Terr. Atmos. Ocean. Sci., Vol. 20, No. 2, 365-375, April 2009 doi: 10.3319/TAO.2008.02.25.01(A)

Alternative Approaches to Estimating the Linear Propagator and Finite-Time Growth Rates from Data

Yung-An Lee* and Zih-Ting Hung

Institute of Atmospheric Physics, National Central University, Chung-Li, Taiwan, ROC

Received 3 September 2007, accepted 25 February 2008

ABSTRACT

The propagator of a linear model plays a central role in empirical normal mode and finite-time instability problems. Its estimation will affect whether the linear stability characteristics of the corresponding dynamic system can be properly extracted. In this study, we introduce two alternative methods for estimating the linear propagator and finite-time growth rates from data. The first is the generalized singular value decomposition (GSVD) and the second is the singular value decomposition combined with the cosine-sine decomposition (SVD-CSD). Both methods linearize the relation between the predictor and the predictand by decomposing them to have a common evolution structure and then make the estimations. Thus, the linear propagator and the associated singular vectors can be simultaneously derived. The GSVD clearly reveals the connection between the finite-time amplitude growth rates and the singular values of the propagator. However, it can only be applied in situations when given data have more state variables than observations. Furthermore, it generally encounters an over-fitting problem when data are contaminated by noise. To fix these two drawbacks, the GSVD is generalized to the SVD-CSD to include data filtering capability. Therefore, it has more flexibility in dealing with general data situations. These two methods as well as the Yule-Walker equation were applied to two synthetic datasets and the Kaplan's sea surface temperature anomalies (SSTA) to evaluate their performance. The results show that, because of linearization and flexible filtering capabilities, the propagator and its associated properties could be more accurately estimated with the SVD-CSD than other methods.

Key words: Generalized singular value decomposition, Finite-time instability, Linearization

Citation: Lee, Y. A. and Z. T. Hung, 2009: Alternative approaches to estimating the linear propagator and finite-time growth rates from data. Terr. Atmos. Ocean. Sci., 20, 365-375, doi: 10.3319/TAO.2008.02.25.01(A)

1. INTRODUCTION

Waves are common phenomena in the atmosphere and the ocean. The mechanisms of the generation, evolution, and predictability of such phenomena are strongly dependent on the stability characteristics of these waves. Therefore, the ability to correctly extract wave stability characteristics from observed data is very crucial to our understanding of the dynamics of the corresponding systems. Currently, empirical stability studies are mainly carried out through linear inverse modeling of the observed spatial-temporal data. In other words, one fits an observed dataset that has m variables of $n + \tau$ temporal length to a discrete form of the linear advection equation,

$$\mathbf{Y}(t + \tau) = \mathbf{A}\mathbf{Y}(t) \quad (1)$$

where $\mathbf{Y}(t)$ ^{m} ^{n} is the first n temporal length of the m variables (predictor matrix), $\mathbf{Y}(t + \tau)$ is the same variables at τ period later (predictand matrix), and $\mathbf{A}$ is the linear propagator. The principal oscillation pattern analysis (POP; von Storch et al. 1995; von Storch and Zwiers 1998) applies an eigenvalue-eigenfunction analysis to the linear propagator to extract the dominant normal mode oscillations and the associated stability characteristics of the corresponding system. However, normal mode instability is not the only instability mechanism that operates in a fluid system. The interactions among waves may also cause amplitudes of some perturbations to growth temporarily even in a system that is normal mode stable (Farrell and Ioannou 1996a, b). Because most observational data are

* *Corresponding author*
E-mail: yalee@atm.ncu.edu.tw

nearly stationary, it is hard to associate any normal mode exponential growth behavior with them. Therefore, the finite-time instability is widely accepted to play an equal or more important role as the normal mode instability in cyclone genesis, predictability, and data assimilation studies. In such studies, the singular value decomposition (SVD; Golub and Van Loan 1996) is applied to the linear propagator to extract the dominant singular vectors and the associated finite-time growth rates. Because the linear propagator plays a central role in empirical normal mode and finite-time instability problems, how to estimate it from the data will determine whether the stability characteristics of the corresponding system can be properly extracted.

Conventionally, the linear propagator **A** is estimated using the Yule-Walker equation (Brockwell and Davies 1991):

$$\mathbf{A} = \{\mathbf{Y}(t+\tau)\,\mathbf{Y}^T(t)\}\,\{\mathbf{Y}(t)\,\mathbf{Y}^T(t)\}^{-1} \quad (2)$$

where superscripts T and -1 denote the matrix transpose and matrix inverse respectively. However, the predictor and the predictand are generally not perfectly linearly correlated; i.e., their lag τ autocorrelations are always less than 1. Consequently, the linear propagator so derived generally underestimates the linear instability of the corresponding dynamic system. For example, it is well known that POP analysis can yield only damped oscillation patterns (von Storch et al. 1995). This means that using the Yule-Walker equation to estimate the propagator, the observed variability cannot be fully maintained and Eq. (1) becomes a stable linear system. Because of this, some kind of forcing term must be added to Eq. (1) to maintain the observed variability in the linear inverse model (Penland and Magorian 1993; Penland and Sardeshmukh 1995) and the Markov model (Xue et al. 1994; Xue et al. 2000). Therefore, it is desirable to know whether an initial linearization of the original data, before the estimation of the linear propagator, can allow the linear stability characteristics of the corresponding system to be more properly extracted. Furthermore, although it is well known that the singular values of a linear propagator represent the non-modal averaged amplification or decaying rates (finite-time growth rates) of the corresponding linear dynamic system in τ period, the connection between them is generally established through some predefined matrix norms. Various studies (Palmer et al. 1998; Kim and Morgan 2002) have shown that different choices of norms will yield quite different singular vectors. This seemingly non-uniqueness of singular vectors has added considerably complexity to the dynamic interpretation of the corresponding system. Therefore, it is also desirable to establish a more direct and clearer connection between the singular values of a linear propagator and the finite-time growth rates.

In this study, we introduce two alternative approaches, which originate from the field of matrix computation, to estimate the propagator, finite-time growth rates and the associated singular vectors from the data. Rather than using the autocovariance and the lag-τ covariance matrices of the observed data to estimate the propagator, these approaches first linearize the relation between the predictor and the predictand by decomposing them to have the same common evolution matrix, and then make the estimations. Furthermore, these approaches establish a direct connection between the singular values of a linear propagator and the finite-time growth rates of the predictand singular vectors. The remainder of this paper is organized as follows. Section 2 describes the methodology and properties of the generalized singular value decomposition (GSVD; Golub and Van Loan 1996). Section 3 describes the methodology and properties of the singular value decomposition (SVD) combined with the cosine-sine decomposition (CSD; Golub and Van Loan 1996). To compare and evaluate the relative advantages of different approaches, the Yule-Walker equation, the GSVD and the SVD-CSD are applied to two synthetic datasets and an observed dataset to estimate the corresponding linear propagators and the associated properties. Section 4 shows the results for two synthetic datasets consisting of 10 normal mode oscillations plus two different levels of noise. Section 5 shows results for the monthly mean sea surface temperature anomalies (SSTA, Kaplan et al. 1998). Section 6 discusses and summarizes this study.

