the first DMD mode Ψ_1.

The phases $\angle\Psi_j, j = 1,2,3,4$ are presented in Figure 3. There are two persistent phases clearly depicted in Figures 3(a)-(b) associated with the trend components. The phase of the non-trend component Ψ_4 is shown in Figure 3(d) for comparison where it is seen that the phase behaves in a non-stationary manner. From these results, we learn that the phase of DMD modes can distinctively separate trend component from non-trend. Thus, the phase of DMD modes can be used to identify persistent patterns in the time series that capture the global trend of a time series.

The DMD modes associated with the trend component are grouped and then used to reconstruct the original time series in Figure 2(a). A comparison of the reconstructed time series and the original time series is shown in Figure 4. The presence of the time stamp delay can be observed in Figure 4(a) where the reconstructed time series can be seen to be shifted to the right. To resolve this issue, we determined the time stamp τ empirically by choosing τ for which the mean square error (MSE) of the reconstructed time series is minimum. The MSE is calculated as

$$(MSE)_r = \frac{\sum_{i=1}^{T}(y_i - \tilde{y}_{i+\tau})^2}{\sum_{i=1}^{T}(y_i)^2}, \qquad (21)$$

where y_i refers to the ith true data and $\tilde{y}_{i+\tau}$ is the estimate of y_i. With a choice of $\tau = L = 15$ we are able to reduce the MSE from 0.0150 to 0.0116 (see Figure 2(b)).

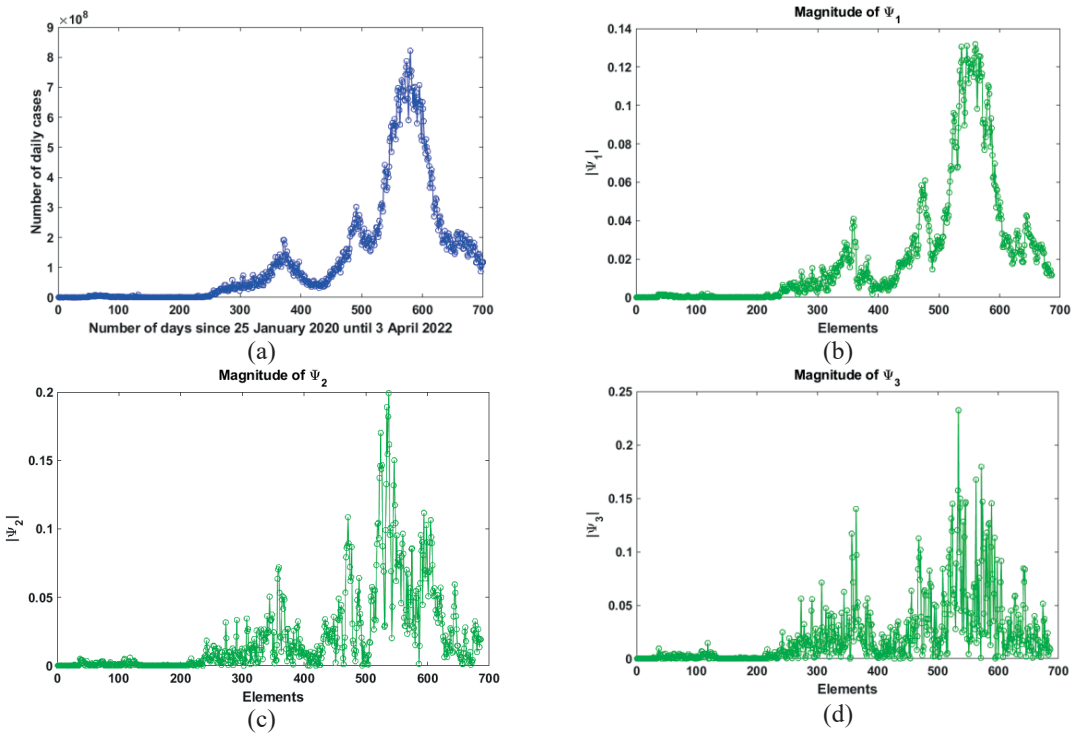


Figure 2. Original time series (number of daily cases) (a), Magnitude of Ψ_1 (b), Magnitude of Ψ_2 (c) and Magnitude of Ψ_3 (d).

