APCC
APEC CLIMATE CENTER

RESEARCH REPORT

APCC
APEC CLIMATE CENTER

PREFACE

It is our pleasure to present this report on the APEC Climate Center (APCC)'s research activities in 2013, which has been a very productive year for our Center.

APCC has expanded its research scope, in response to regional societal and scientific needs. While building expertise in climate prediction remains a priority, we are extending our reach to include policy-relevant climate applications and value-added climate information products.

APCC has accelerated efforts to better our service to the region. As one of the main services provided by APCC, the MME 3-month prediction information has been productively applied by scientists in developing countries that are unable to produce their own prediction information. Furthermore, in order to better prepare for climate-related hazards in a timely manner, APCC launched its 6-month MME prediction service in September 2013. We also began to release forecasts of the Boreal Summer Intraseasonal Oscillation (BSISO), starting from July 2013, as the world's first operational BSISO forecast service. Our researchers also achieved great success in publishing their papers in noted academic journals. Dr. Ok-Yeon Kim, for example, published a paper in *Climate Dynamics* and her research was later selected as one of the Research Highlights by another distinguished journal, *Nature Climate Change*. The following research report provides more information about our research outcomes from 2013.

We will continue to promote the best use of our research outcomes in various scientific and application areas. Our successes and achievements would not have been possible without the support of our valued partners. In this regard, I extend my thanks to you and I hope you enjoy this 2013 Research Report.

Chin-Seung Chung
Director, APEC Climate Center

CONTENTS

Development of Scaled SVD Analysis and Related Methods with Focus on Application to Tropical-Extratropical Teleconnections

Dr. Erik Swenson | Climate Prediction Team

APCC
RESEARCH REPORT 2013-02

Development of Scaled SVD Analysis and Related Methods with Focus on Application to Tropical-Extratropical Teleconnections

Dr. Erik Swenson | Climate Prediction Team

ABSTRACT

Various multivariate statistical methods have been established and proven useful for representing linear relationships in datasets. Popular methods include Canonical Correlation Analysis (CCA), Maximum Covariance Analysis (MCA), and Redundancy Analysis (RDA) that are optimal at maximizing squared correlation, squared covariance, and variance explained, respectively. Considering that such measures of relation vary only in the representation of amplitude, they may be generalized into one particular generic form introduced here as scaled covariance. Scaled covariance is equivalent to covariance between scaled data, with the scaling transformation acting to apply a specified degree of normalization and decorrelation. For intermediate scaling values, scaling may be referred to as partial whitening, a transformation that may provide a new and highly useful noise pre-filter for covariance.

Extending to multiple dimensions, Scaled Singular Value Decomposition Analysis (SSVD) provides a continuum of methods linking CCA, MCA and RDA that maximizes sample squared scaled covariance. SSVD is derived and its properties are discussed. In order to understand the behavior of SSVD in terms of its ability to isolate coupled signals as a function of scaling values, SSVD is applied to randomly-generated synthetic datasets with known coupled signals embedded in background noise. From this, it is demonstrated that intermediate SSVD solutions tend to yield higher cross-validated measures of relation in contrast to their counterparts that are optimal in training data (CCA, MCA and RDA), consistent with higher skill in isolating the known signal. Given that MCA maximizes the *square* of covariance leading to an amplitude bias, correcting this bias with SSVD yields stronger relationships that better resemble the known signals. Also, values of scaling associated with highest skill are shown to be proportionate to the signal-to-noise ratio. It follows that provided with some crude prior expectation of the signal-to-noise ratio, one may appropriately choose scaling values. Interestingly, very small but nonzero scaling values (nearing those of CCA but not CCA) are also found to produce strong results.

Following the synthetic experiments, SSVD is applied to examine tropical-extratropical teleconnections, and SSVD solutions are shown to produce robust relationships for ENSO and ENSO Modoki teleconnections that are stronger compared with those estimated by traditional methods after cross-validation. It is found that SSVD solutions using very small scaling values yield the strongest relationships. Following this, an additional application of SSVD to statistical downscaling is investigated with the goal of enhancing skill of MME seasonal prediction of precipitation over the Maritime Continent. Different SSVD solutions are shown to produce a range of skill, with some solutions yielding the highest skill in contrast to other methods, providing an improvement in prediction skill for least the inter-annual component of precipitation variability. Further application to statistical downscaling is proposed for future study, for which SSVD and related pattern-based methods show some promise.