2. THE GENERALIZED SINGULAR VALUE DECOMPOSITION METHOD (GSVD)

In the matrix computation field, there is a well-known theorem called generalized singular value decomposition (GSVD; theorem 8.74 in Golub and Van Loan 1996). It states that, if $\mathbf{F}$ $^{k\ q}$, $\mathbf{G}$ $^{p\ q}$ and $k, p \quad q$, then there exist two orthogonal matrices, $\mathbf{U}$ $^{k\ k}$ and $\mathbf{V}$ $^{p\ p}$, and an invertible matrix, $\mathbf{X}$ $^{q\ q}$, such that:

$$\mathbf{U}^T\mathbf{F}\mathbf{X} = \mathbf{C} = diag(c_1, \quad, c_q) \quad 0 \quad c_1 \quad c_2 \qquad c_q \quad 1 \quad (3)$$

$$\mathbf{V}^T\mathbf{G}\mathbf{X} = \mathbf{S} = diag(s_1, \quad, s_q) \quad 1 \quad s_1 \quad s_2 \qquad s_q \quad 0 \quad (4)$$

where $\mathbf{C}^T\mathbf{C} + \mathbf{S}^T\mathbf{S} = \mathbf{I}_q$ and $\mathbf{I}_q$ is a $q \quad q$ identity matrix. Therefore, a generalized eigenvalue problem (i.e., $\mathbf{G}\mathbf{x} = \lambda\mathbf{F}\mathbf{x}$), can be directly solved without the need for forming $\mathbf{G}^T\mathbf{G}$ and $\mathbf{F}^T\mathbf{F}$. Based on this theorem, the proposed GSVD method for estimating the linear propagator and the associated properties can be described as follows. For the linear model of Eq. (1), if $m \quad n$, the predictor and the predictand matrices can be decomposed using the GSVD as:

$$\mathbf{Y}(t) = \mathbf{U}\mathbf{C}\mathbf{X}^{-1} \quad (5)$$

$$\mathbf{Y}(t+\tau) = \mathbf{V}\mathbf{S}\mathbf{X}^{-1} \quad (6)$$

Because $\mathbf{Y}(t)$ and $\mathbf{Y}(t+\tau)$ have been rearranged to have the same temporal evolution structure, $\mathbf{X}^{-1}$ n n, the correlation coefficient between the time series of each corresponding column vector of $\mathbf{U}$ m m and $\mathbf{V}$ m m becomes unity. Therefore, the use of GSVD linearizes the relation between the predictor and the predictand. The substitution of Eqs. (5) and (6) into Eq. (1) with some matrix manipulations then yields:

$$\mathbf{A} = \mathbf{VSC}^{-1}\mathbf{U}^T = \mathbf{V}\ \mathbf{U}^T \tag{7}$$

where

$$= diag(\sigma_1, \quad, \sigma_n) = \mathbf{SC}^{-1} = diag(s_1/c_1, \quad, s_n/c_n) \tag{8}$$

is a diagonal matrix.

From Eqs. (5), (6), and (7), it is clear that the linear propagator can be directly derived from the GSVD decomposition of the predictor and the predictand without the need to form $\mathbf{Y}(t)\ \mathbf{Y}^T(t)$ and $\mathbf{Y}(t+\tau)\mathbf{Y}^T(t)$ first. Furthermore, it is noted that the left and right singular vectors derived from the SVD of a given matrix are unique up to multiplication of a column of the left singular vectors by a unit phase factor and simultaneous multiplication of the corresponding column of the right singular vectors by the same unit phase factor if all the singular values are non-degenerate and non-zero. As Eq. (7) is also the SVD of **A**, therefore if the same conditions are met, the use of GSVD decomposition allows the singular values and the left and right singular vectors of the propagator to be simultaneously derived. Because all the singular values derived from observed data are generally distinct and non-zero, therefore the linear propagator and the associated singular vectors derived from the GSVD are also generally unique. The uniqueness of the GSVD results can also be readily checked by direct calculation. For example, one can randomly generate a linear propagator **A**, then use Eq. (1) and an arbitrary initial condition to generate the corresponding $\mathbf{Y}(t)$ and $\mathbf{Y}(t+\tau)$. By applying the SVD to **A** and the GSVD to $\mathbf{Y}(t)$ and $\mathbf{Y}(t+\tau)$, one can easily see that the GSVD derived linear propagator and singular values are the same as **A** and its associated singular values, while the GSVD derived singular vectors are also the same as those of **A** except for possible sign reversal among some of the vectors.

Moreover, we note in Eq. (8) that the i-th singular value σ_i is expressed in terms of the ratio between s_i and c_i, which are the normalized amplitude measures of the i-th singular vector of $\mathbf{Y}(t+\tau)$ and $\mathbf{Y}(t)$, respectively. Therefore, the i-th singular value of the propagator directly and clearly represents the non-modal averaged amplitude amplification or decay rates of the corresponding linear system in period in terms of the L2 norm of the state vector, $\mathbf{Y}(t)$. Because different variables generally have different spectral characteristics, it is likely that using different state variable to describe the same linear system will yield somewhat different linear propagators and associated singular vectors. This feature may partially explain why different choices of norms in previous studies yielded quite different singular vectors. Another significant feature of the method is the exact equality between $\mathbf{Y}(t+\tau)$ and $\mathbf{AY}(t)$; i.e., $\mathbf{Y}(t+\tau) = \mathbf{VSX}^{-1} = \mathbf{VSC}^{-1}\mathbf{U}^T\mathbf{UCX}^{-1} = \mathbf{AY}(t)$. This implies that the variability of $\mathbf{Y}(t+\ )$ can be completely reproduced by $\mathbf{AY}(t)$. Hence, when using the GSVD to estimate the linear propagator, there is no need to add any unknown forcing term to Eq. (1) to maintain the observed variability.

3. THE SINGULAR VALUE DECOMPOSITION COMBINED WITH THE COSINE-SINE DECOMPOSITION METHOD (SVD-CSD)

The above derivation shows that, when data have more grid points than observations, the GSVD provides a direct way to estimate the linear propagator and its associated singular vectors. However, the GSVD theorem requires that k, p q, therefore it cannot be applied to data with fewer variables than observations (i.e., $m < n$). More importantly, all data are more or less contaminated by noise. Therefore, when m n, both the GSVD and the Yule-Walker equation are very likely to over-fit the linear relation between the predictor and the predictand and lead to wrong estimations. To prevent such an over-fitting problem, it is a common practice to use only the dominant principal components (PCs) from the principal component analysis (PCA; Preisendorfer and Mobley 1988) of the original data to estimate the linear propagator. In such cases, the number of retained PCs will generally be much less than the observations. Therefore, we need to develop another approach to estimate the linear propagator and the associated properties for $m < n$ cases.

The SVD-CSD method is basically an extension of the GSVD method. In line with the derivation of the GSVD theorem, we first linearize $\mathbf{Y}(t)$ and $\mathbf{Y}(t+\tau)$, by requiring that they have the same evolution structure. Consequently, the SVD is applied to their joint matrix to yield:

$$\begin{bmatrix} \mathbf{Y}(t) \\ \mathbf{Y}(t+\tau) \end{bmatrix} = \mathbf{Q\Lambda R}^T = \begin{bmatrix} \mathbf{Q}_{d1}\Lambda_d\mathbf{R}_d^T \\ \mathbf{Q}_{d2}\Lambda_d\mathbf{R}_d^T \end{bmatrix} + \begin{bmatrix} \mathbf{Q}_{r1}\Lambda_r\mathbf{R}_r^T \\ \mathbf{Q}_{r2}\Lambda_r\mathbf{R}_r^T \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_d(t) \\ \mathbf{Y}_d(t+\tau) \end{bmatrix} + \begin{bmatrix} \mathbf{Y}_r(t) \\ \mathbf{Y}_r(t+\tau) \end{bmatrix} \tag{9}$$

where $\mathbf{Q}$ $2m$ $2m$ and $\mathbf{R}$ n n are orthogonal, $\mathbf{\Lambda} = diag(\lambda_1, \quad, \lambda_f)$, and f min $(2m, n)$. If m n, the cosine-sine decomposition (CSD; theorem 2.6.2 in Golub and Van Loan 1996) can be used directly to decompose **Q**, yielding the same decompositions of $\mathbf{Y}(t)$ and $\mathbf{Y}(t+\tau)$ as those for Eqs. (5) and (6). However, the CSD can not be directly applied to **Q** when $m < n$. In such cases, **Q**, $\mathbf{\Lambda}$, and **R** need to be de-