3.2 Time Series Reconstruction and Forecasting using DMD

In this section, we compare the performance of DMD and SSA in reconstructing and forecasting the global trend of the Malaysia COVID-19 time series. The time series is divided into two parts; the first part contains data to be used for reconstruction and the second part is treated as future data. Three forecasting points are chosen, namely on 5 October 2020 (255^{th}), 22 April 2021 (454^{th}) and 22 February 2022 (760^{th}).

Based on our analysis in the previous section, we outline the strategies for extracting trend components to be used in the construction of the DMD forecasting model:

Reconstruction stage:

1. Identify the portion of the time series to be used for the reconstruction.
2. With a choice of $L = 15$ identify the indices of DMD modes $im(\omega_j) \approx 0$.
3. Check the phases of the identified DMD components; ensure that they are stationary.
4. Group the DMD trend components and compute Ψ_{G_1}.
5. Construct $\hat{Y}_{trend} = \Psi_{G_1} T^r_{G_1}$.
6. Use diagonal averaging to reconstruct the time series from $\hat{Y}_{trend}$.
7. Compare the reconstructed time series with the original series. Identify the value of the time stamp delay τ empirically by choosing a value that minimizes $(MSE)_r$ in Eqn. (21).

Forecasting stage:

1. Construct $\hat{Y}^f_{trend} = \Psi_{G_1} T^f_{G_1}$.
2. Use diagonal averaging to construct the original time series with F-steps ahead forecasts from $\hat{Y} = [\hat{Y}_{trend} \quad \hat{Y}^f_{trend}]$.

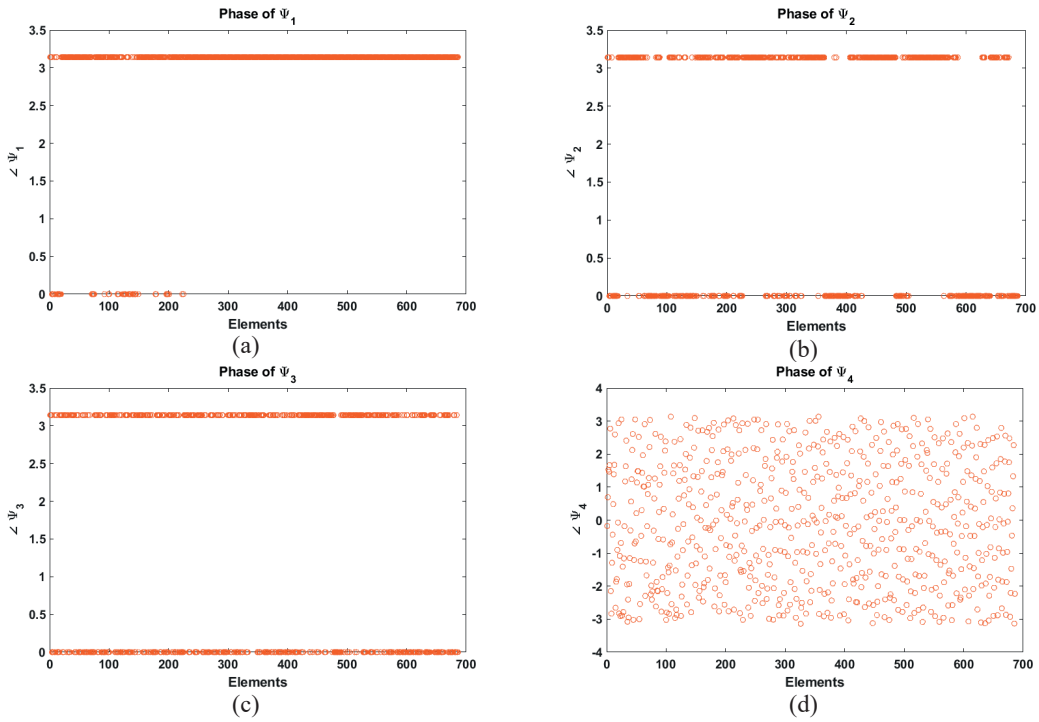


Figure 3. Phase of Ψ_1 (a), Phase of Ψ_2 (b), Phase of Ψ_3 (c) and Phase of Ψ_4 (d).