1. INTRODUCTION

Various multivariate statistical methods are useful for isolating physically meaningful linear relationships between multiple coupled variables. Two popular and well-understood symmetric methods are Canonical Correlation Analysis (CCA) and Maximum Covariance Analysis (MCA) that generate pairs of stationary spatial patterns whose temporal variation produces the highest squared correlation and squared covariance, respectively. In the climate system, such methods have been used to isolate observed inter-annual teleconnections associated with atmospheric interaction with persistent oceanic variation reflected in sea surface temperature (SST), particularly involving the tropical Pacific with the El Niño Southern Oscillation (ENSO) and flavors of ENSO (Kau & Yu 2009, Kug et al. 2009) or ENSO Modoki (Ashok et al. 2007). For ENSO, CCA has been used to demonstrate multiple impacts on North American surface temperature (DelSole & Shukla 2006), and MCA has been used to isolate teleconnections with the Northern Hemispheric (NH) circulation (Wallace et al. 1992, Wu 2010). Ultimately, application of these statistical tools not only aids physical understanding but is practically useful for predictive purposes. For instance, applying these methods for statistical downscaling can often improve prediction at regional scales. Other applications include diagnostic evaluation of dynamical models and paleoclimate reconstruction.

Apart from the symmetric relations of correlation and covariance, asymmetric methods may be preferable as they that focus on the variation of one variable with respect to the other. For instance, Redundancy Analysis (RDA, von Storch & Zwiers 1999) and Partial Least Squares (PLS) regression (Wold 1966) are different methods that both maximize the explained variance of one variable by another.

The numerical optimality of CCA, MCA and RDA stems from their common basis as diagonalizations of coefficient matrices generated from different solutions to multivariate regression (Tippett et al. 2008). As with any regression model, a large degree of statistical overfitting is expected when building a model from a training dataset that is under-sampled. This is often the case when estimating linear relationships of the climate system viewed through reanalysis products from a relatively short observational period. When optimally maximizing any measure of relation, even if

there is adequate convergence to a true relationship, the empirically estimated relationship is expectedly some mixture of the true relationship plus any and all meaningless noise that by chance increases the quantity that is maximized (even slightly). Therefore, sub-optimal techniques that filter out more undesired noise may be more useful when statistical overfitting is an issue. Sub-optimal techniques aimed at isolating relationships that have high correlation, covariance or explained variance must undoubtedly violate the strict constraints required by CCA, MCA and RDA that include optimality across the sample hand-in-hand with orthogonality requirements that strongly affect the representation of higher-order modes. Also, even if true co-existing relationships resemble optimal solutions and happen to satisfy orthogonality constraints, such relationships may not do so within small samples.

Considering the past development and practical use of traditional methods for isolating linear relationships, one basic idea is solidified that the property of individual variables that is worth considering is *amplitude*, a quantity that can play an important role in distinguishing a true relationship from meaningless noise. Amplitude may be represented by standard deviation, in terms of temporal variation either at individual spatial points or of stationary patterns as a whole, such as those represented by Empirical Orthogonal Functions (EOFs) computed from Principal Component Analysis (PCA). When relating two variables according to a particular measure of linear relation, the role of amplitude may be different depending on how standard deviation is treated in the measure. Correlation, covariance and explained variance (maximized by CCA, MCA and RDA, respectively) are measures that differ *only* in the representation of amplitude, and thus may be theoretically linked through measures with partial representation of amplitude.

Using synthetic data, the behavior and systematic bias are illustrated in detail for CCA and MCA in comparative studies by Bretherton et al. (1992) and Cherry (1996). MCA, also referred to as Singular Value Decomposition (SVD) analysis based on the numerical algorithm it uses, often generates solutions with high sub-sample stability given that MCA maximizes squared covariance which accounts for amplitude unlike correlation. Still in most typical cases, this comes at the expense of requiring representation of the linear relationship by a transformation that is strictly orthogonal