composed into dominant (subscript d) and remainder (subscript r) submatrices, based upon how many singular vectors need to be extracted or retained. Because the number of retained singular vectors (N) cannot exceed the number of state variables (m), for a given N m, then $\mathbf{Q}$, $\mathbf{\Lambda}$, and $\mathbf{R}$ in Eq. (9) are further decomposed into $\mathbf{Q}_{d1}$, $\mathbf{Q}_{d2}$ $^{m\ N}$, $\mathbf{Q}_{r1}$, $\mathbf{Q}_{r2}$ $^{m\ (2m\ N)}$, $\mathbf{\Lambda}_d = diag\,(\lambda_1, \lambda_N)$, $\mathbf{\Lambda}_r = diag\,(\lambda_{N+1}, , \lambda_f)$, $\mathbf{R}_d$ $^{n\ N}$, and $\mathbf{R}_r$ $^{n\ (n\ N)}$. In terms of these submatrices, $\mathbf{Y}(t)$ and $\mathbf{Y}(t + \tau)$ can also be separated into $\mathbf{Y}_d(t)$, $\mathbf{Y}_r(t)$ and $\mathbf{Y}_d(t + \tau)$, $\mathbf{Y}_r(t + \tau)$, respectively. The orthogonal relation between $\mathbf{R}_d$ and $\mathbf{R}_r$ implies that $\mathbf{Y}_d(t)$ and $\mathbf{Y}_d(t + \tau)$ are also orthogonal to $\mathbf{Y}_r(t)$ and $\mathbf{Y}_r(t + \tau)$. Hence, Eq. (1) can be divided into a dominant equation and a remainder equation, i.e.,

$$\mathbf{Y}_d(t + \tau) = \mathbf{A}_d\mathbf{Y}_d(s) \tag{10}$$

$$\mathbf{Y}_r(t + \tau) = \mathbf{A}_r\mathbf{Y}_r(s) \tag{11}$$

Due to the fact that $\mathbf{Y}_d(t)$ and $\mathbf{Y}_d(t + \tau)$ represent the dominant linear covariability of $\mathbf{Y}(t)$ and $\mathbf{Y}(t + \tau)$, Eq. (10) can be viewed as a filtered version of Eq. (1). The CSD is then used to decompose $\mathbf{Q}_{d1}$ and $\mathbf{Q}_{d2}$ to become:

$$\begin{bmatrix}\mathbf{Q}_{d1}\\ \mathbf{Q}_{d2}\end{bmatrix} = \begin{bmatrix}\mathbf{U}_{d1}\mathbf{C}_d\mathbf{V}_d^{\,T}\\ \mathbf{U}_{d2}\mathbf{S}_d\mathbf{V}_d^{\,T}\end{bmatrix} \tag{12}$$

where $\mathbf{U}_{d1}$, $\mathbf{U}_{d2}$ $^{m\ m}$, $\mathbf{V}_d$ $^{N\ N}$ are orthogonal matrices, $\mathbf{C}_S$ and $\mathbf{S}_S$ diagonal with $\mathbf{C}_d^{\,T}\mathbf{C}_d + \mathbf{S}_d^{\,T}\mathbf{S}_d = \mathbf{I}_N$. With these decompositions, $\mathbf{Y}_d(t)$, $\mathbf{Y}_d(t + \tau)$ and $\mathbf{A}_d$ can be written, similar to Eqs. (5), (6), and (7), as:

$$\mathbf{Y}_d(t) = \mathbf{U}_{d1}\mathbf{C}_d\mathbf{V}_d^{\,T}\mathbf{\Lambda}_d\mathbf{R}_d^{\,T} = \mathbf{U}_{d1}\mathbf{C}_d \quad \mathbf{X}_d^{\,1} \tag{13}$$

$$\mathbf{Y}_d(t + \tau) = \mathbf{U}_{d2}\mathbf{S}_d\mathbf{V}_d^{\,T}\mathbf{\Lambda}_d\mathbf{R}_d^{\,T} = \mathbf{U}_{d2}\mathbf{S}_d\mathbf{X}_d^{\,1} \tag{14}$$

and

$$\mathbf{A}_d = \mathbf{U}_{d2}\mathbf{S}_d\mathbf{C}_d^{\,1}\mathbf{U}_{d1}^{\,T} \tag{15}$$

Consequently, when data have more state variables than observations, the SVD-CSD still allows for a direct estimation of the linear propagator and the associated properties from the filtered predictor $[\mathbf{Y}_d(t)]$ and predictand $[\mathbf{Y}_d(t + \tau)]$. Similarly, because the variability of $\mathbf{Y}_d(t + \tau)$ can be completely reproduced by $\mathbf{A}_d\mathbf{Y}_d(t)$, one also does not need to add any unknown forcing term $\mathcal{N}$ to Eq. (10) to maintain the observed variability.

The SVD-CSD can recover the GSVD results when $m > n$ and can be applied to $m < n$ cases. Therefore, although slightly more complicated to implement, the SVD-CSD has a wider range of applicability. Furthermore, if some kind of filtering needs to be applied to the original data, the SVD-CSD has two options to choose from. The first is the number of retained PCs (m); the second is the number of retained singular vectors (N). The use of a specific m to filter data, as in the conventional data analysis, is based on the explained variance of the original data. However, this kind of data filtering does not differentiate whether the predictor and the predictand are linearly related or not. Therefore, some important linear covariability between the predictor and the predictand may be filtered out. On the other hand, if all the PCs of the original data are retained, the SVD-CSD can still allow data filtering by selecting $N < m$. In such cases, one can assure that the dominant linear covariability between the predictor and the predictand will not be unintentionally filtered out. As for the Yule-Walker equation, because the retained singular modes must be the same as the number of state variables, data filtering can only be applied by choosing a subset of PCs (i.e., choosing a specified m and $N = m$). Therefore, the SVD-CSD offers a more selective data filtering mechanism for the estimation of the linear propagator and the associated properties.

4. THE SYNTHETIC NORMAL MODE OSCILLATIONS

In the above derivations, we learned that both the GSVD and the SVD-CSD are capable of deriving the linear propagator and the associated properties directly from the predictor and the predictand matrices. However, their performance in estimating the linear propagator and the associated singular values should be compared with the conventional Yule-Walker equation to evaluate their usefulness. In this section, we apply all these methods to two synthetic datasets with known normal mode growth rates to see if they can faithfully extract the stability characteristics from the data. These synthetic datasets are constructed as a superposition of 10 normal mode oscillations contaminated by two different levels of random noise on the global tropics between 30 S - 30 N with 5 5 resolution, that is:

$$z(x, y, t) \quad \sum_{j\ 1}^{10} \cos(k_j x \quad l_j y \quad \sigma_j t)e^{\lambda_j t} \quad \mathcal{N}(0, \alpha) \tag{16}$$

where $\mathcal{N}(0, \alpha)$ denotes random noise with zero mean and standard deviation $\alpha = 0.1$ or $\alpha = 1$, k_j $\frac{j\pi}{180}$, $l_j = \frac{[(1\quad 2(j\quad 1)]\pi}{180}$, σ_j $\frac{[1\quad 3(j\quad 1)]\pi}{600}$ with $j = 1, \ldots, 10$, $\lambda_j = -0, -0.001, , -0.009$ is the growth rate of the jth normal mode, $x = 0, 5, , 350, 355$, $y = -30, -25, , 25, 30$ and $t = 1, 2, , 600$. It is noted that each of the dataset consists of 936 time series ($m = 936$), with 600 time steps ($n = 600$) in each series, therefore the GSVD restriction is satisfied (i.e., m n). Furthermore, because these 10 modes are

orthogonal and each λ_j is small, the finite-time growth rate can be approximately taken as $1 + \lambda_j\tau$ if τ is also small. Therefore, the estimated finite-time growth rates from the GSVD, the SVD-CSD and the Yule-Walker equation can be compared to these known values to evaluate their validity.