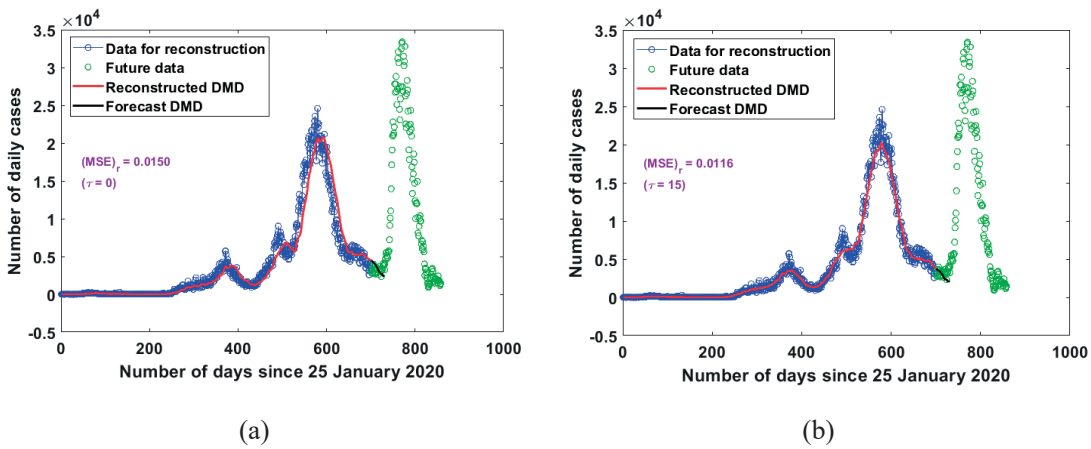


Figure 4. Reconstruction using time stamp delay, $\tau = 0$ (a) and reconstruction using time stamp delay, $\tau = 15 = L$ (b).

3. Adjust the time index of the constructed time series $n \to n + \tau$.

For SSA reconstruction and forecast, components that correspond to the first two principal singular values are used. To evaluate the performance of both methods, the reconstruction MSE and forecast MSE are computed. For DMD the reconstruction MSE is calculated according to Eqn. (21) whereas for SSA, the reconstruction MSE is calculated according to the formula

$$(MSE)_r^{SSA} = \frac{\sum_{i=1}^{T}\left(y_i - \tilde{y}_i^{(SSA)}\right)^2}{\sum_{i=1}^{T}(y_i)^2}, \qquad (22)$$

where $\tilde{y}_i^{(SSA)}$ is the reconstructed time series from SSA components. The forecast MSE are calculated as follows:

For DMD:

$$(MSE)_f^{DMD} = \frac{\sum_{i=1}^{T}\left(y_i - \tilde{y}_{i+\tau}^{f}\right)^2}{\sum_{i=1}^{T}(y_i)^2}, \qquad (23)$$

For SSA:

$$(MSE)_f^{SSA} = \frac{\sum_{i=1}^{T}\left(y_i - \tilde{y}_i^{f(SSA)}\right)^2}{\sum_{i=1}^{T}(y_i)^2}, \qquad (24)$$

where $\tilde{y}_{i+\tau}^{f}$ and $\tilde{y}_i^{f(SSA)}$ are the forecasts from DMD and SSA respectively.

At each forecasting points, a 30-days forecast ($F = 30$) is generated using both DMD and SSA and the results are shown in Figure 5. At the first forecasting point (Figure 5(a) (DMD) and Figure 5(b) (SSA)), we can see both methods are able to detect the increasing trend in the time series. The forecast generated by SSA follows closely the direction of maximum variance of data thus producing a good MSE for the reconstruction. However, the sudden change in the direction of the time series is not detected in its forecast. The DMD forecast appears to be more modest and follows the long-term global trend. At the second forecasting point (Figure 5(c) (DMD) and Figure 5(d) (SSA)), we observe a similar behaviour with SSA, where the forecast tends to follow the direction of maximum variance but in general a bit slow to detect sudden change in the global trend of the time series. The SSA forecasts at this point are showing a slowdown in transmission and nearing a peak. On the other hand, DMD forecast do not generally follow the local trend but instead agrees better with global trend that observes an increase in transmission. An interesting observation is found at the third forecasting point (Figure 5(e) (DMD) and Figure 5(f) (SSA)). The 30-day forecasts are produced near a peak of the time series. Both methods managed to capture the peak of the pandemic, however DMD manage to detect the peak with a much better resolution (in magnitude as well in time) compared to SSA.