(Newman & Sardeshmukh 1995). This is not the case for traditional CCA, however traditional CCA applied to under-sampled data is generally highly sensitive to noise, and noise can be inflated substantially when computing inverses for CCA leading to solutions sometimes dominated by artificial overfitting. The only additional constraint that has been used to address this practical problem with traditional CCA is that proposed by Barnett & Preisendorfer (1987, hereafter BP) for which EOF truncation is applied as a pre-filter before CCA is computed. From the viewpoint of linear regression, EOF truncation provides a means of predictor selection and reduction, generally a critical step for obtaining useful results in under-sampled data. Although, results of CCA are typically highly sensitive to the choice of truncation number which itself is often chosen subjectively, and important relationship information is neglected entirely if it is represented in higher-order EOF modes. Is this a fair approach or simply a practical fix? Also, considering the high degree of spatial autocorrelation for typical geophysical variables, spatial scale is often coincident with variance such that dominant modes of variability span larger scales, giving EOFs a scale-selective tendency. In line with this idea, a second prior constraint involves spatial smoothing (e.g. ridge regression, Hastie et al. 1995).

In this study, Scaled Singular Value Decomposition Analysis (SSVD) is introduced as a new technique for linearly relating two variables by isolating sets of multivariate relationships. Complementary to correlation, amplitude is arguably the most important property of individual variables to consider when distinguishing meaningful relationships, and with this in mind SSVD is developed as a continuum of methods that vary only in terms specified role of amplitude. Univariate measures of relation given by correlation, covariance and explained variance differ only how amplitude is considered, and thus may be represented by a single measure referred to as scaled covariance, or covariance between scaled variables. As opposed to the whitening transformation, scaling, or partial whitening, is a transformation that results in only partial decorrelation of differing spatial points and partial normalization of individual spatial points. As a prior noise filter, this is suggested as a more objective approach in contrast to BP's EOF truncation applied with CCA. This approach can theoretically alleviate sensitivity to noise that pre-whitening can have with CCA. As SSVD maximizes scaled covariance, SSVD equivalently recovers CCA, MCA and RDA as special cases.

Scaled covariance is introduced in Section 2 as generic measure of relation bridging traditional measures. Section 3 presents the formulation and properties of SSVD. Section 4 then applies SSVD towards isolating known coupled signals embedded in synthetic datasets, contrasting results with CCA, MCA and RDA. In Section 5, SSVD is applied to isolate the leading two wintertime relationships observed between tropical Pacific precipitation and NH circulation, namely the ENSO and ENSO Modoki teleconnection signals. Section 6 briefly demonstrates an application of SSVD and related methods to statistical downscaling, and Section 7 provides a summary with a final discussion.

2. SCALED COVARIANCE

In this section, scaled covariance is introduced as a generic measure of linear relation whose form takes into consideration a range of amplitude weighting for individual variables. First it is derived and related to common measures of relation in its univariate form, and then it is derived in its multivariate form. As the name alludes, scaled covariance may be simply understood as covariance between two scaled variables. The prior scaling, or partial whitening transformation, is introduced here and related to the whitening transformation. Scaled covariance recovers correlation, covariance, and explained variance for which it provides a continuous link.

Standard measures of linear relation include correlation, covariance and explained variance, measures that act to relate two variables *x* and *y* in different manners. The notation used here is identical to Tippett et al. 2008, (hereafter T2008, see equations 1-5) who also revisit univariate measures or relation as a basis for multivariate regression. The most common measure is correlation between *x* and *y* ($\rho_{x,y}$) which is given by $\rho_{x,y}=\sigma_x^{-1} < xy > \sigma_y^{-1}$, with $<>$ denoting the expectation and standard deviations given by $\sigma_x = \sqrt{< x^2 >}$ and $\sigma_y = \sqrt{< y^2 >}$. Retaining standard deviation (σ_x and σ_y), the covariance between *x* and *y* is given by $< xy > = \sigma_x \rho_{x,y} \sigma_y$, and the square root of variance of *y* explained by *x* is given

by $\sigma_x^{-1}|< xy >|=|\rho_{x,y}|\sigma_y$. It is important to note that such measures all involve correlation and differ only in how amplitude, or standard deviation, is considered. One form that encompasses these measures of relation may be constructed by simply assigning exponents to individual standard deviations of *x* and *y*, given by parameters α and β, respectively. This results from computing the covariance between transformed variables *x** and *y** as follows:

$$\langle x^* y^* \rangle = {\sigma_x}^{\alpha} \rho_{x,y} {\sigma_y}^{\beta} \tag{2.1}$$

$$\text{with} \quad \begin{aligned} x^* &= {\sigma_x}^{\alpha-1} x \\ y^* &= {\sigma_y}^{\beta-1} y \end{aligned}$$