Results from the weakly random noise contaminated synthetic dataset (i.e., $\alpha = 0.1$) are shown in Figs. 1 and 2. In Fig. 1a, one clearly notes that the explained variance of the first 20 PCs are well above those of the rest of PCs. Because the signal part of the synthetic time series constitutes 10 normal mode oscillations and each normal mode can be decomposed into a pair of equal variance PCs, this result indicates that the PCA is very effective in differentiating signal from noise in the weakly random noise contaminated data. Figure 1b shows the estimated lag-1 ($\tau = 1$) finite-time growth rates using both the GSVD and the Yule-Walker equation methods. Because the GSVD and the Yule-Walker equation yield $\mathbf{A}^{600\ 600}$ and $\mathbf{A}^{936\ 936}$, respectively, the maximum number of singular vectors can be extracted from them are 600 and 936, respectively. However, the Yule-Walker equation needs to calculate the inverse of the auto covariance matrix. When data have fewer observations than grid points, the resultant auto covariance matrix is singular. Hence, with the Yule-Walker equation, one encounters the rank deficiency problem in matrix inversion and yields unrealistically large finite-time growth rates compared to the GSVD. Conversely, results from the GSVD showed a nearly anti-symmetrical distribution between the growing and decaying modes. Because **C** and **S**, as shown in Eqs. (5) and (6), are arranged in ascending and descending order between 0 and 1, respectively, this anti-symmetrical distribution is what one would expect when the variability of $\mathbf{Y}(t + 1)$ can be faithfully reproduced by $\mathbf{AY}(t)$. Nevertheless, because the GSVD does not have any data filtering mechanism to differentiate signal from noise, it inevitably overestimates the linear relation between the predictor and the predictand. Therefore it is also incapable of yielding the correct results (i.e., the solid line significantly deviates from the diamond symbols).

Figure 2 shows the SVD-CSD and the Yule-Walker equation estimated lag-1 growth rates for PCA filtered data. The SVD-CSD results are calculated using 20 retained PCs ($m = 20$) and retained singular vectors, $N = 5, 10, 15, 20$, respectively. The Yule-Walker equation results are calculated using retained PCs $m = 5, 10, 15, 20$, respectively. One observes clearly that the Yule-Walker equation always overestimates the decay rates, while the SVD-CSD yields results overall quite similar to the true decay rates. However, when

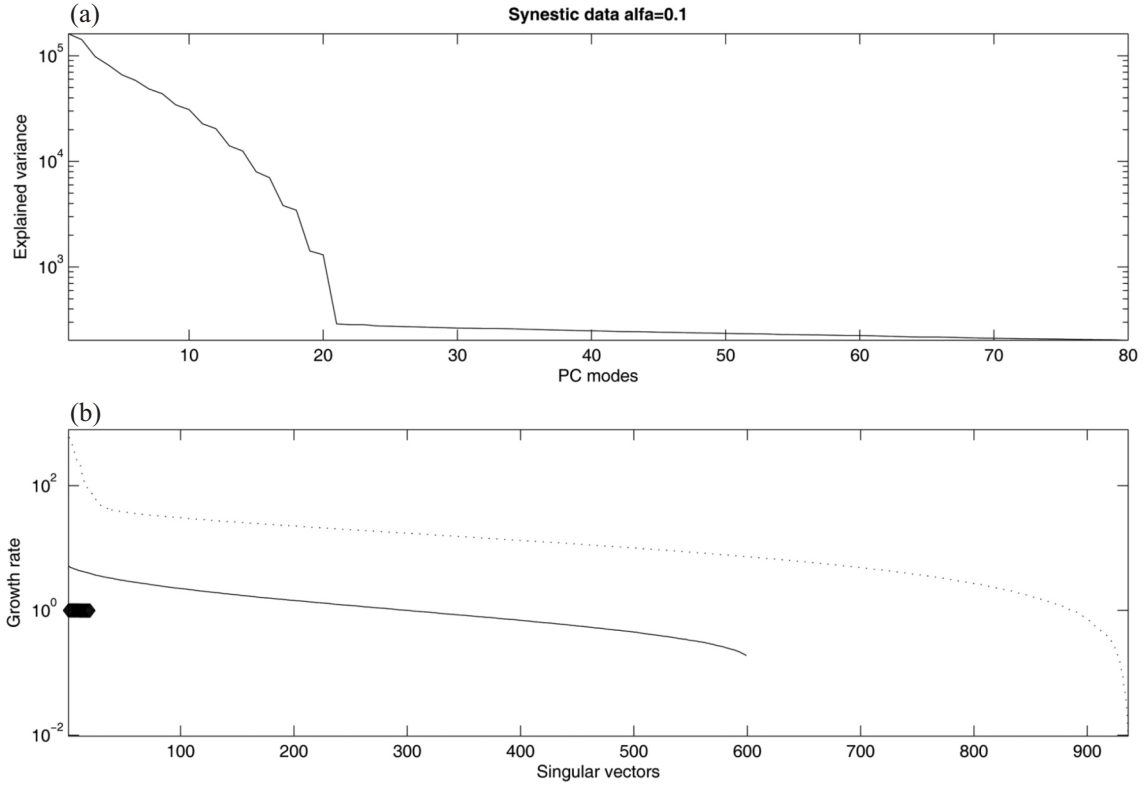

Fig. 1. The PCA and the estimated lag-1 ($\tau = 1$) finite-time growth rates results of the weakly random noise contaminated synthetic dataset (standard deviation $\alpha = 0.1$). (a) is the explained variance as a function of PC modes from PCA of the data. (b) is the estimated finite-time growth rates as a function of the singular vectors using the Yule-Walker equation (dashed line) and the GSVD (thick solid line), respectively. The diamond symbols represent the lag-1 finite-time growth rates derived from the normal mode growth rates of the synthetic data (i.e., $1 + \lambda_j$).

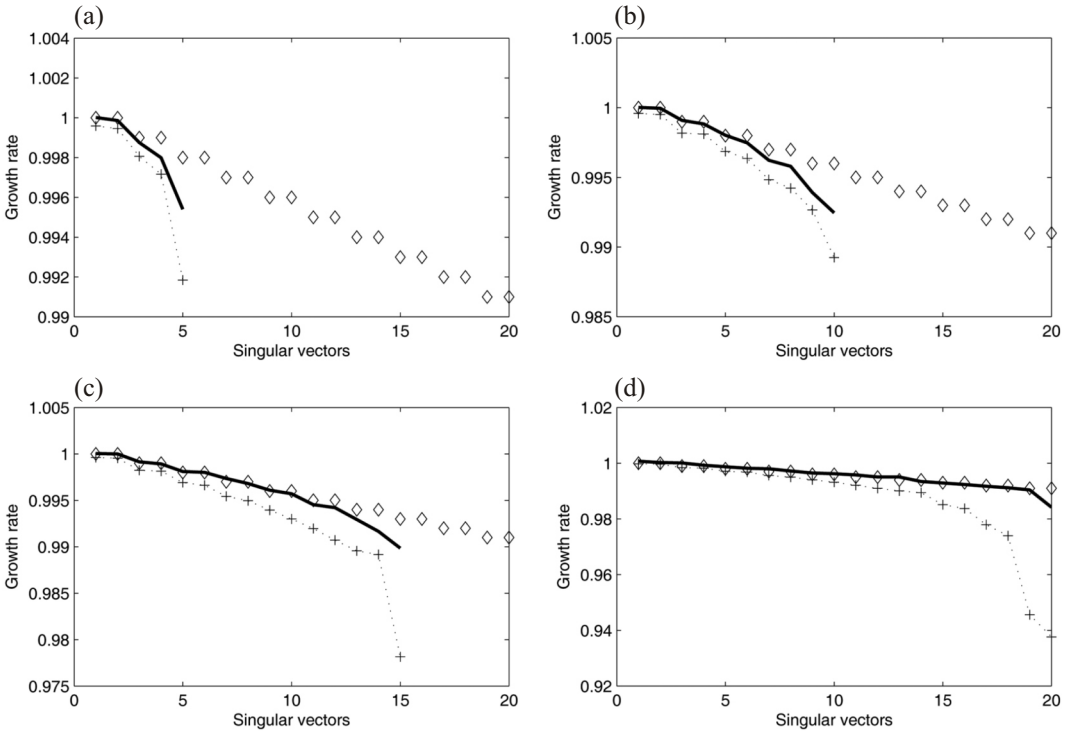

Fig. 2. The estimated lag-1 finite-time growth rates results of the weakly random noise contaminated synthetic dataset using both the SVD-CSD (thick solid line) and the Yule-Walker equation (dashed line) methods. (a), (b), (c), and (d) correspond to results from the SVD-CSD using 20 PCs (m = 20) and retained singular vectors N = 5, 10, 15, 20 and results from the Yule-Walker equation using 5, 10, 15, and 20 PCs (m = 5, 10, 15, 20), respectively. Similarly, The diamond symbols in each panel represent the lag-1 finite-time growth rates derived from the normal mode growth rates of the synthetic data (i.e., $1 + \lambda_j$).