The reconstruction MSE and forecast MSE for the results in Figure 5 are compared in Table 1. It is observed that the SSA method consistently records lower reconstruction MSE compared to DMD. However, the forecast MSE for DMD appears to be more controlled and consistent compared to SSA.

To investigate the generalization capability of the DMD and SSA for different window size, $3 \le L \le 15$, the values of reconstruction MSE and forecast MSE are computed for window sizes and presented in Table 2. From Table 2, it is observed that the SSA method consistently records lower reconstruction MSE compared to DMD but the forecast MSE for SSA is consistently higher than that of DMD. The plot of the ratios $(\boldsymbol{MSE})_r/(\boldsymbol{MSE})_f$ for DMD, and $(\boldsymbol{MSE})_r^{\boldsymbol{SSA}}/(\boldsymbol{MSE})_f^{\boldsymbol{SSA}}$ for SSA as shown in Figure 6 provides a clearer picture of the generalization ability of SSA and DMD. For all window size $3 \le L \le 15$, DMD shows better generalization ability compared to SSA where the reconstruction MSE is more comparable to the forecast MSE. This observation is consistent with the observation in Figure 5 where DMD has more tendency to observe global trend of the time series compared to SSA.

3.3 Generalization Capability of SSA and DMD with Increasing Window Size

To investigate the generalization capability of

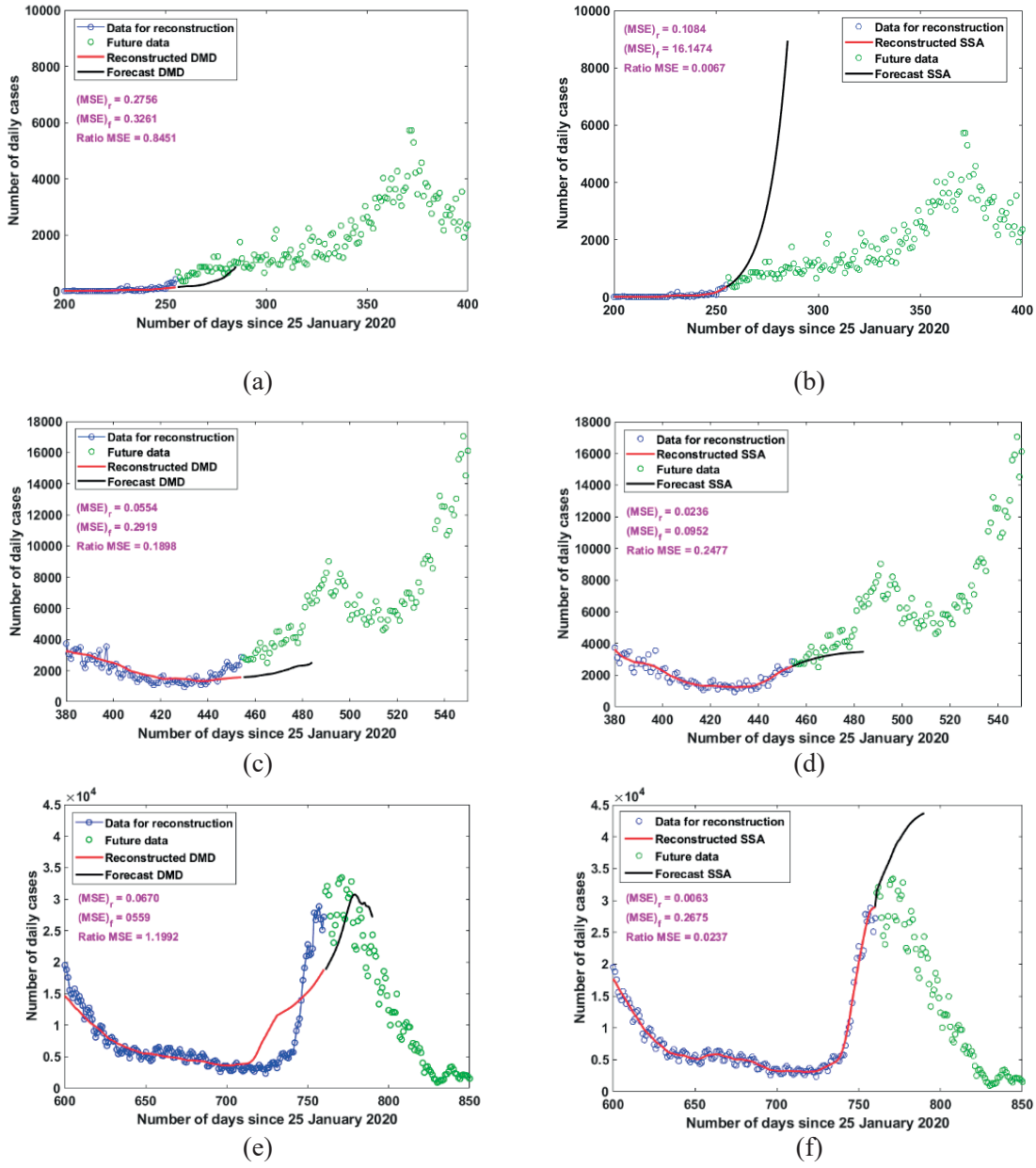


Figure 5. DMD forecast on 5 October 2020 (255^{th}) (a), SSA forecast on 5 October 2020 (255^{th}) (b), DMD forecast on 22 April 2021 (454^{th}) (c), SSA forecast on 22 April 2021 (454^{th}) (d), DMD forecast on 22 February 2022 (760^{th}) (e) and SSA forecast on 22 February 2022 (760^{th}) (f).

Table 1. Reconstruction MSE and forecast MSE of DMD and SSA at forecasting point 255, 454 and 760.

Forecast point	DMD			SSA		