the retained singular vectors are equal to the retained PCs (i.e., $m = N = 20$), the growth rates of the first few singular vectors tend to be overestimated. Note that the retained PCs, although being PCA filtered, are still not noise-free. When $m = N$, because the SVD-CSD is unable to apply further filtering to the data, the overestimation of the linear relation between the predictor and the predictand inevitably leads to the overestimation of the growth rates.

Figures 3 and 4 show results from the moderately random noise contaminated synthetic dataset (i.e., α = 1). In Fig. 3a, one notes that all PCs of the α = 1 data have greater explained variance than those of the α = 0.1 data. Furthermore, signal and noise parts of the spectra are not as clearly separated as those of the α = 0.1 case. These results indicate that the PCA is ineffective in differentiating signal from noise in moderately random noise contaminated data. Similar to Fig. 1b, Fig. 3b also shows clearly that both the GSVD and the Yule-Walker equation methods are unable to estimate the finite-time growth rates correctly from original data. These results strongly suggest that one always needs to apply some kinds of data filtering to correctly estimate the linear propagator and the associated properties.

Figure 4 shows the same as Fig. 2, except for results of the α = 1 case. One notes that the overestimation of the decay rates by the Yule-Walker equation is more serious than that in the α = 0.1 case. On the other hand, the SVD-CSD still yields results quite similar to the true decay rates for retained singular vectors up to 15. When $m = N = 20$, because the noise level is relatively high and the SVD-CSD is unable to apply further filtering to the data, the overestimation of the linear relation between the predictor and the predictand then leads to a more serious overestimation of the growth rates for the first few singular vectors. These results clearly suggest that the selective data filtering capability of SVD-CSD can yield a more correct estimation of the propagator and the associated finite-time growth rates than can the Yule-Walker equation. It is noted that similar analyses were also applied to various values of τ. Their results (not shown) are quite similar to those of τ = 1. Therefore, the above conclusions are valid not just for τ = 1 only.

5. SEA SURFACE TEMPERATURE ANOMALIES

In this section, all three methods are applied to Kaplan's SSTA data (Kaplan et al. 1998) to evaluate their performance. Kaplan's SSTA data consist of the global gridded (5 5) monthly mean SSTA from January 1856 to March 2003 (they can be found in http://ingrid.ldeo.columbia.edu/SOURCES/.KAPLAN/.EXTENDED). It was constructed using optimal estimation in the space of 80 PCs to interpo-

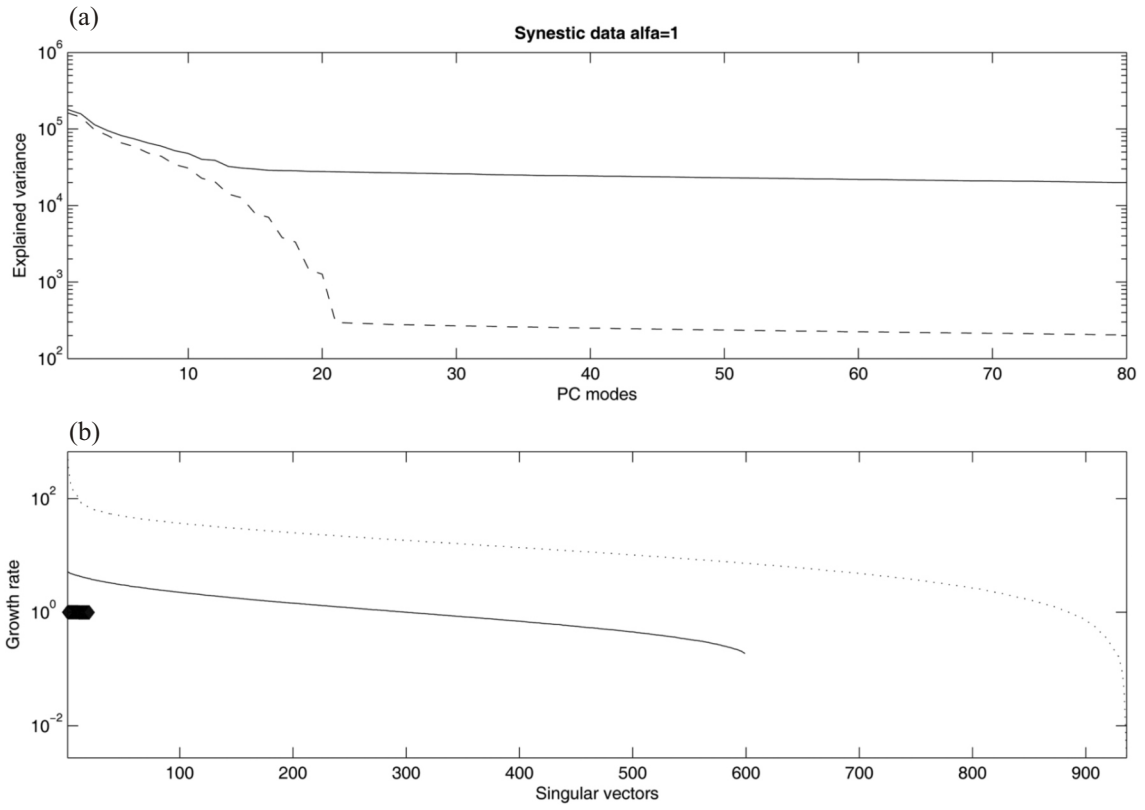

Fig. 3. The same as Fig. 1, except for results of the moderately random noise contaminated synthetic data (standard deviation $\alpha = 1$). For comparison, the PCA result of the $\alpha = 0.1$ (dashed line) is also shown in (a).

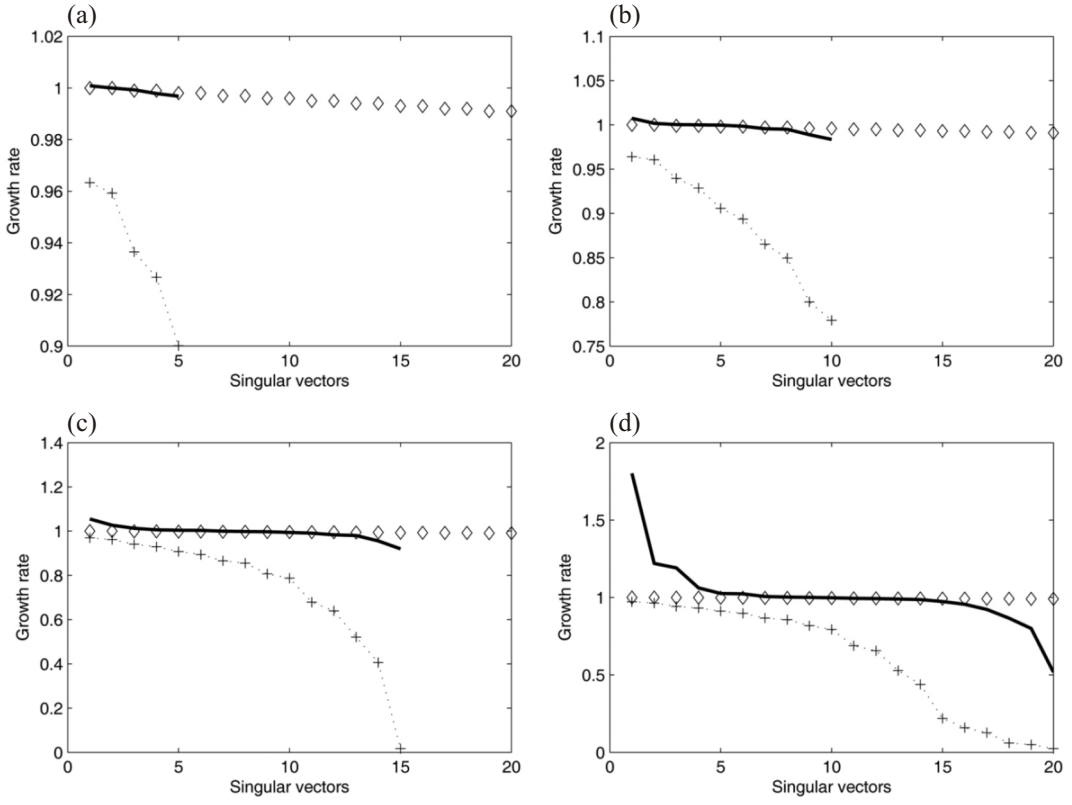

Fig. 4. The same as Fig. 2, except for results of the moderately random noise contaminated synthetic data (standard deviation $\alpha = 1$).

late ship observations from the UK Met Office database (Parker et al. 1994). The data after 1981 represents the projection of the NCEP OI analysis (which combines ship observations with remote sensing data) by Reynolds and Smith (1994) on the same set of 80 PCs. In this study, only tropical ocean (27.5 S - 27.5 N) data between the years 1950 and 2002 are used.

The results (not shown) from applying the GSVD and the Yule-Walker equation to the original data are similar to those in Fig. 1 and still are unable to yield correct estimations. The estimated maximum lag-1 finite-time growth rates for the PCA filtered Kaplan's SSTA from both the SVD-CSD (m = 80) and the Yule-Walker equation as a function of the retained singular modes are shown in Fig. 5. Results from both the Yule-Walker equation and the SVD-CSD show monotonically increasing growth rates with retained singular vectors. Nevertheless, those from the Yule-Walker equation have lower growth rates than those for the SVD-CSD. As can be seen in Fig. 5b, no instability can be found in the results from the Yule-Walker equation for retained singular vectors less than 4. Does the Yule-Walker equation underestimate the growth rates? Figure 6 shows the estimated finite-time growth rates for the PCA filtered Kaplan's SSTA from both the SVD-CSD and the Yule-Walker equation as a function of the singular vectors for (a) 5, (b) 10, (c) 15, and (d) 20 retained singular vectors, respectively. Note that the growth rate curves from the Yule-Walker equation are predominantly asymmetrical towards the decay states (growth rates < 1), while those from the SVD-CSD are more anti-symmetrical about the neutral state (growth rate = 1). The results of the Yule-Walker equation indicate that the variability of $\mathbf{Y}(t + 1)$ cannot be well maintained by $\mathbf{AY}(t)$. To show this is indeed the case, the ratios between the total variance of $\mathbf{AY}(t)$ and $\mathbf{Y}(t + 1)$ are calculated. They are 0.96, 0.94, 0.93, and 0.92, respectively. These results show the variability of $\mathbf{Y}(t + 1)$ is not fully recovered by $\mathbf{AY}(t)$ when the Yule-Walker equation is used to estimate $\mathbf{A}$. Furthermore, the damping of the variability of $\mathbf{Y}(t + 1)$ increases as more PCs are used. Conversely, the anti-symmetrical distribution results of the SVD-CSD are what one would expect when the variability of $\mathbf{Y}_d(t + 1)$ can be faithfully reproduced by $\mathbf{A}_d\mathbf{Y}_d(t)$. These results clearly suggest that, due to the incapability of differentiating linear related and unrelated variability between the predictor and the pre-

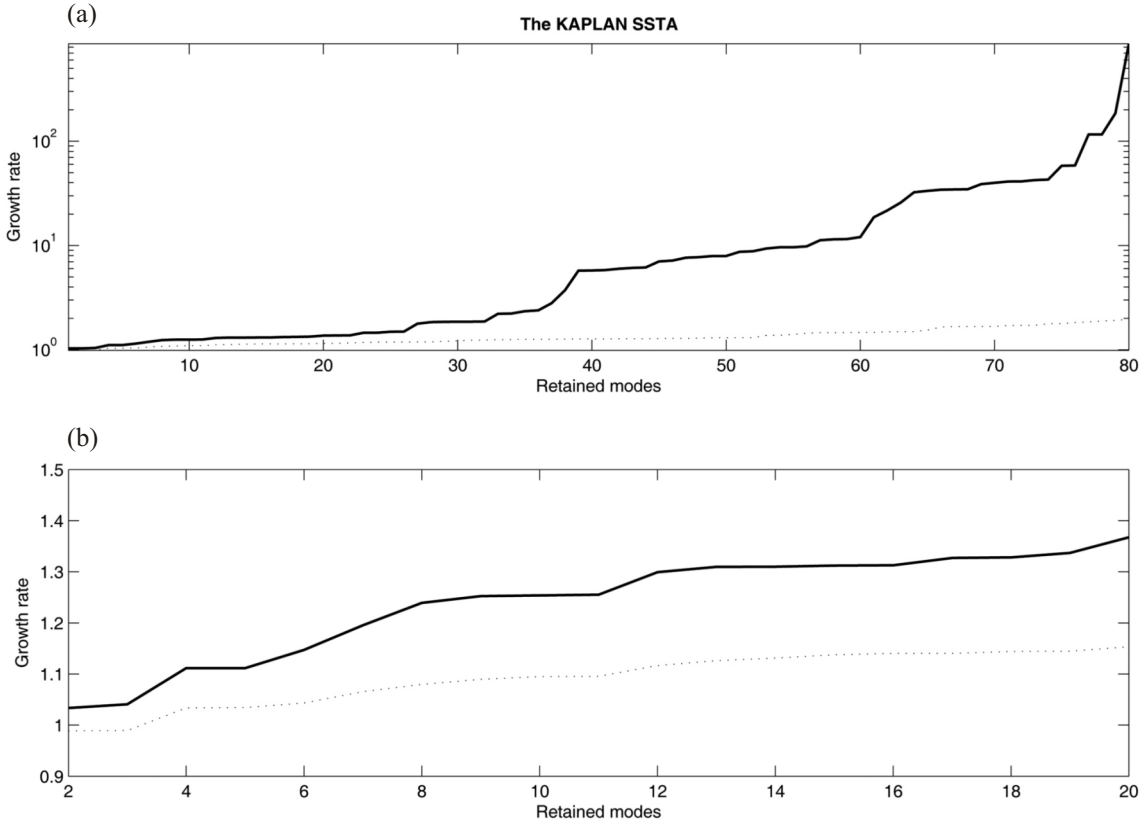

Fig. 5. The estimated maximum lag-1 finite-time growth rates for the PCA filtered Kaplan SSTA as a function of the retained singular vectors using both the SVD-CSD and the Yule-Walker equation methods. The thick solid line corresponds to the results obtained by applying the SVD-CSD to the first 80 PCs. The dashed line corresponds to the results obtained by applying the Yule-Walker equation to the same number of PCs as used for the retained singular vectors. (a) shows results for all retained singular vectors. (b) shows a blown-up view of (a), for the first 20 retained singular vectors.