	$(MSE)_r$	$(MSE)_f$	$\frac{(MSE)_r}{(MSE)_f}$	$(MSE)_r^{SSA}$	$(MSE)_f^{SSA}$	$\frac{(MSE)_r^{SSA}}{(MSE)_f^{SSA}}$
255	0.2756	0.3261	0.8451	0.1084	16.1474	0.0067
454	0.0554	0.2919	0.1898	0.0236	0.0952	0.2477
760	0.0670	0.0559	1.1992	0.0063	0.2675	0.0237

Table 2. Reconstruction MSE and forecast MSE of DMD and SSA for several window size, *L*.

L	DMD			SSA		
	$(MSE)_r$	$(MSE)_f$	$\frac{(MSE)_r}{(MSE)_f}$	$(MSE)_r$	$(MSE)_f$	$\frac{(MSE)_r}{(MSE)_f}$
3	0.0337	0.0343	0.9826	0.0025	0.0229	0.1075
4	0.0309	0.0365	0.8477	0.0039	0.0188	0.2062
5	0.0284	0.0481	0.5906	0.0052	0.0184	0.2815
6	0.0279	0.0649	0.4298	0.0060	0.0183	0.3288
7	0.0268	0.0654	0.4100	0.0064	0.0212	0.3033
8	0.0257	0.0652	0.3949	0.0065	0.0261	0.2495
9	0.0213	0.0494	0.4314	0.0048	0.0225	0.2126
10	0.0166	0.0518	0.3202	0.0056	0.1528	0.0368
11	0.0150	0.0635	0.2358	0.0056	0.5699	0.0098
12	0.0132	0.1057	0.1252	0.0054	0.4273	0.0126
13	0.0128	0.1188	0.1077	0.0055	0.2934	0.0187
14	0.0130	0.0965	0.1346	0.0057	0.2760	0.0208
15	0.0134	0.0793	0.1690	0.0061	0.2853	0.0213

the DMD and SSA for different window size, the values of reconstruction MSE and forecast MSE are computed for window sizes $3 \leq L \leq 15$ and presented in Table 2. From Table 2, it is observed that the SSA method consistently records lower reconstruction MSE compared to DMD but the forecast MSE for SSA is consistently higher than that of DMD. The plot of the ratios $(MSE)_r/(MSE)_f$ for DMD, and $(MSE)_r^{SSA}/(MSE)_f^{SSA}$ for SSA as shown in Figure 6 provides a clearer picture of the generalization ability of SSA and DMD. For all window size $3 \leq L \leq 15$, DMD shows better generalization ability compared to SSA where the reconstruction MSE is more comparable to the forecast MSE. This observation is consistent with the observation in Figure 3 where DMD has more tendency to observe global trend of the time series compared to SSA.