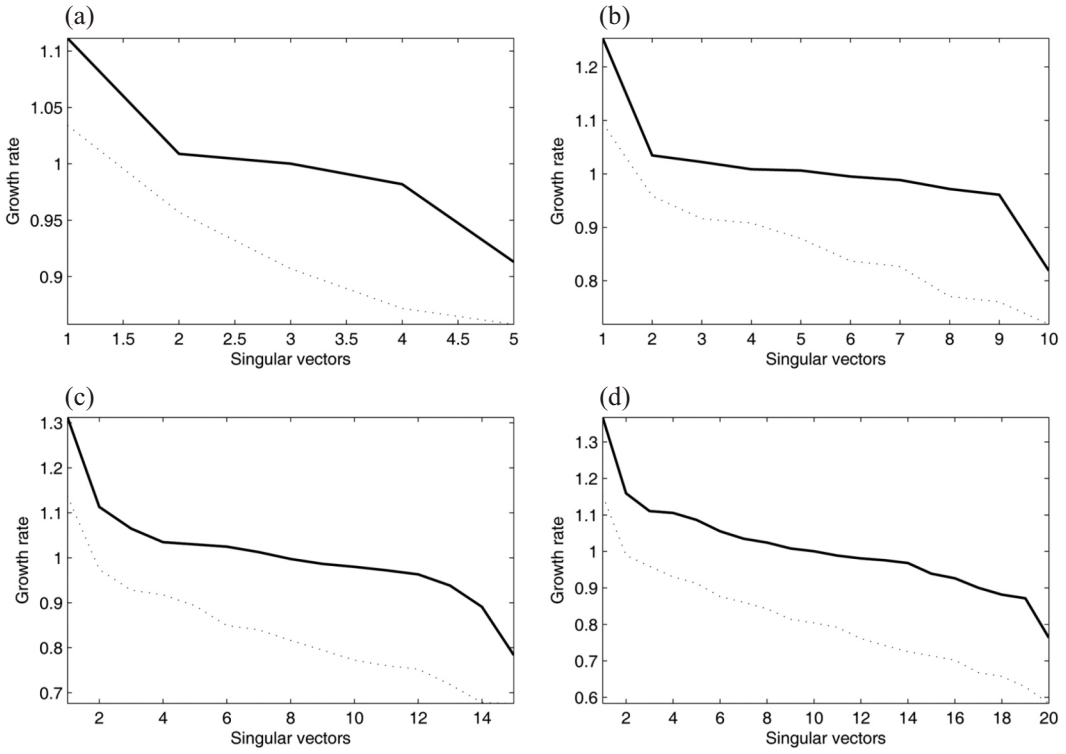

Fig. 6. The estimated finite-time growth rates for the PCA filtered Kaplan SSTA as a function of the singular vectors for: (a) 5, (b) 10, (c) 15, and (d) 20 retained singular vectors, respectively. The thick solid and dashed lines correspond to the results obtained by applying the SVD-CSD to the first 80 PCs, and from the Yule-Walker equation.

dictand, the Yule-Walker equation may overly underestimate the growth rates. As for different choice of τ (results are not shown), although the estimated growth rates are generally larger than those of $\tau = 1$, the conclusions are similar.

The differences between these two approaches are shown not only in the growth rates but also in the singular vectors. Because the original data has been PCA filtered, the singular vectors so derived are linear combinations of the retained PC modes. Figure 7 shows the optimal singular vectors (the most unstable singular vectors) derived by the Yule-Walker equation as a function of the retained PC modes for $m = N = 2, 4, \ldots, 20$, respectively. As all the corresponding optimal growth rates are close to unity, the patterns between the predictor and the predictand are very similar for each given m. Furthermore, the optimal mode constituents increase as more PCs are retained. However, the relative contribution of each PC mode to the optimal mode does not show a dramatic change with the increase of the retained PCs. This is because the Yule-Walker equation is covariance based. The singular vectors so derived not only depend on the linear relation between the predictor and the predictand but also on their variance. Since most variance of the predictor and the predictand is explained by the first few PCs, the covariance structure between the predictor and the predictand is also primarily controlled by these few PCs. Therefore, even though the linear relationship may change if more PCs are retained, the optimal modes so derived are relatively insensitive to the variation of the retained PCs.

Figure 8 shows the same plots as Fig. 7, except for results from the SVD-CSD with $m = 80$. The resemblance between patterns of the predictor and the predictand can still be clearly observed. However, because they are derived using 80 retained PCs, the optimal mode constituents are no longer restricted only to PC modes less than the retained singular vectors. Furthermore, the relative contribution of each PC mode to the optimal mode gradually shifts toward higher PC modes, as more singular vectors are retained (*i.e.*, N increases). Because the SVD-CSD has linearized the relation between the predictor and the predictand before estimating the linear propagator, the optimal modes thus derived are chosen solely according to which combination of the empirical orthogonal functions [EOFs, i.e., column vectors of **Q** in Eq. (9)] of the joint predictor and the predictand matrices will yield the largest singular value. Therefore, the structures of the SVD-CSD derived optimal modes depend strongly on how many singular vectors (or equivalently, how many column vectors of **Q**) are retained.

6. SUMMARY AND DISCUSSION

The propagator of a linear model plays a very important

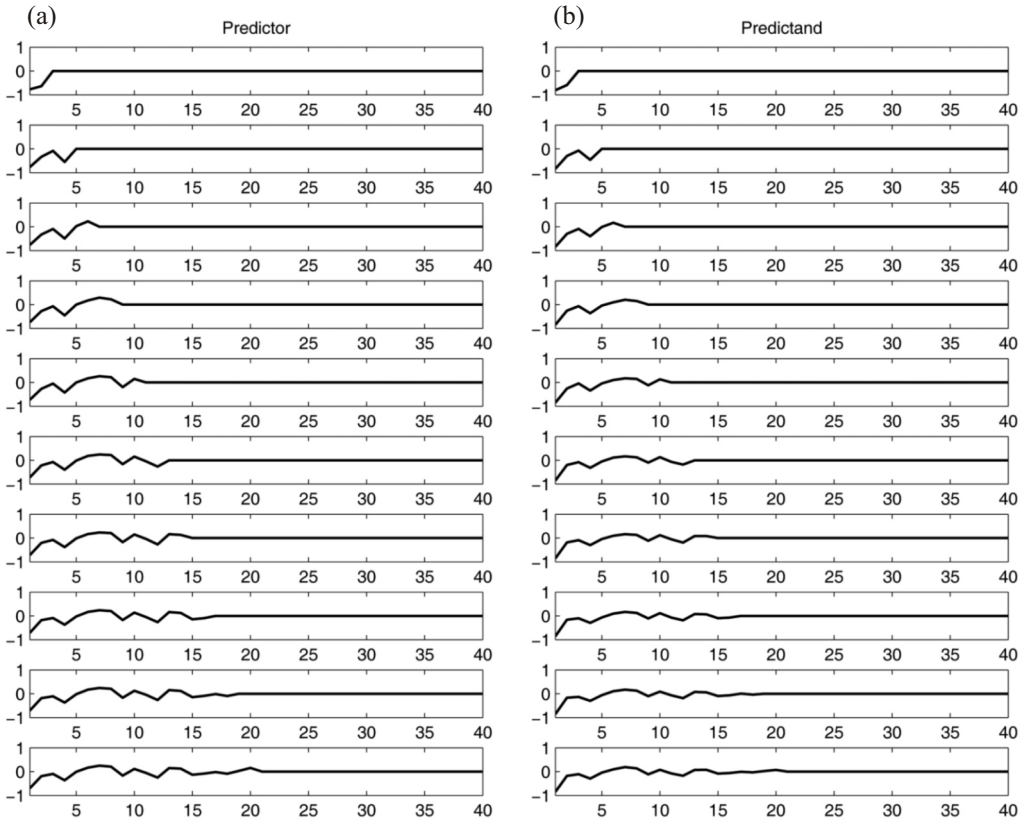

Fig. 7. The optimal singular vectors for the Kaplan's SSTA derived from the Yule-Walker equation as a function of PC modes, from top to bottom, for 2, 4,…, 20 and retained singular vectors, respectively. (a) and (b) show results for the predictor and the predictand, respectively.