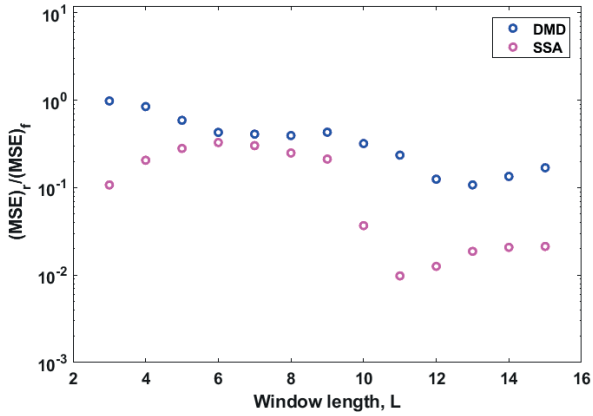


Figure 6. Ratio of reconstruction MSE to the forecast MSE for DMD and SSA model respectively.

4. Discussion and conclusions

In this study, we investigate the applicability of DMD in capturing the global trend in the COVID-19 time series of Malaysia. We outlined the strategies to extract trend component in the time series and to empirically resolving the time stamp delay prior to constructing the forecasting model. Trend components correspond to slow varying DMD modes and these are identified as components whose frequencies have almost zero imaginary part. It is also found that persistent patterns in the time series are correspond well with the DMD modes with persistent stationary phases.

The observations in Figure 5 highlight the following fundamental differences in the capabilities of DMD and SSA:

• Because SSA forecasting model is constructed directly from principal components from the SVD, we found that SSA reconstructs the time series along the direction of maximum variance, and this observation is supported by the consistently low reconstruction MSE. However this attribute limits the capability of SSA to detect sudden changes in the direction of the time series which can result in poor generalization capability.

• When extracting trends from DMD components, high frequency components are strategically ignored and only slow varying components are used for reconstruction. Information on the phase of DMD modes allows better identification of persistent patterns that govern future behaviour of the time series Therefore, with DMD, the emphasize is to capture the global trend, thus its generalization capability is expected to be better than SSA.

• the ability of DMD to isolate persistent pattern allows it to capture temporal events much better compared to SSA. This attribute is observed in Figure 5(c) where DMD managed to detect the peak of the pandemic with a much better resolution (in magnitude as well in time) compared to SSA.

The method and analysis presented in this paper underline the intrinsic properties of DMD that can help to uncover meaningful insights in data-driven analysis of dynamical time series such as epidemic. The results are confined to the COVID-19 time series of Malaysia but has the potential to be explored further for different types of dynamical time series such as environmental time series, biomedical time series, ecological time series and others.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Higher Education Malaysia for the financial support under the FRGS grant with Project Code: FRGS/1/2020/ICT06/USM/02/1.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

REFERENCES

[1] Osmanzade A., *Singular Spectrum Analysis Forecasting for Financial Time Series*, PhD Thesis, University of Tartu, Tartu, 2017.

[2] Sanei S. and Hassani H., *Singular Spectrum Analysis of Biomedical Signals,* CRC press, 2015.

[3] Mohammadi S.M., Kouchaki S., Ghavami M. and Sanei S., *J. Neurosci.,* 2016; **273**: 96–106. DOI 10.1016/j.jneumeth.2016.08.008.

[4] Kouchaki S., Sanei S., Arbon E.L. and Dijk D.J., *IEEE T. Neur. Sys. Reh.,* 2014; **23**: 1-9. DOI 10.1109/TNSRE.2014.2329557.

[5] Ghil M., Allen M., Dettinger M., Ide K., Kondrashov D., Mann M., et al., *Rev. Geophys.,* 2002; **40**: 3-1-3-41. DOI 10.1029/2000RG000092.

[6] Othman N.A.B., Aziz A.A.A., Ahmad N.A., Mohd M.H. and Adam S.I.M., *Matematika: Malay. J. Ind. Appl. Math.,* 2021; **37**: 121-134.

[7] Ahmad N.A., Mohd M.H., Musa K.I., Abdullah J.M. and Othman N.A., *Malays. J. Med. Sci.,* 2021; **28(5)**: 1-9. DOI 10.21315/mjms2021.28.5.1.

[8] Golyandina N. and Zhigljavsky A., *Singular Spectrum Analysis for Time Series,* Springer, 2020.

[9] Deng C., *Time Series Decomposition using Singular Spectrum Analysis,* Electronic Theses and Dissertations, 2014.

[10] Benvenuto D., Giovanetti M., Vassallo L., Angeletti S. and Ciccozzi M., *Data Br.*, 2020; **29**: 105340. DOI 10.1016/j.dib.2020.105340.