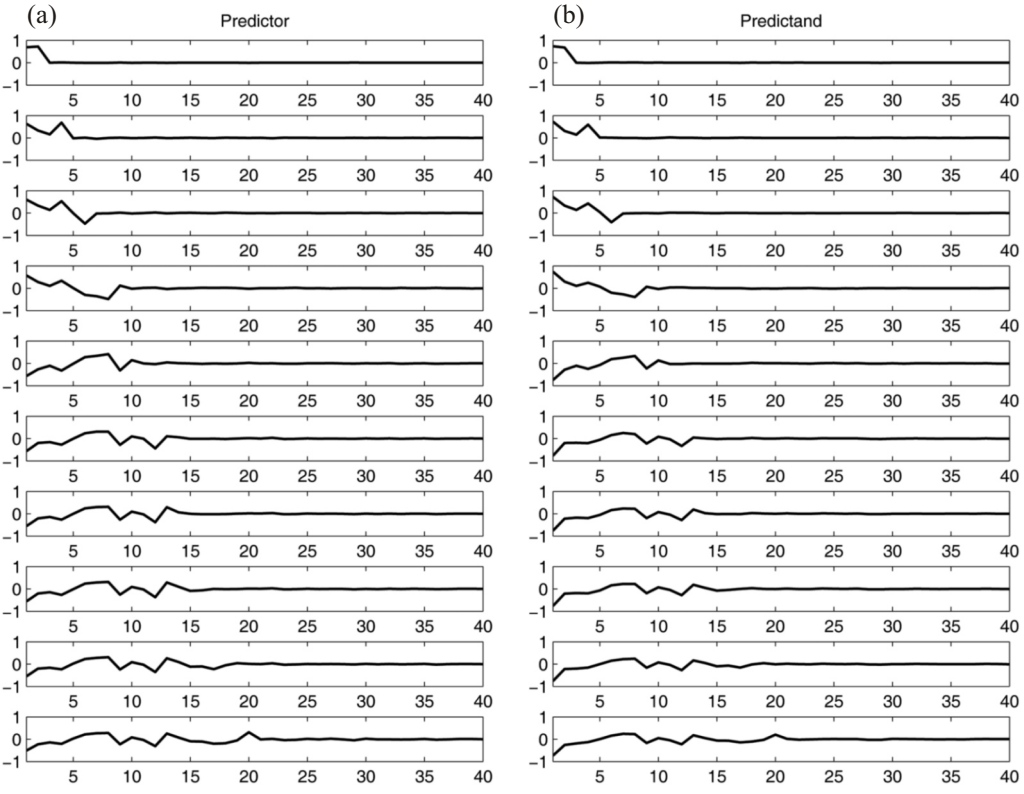

Fig. 8. The same as Fig. 5, except for results obtained by applying the SVD-CSD to the first 80 PCs of the Kaplan's SSTA.

role in normal mode and finite-time instability problems. Its estimation will affect whether the linear stability characteristics of the corresponding dynamic system can be properly extracted. Conventionally, the propagator is estimated using the Yule-Walker equation with the auto and lag covariance matrices of the predictor and the predictand. However, because nonlinear and noise information of the predictor and the predictand may also be included in forming these covariance matrices, the linear propagator thereby derived is likely to underestimate the linear relationship between them. Therefore, in this study the GSVD and SVD-CSD methods have been introduced as alternatives to the Yule-Walker equation to estimate the linear propagator and its associated properties for a linear model. In accord with the basic concept of a linear model, both methods linearize the relation between the predictor and the predictand by decomposing them to have a common evolution structure and then make the estimations. With these decompositions, the linear propagator and the associated singular vectors can be simultaneously derived. Furthermore, the connection between the finite-time amplitude growth rates and the singular values of the propagator are clearly established. Both the GSVD and the SVD-CSD, together with the Yule-Walker equation, are applied to two synthetic datasets and Kaplan's SSTA datasets to evaluate their respective performances. The results show that, because of the linearization and flexible filtering capabilities, the SVD-CSD allows the propagator, the finite-time growth rates, and the associated singular vectors to be more appropriately estimated.

It is noted that the application of the SVD-CSD is not restricted to linear models where the predictor and the predictand use the same field variables. Therefore, it can be applied not only to linear statistical prediction problem, but also to bias correction or statistical downscaling problems. As for data assimilation and related problems in numerical weather prediction, because the linear propagator is time dependent, the applicability of the SVD-CSD to these problems is not clear and needs further studies.

Acknowledgements I would like to thank the anonymous reviewers and the editor for their critical review of this paper. Their comments were essential to improving the final version of this paper. This study was supported by the National Science Council of Taiwan through grants NSC 90-2111-M-008-061 and NSC91-2111-M-008-011.

REFERENCES

Brockwell, P. J. and R. A. Davis, 1991: Time Series: Theory and Methods, Second edition, Springer-Verlag, 577 pp.

Farrell, B. F. and P. J. Ioannou, 1996a: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci*., **53**, 2025-2040, doi: 10.1175/1520-0469(1996)053<2025:GSTPIA>2.0.CO;2. [Link]

Farrell, B. F. and P. J. Ioannou, 1996b: Generalized stability theory. Part II: Nonautonomous operators. *J. Atmos. Sci*., **53**, 2041-2053, doi: 10.1175/1520-0469(1996)053<2041:GSTPIN>2.0.CO;2. [Link]

Golub, G. H. and C. F. Van Loan, 1996: Matrix Computations, Johns Hopkins Univ. Press, Baltimore, 694 pp.

Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856-1991. *J. Geophys. Res*., **103**, 18567-18589, doi: 10.1029/97JC01736. [Link]

Kim, H. M. and M. C. Morgan, 2002: Dependence of singular vector structure and evolution on the choice of norm. *J. Atmos. Sci*., **59**, 3099-3116, doi: 10.1175/1520-0469(2002)059<3099:DOSVSA>2.0.CO;2. [Link]

Palmer, T. N., R. Gelaro, J. Barkmeuer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci*., **55**, 633-653, doi: 10.1175/1520-0469(1998)055<0633:SVMAAO>2.0.CO;2. [Link]

Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan, 1994: Interdecadal changes of surface temperature since the late nineteenth century. *J. Geophys. Res*., **99**, 14373-14399, doi: 10.1029/94JD00548. [Link]

Penland, C. and T. Magorian, 1993: Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Climate*, **6**, 1067-1076, doi: 10.1175/1520-0442(1993)006<1067:PONSST>2.0.CO;2. [Link]

Penland, C. and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999-2024, doi: 10.1175/1520-0442(1995)008<1999:TOGOTS>2.0.CO;2. [Link]

Preisendorfer, R. W. and C. D. Mobley, 1988: Principal Component Analysis in Meteorology and Oceanography. Elsevier, 425 pp.

Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, **7**, 929-948, doi: 10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2. [Link]

von Storch, H. and F. W. Zwiers, 1998: Statistical analysis in climate research, Cambridge University Press, 484 pp.

von Storch, H., G. Bürger, R. Schnur, and J. S. von Storch, 1995: Principal oscillation patterns: A review. *J. Climate*, **8**, 377-400, doi: 10.1175/1520-0442(1995)008<0377:POPAR>2.0.CO;2. [Link]

Xue, Y., M. A. Cane, S. E. Zebiak, and M. B. Blumenthal, 1994: On the prediction of ENSO: A study with a low-order Markov model. *Tellus*, **46A**, 512-528, doi: 10.1034/j.1600-0870.1994.00013.x. [Link]

Xue, Y., A. Leetmaa, and M. Ji, 2000: ENSO prediction with Markov models: The impact of sea level. *J. Climate*, **13**, 849-871, doi: 10.1175/1520-0442(2000)013<0849:EPWMMT>2.0.CO;2. [Link]