[11] Schmid P. and Sesterhenn J., *J. Fluid Mech.,* 2010; **656:** 5-28. DOI 10.1017/S0022112010001217.

[12] Rowley C.W., Mezić I., Bagheri S., Schlatter P. and Henningson D.S., *J. Fluid Mech.,* 2009; **641**: 115–127. DOI 10.1017/S0022112009992059.

[13] Schmid P.J., Li L., Juniper M.P. and Pust O., *Theor. Comp. Fluid Dyn.*, 2011; **25**: 249–259. DOI 10.1007/s00162-010-0203-9.

[14] Kutz J.N., Brunton S.L., Brunton B.W. and Proctor J.L., *Dynamic Mode Decomposition: Data-driven Modeling of Complex Systems*, SIAM, 2016.

[15] Brunton S.L. and Kutz J.N., *Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control,* Cambridge University Press, 2022.

[16] Grosek J. and Kutz J.N., *Dynamic Mode Decomposition for Real-time Background/ Foreground Separation in Video*, Computer Vision and Pattern Recognition, 2014.

[17] Tirunagari S., Poh N., Bober M., and Windridge D., *2016 International Conference onImage, Vision and Computing (ICIVC),* IEEE, 2016; 46–50. DOI 10.1109/ICIVC.2016.7571272.

[18] Erichson N.B. and Donovan C., *Comput. Vis. Image Und.,* 2016; **146**: 40–50. DOI 10.1016/j.cviu.2016.02.005.

[19] Tirunagari S., Poh N., Windridge D., Iorliam A., Suki N. and Ho A.T., *IEEE T. Inf. Foren. Sec.*, 2015; 10:762-777. DOI 10.1109/TIFS.2015.2406533.

[20] Niyigena Ingabire H., Wu K., Toluwani Amos J., He S., Peng X., Wang W., et al., *IEEE J. Biomed. Health,* 2022 May; **26(5)** :2124-2135. DOI 10.1109/JBHI.2021.3130275.

[21] Berger E., Sastuba M., Vogt D., Jung B., and Amor H.B., *Proceeding of the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, Edinburgh, UK, 25-29 August 2014; 593–600. DOI 10.1109/ROMAN.2014.6926317.

[22] Brunton B.W., Johnson L.A., Ojemann J.G. and Kutz J.N., *J. Neurosci.*, 2016; **258**: 1–15. DOI 10.1016/j.jneumeth.2015.10.010.

[23] Koopman B.O., *P. Nat. Acad. Sci. USA.*, 15 March 1932; 255-263. DOI 10.1073/pnas.17.5.315.

[24] Mezic I., *Nonlinear Dyn.*, 2005; **41**: 309–325. DOI 10.1007/s11071-005-2824-x.

[25] Ahmad N.A. and Javed S., *Springer Proceedings in Mathematics and Statistics*, Singapore, 11 June 2021; 309–329.

[26] Golyandina N., Nekrutin V. and Zhigljavsky A., *Analysis of Time Series Structure: SSA and Related Techniques*, Boca Raton, FL: Chapman & Hall/CRC, 2001.

[27] Tirunagari S., Kouchaki S., Poh N., Bober M., and Windridge D., *Dynamic Mode Decomposition for Univariate Time Series: Analysing Trends and Forecasting*, Computer Science, 2017.

[28] Yuan Y., Zhou K., Zhou W., Wen X. and Liu Y., *Phys. Fluids*, 2021; **33**: 095109. DOI 10.1063/5.0064867.

[29] Climaco P., Garcke J. and Teran R.I., *Multi-resolution Dynamic Mode Decomposition for Damage Detection in Wind Turbine Gearboxes*, Signal Processing, 2021.

[30] Anderson R.M. and May R.M., *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1992.

[31] Schmid P.J., *Annu. Rev. Fluid Mech.*, 2022; **54**: 225-254. DOI 10.1146/annurev-fluid-030121-015835.

[32] Tu J.H., Rowley C.W., Luchtenburg D.M., Brunton S.L. and Kutz J.N., *J. Comput. Dyn.,* 2014; **1**: 391–421. DOI 10.3934/jcd.2014.1.391.

[33] Kalantari M., *Chaos Soliton. Fract.*, 2021; **142**: 110547. DOI 10.1016/j.chaos.2020.110547.