As a generalized measure of linear relation, Eq. 2.1 is referred to as *scaled* covariance. The transformations that produce *x** and *y** act to reassign amplitude of *x* and *y* based on powers of σ_x and σ_y specified by α and β, respectively, an operation that may be simply referred to as *scaling*. Assuming stationarity, scaling is linear for fixed values of α and β. Also, note that normalized *x** and *x* are equivalent with $x^*/\sqrt{< x^{*2} >} = \sigma_x^{-1} x$, and similarly for *y** and *y*. Correlation results from choosing α=β=0, covariance results from choosing α=β=1, and the square root of explained variance results from choosing α=0 and β=1 followed by taking the absolute value. Similar as T2008, the formation of scaled covariance as regression coefficient between transformed variables is demonstrated in Appendix B.

The sample estimate of scaled covariance may be generalized to multiple dimensions, but first it is necessary to examine covariance itself. Given two variables *X* and *Y*, now matrices of anomalies with dimensions of space × time, assume fields have uniform spatial weighting and have absorbed a factor of $1/\sqrt{n}$ in order to simplify representations of covariances and cross-covariance. Then, individual covariances for *X* and *Y* are given by XX^T and YY^T, respectively, and cross-covariance is given by XY^T. For *X*, values of diagonal (off-diagonal) elements of the positive-definite matrix XX^T give the variances at individual spatial points (covariances between differing spatial points). The corresponding sample standard deviations result from applying a matrix square root with $(XX^T)^{1/2}$, defined by $XX^T=(XX^T)^{1/2}(XX^T)^{1/2}$. It follows that an important transformation for *X* results through multiplication by the inverse

of its matrix square root $(XX^T)^{-1/2}$, a multivariate form of normalization called the *whitening transformation* (Delsole & Tippett 2007). The whitening transformation both standardizes the data (normalization at individual spatial points) *and* removes covariance between differing spatial points, given by multiplication with the diagonal and off-diagonal elements of $(XX^T)^{-1/2}$, respectively. Considering the whitened variable $\tilde{X}=(XX^T)^{-1/2}X$, these points are evident in the fact that $\tilde{X}\tilde{X}^T=(XX^T)^{-1/2}XX^T(XX^T)^{-1/2}=I$. Note that only standardization results when multiplying by $(Diag\ XX^T)^{-1/2}$, a diagonal matrix with values in diagonal elements equivalent to those of the inverse matrix square root.

Analogous to univariate measures of relation, multivariate forms of sample correlation, covariance and explained variance may be produced by applying the whitening transformation to both, none or one of the variables, respectively. For instance, cross-correlation may be given by $(XX^T)^{-1/2}XY^T(YY^T)^{-1/2}$. Seeking a single form that encompasses these measures of relation may be done by first introducing a new more generalized transformation that continuously bridges the whitening transformation with the identity matrix (no transformation). Such a transformation may be represented for X and Y by T_x and T_y as functions of parameters α and β, respectively, as follows:

$$T_x = (XX^T)^{(\alpha-1)/2} \tag{2.2}$$

$$T_y = (YY^T)^{(\beta-1)/2} \tag{2.3}$$

Applying transformations of Eqs. 2.2 & 2.3 produces transformed variables $X^*=T_xX$ and $Y^*=T_yY$, respectively. As a general form, both operations may be referred to as the *scaling* transformation (in multiple dimensions). More specifically, for fractional values of α and β this may be called the *partial whitening* transformation, a term borrowed from an analogous procedure involving spherical harmonics as computed by Donohue et al. (2007). For X, as α increases from zero to one, T_x continuously transitions from $(XX^T)^{-1/2}$ (the whitening transformation) to the identity matrix, respectively, given that $(XX^T)^0=I$. Partial whitening acts to partially normalize X at individual spatial points (degree of amplitude reduction) and only partially decorrelate differing spatial points.

In theory, partially whitening can alleviate issues involved with whitening, given that the necessary computation of the inverse is often highly sensitive to noise in *X*, especially if *X* is under-sampled. The full inverse may have artificially very high covariance values (particularly if $(XX^T)^{-1/2}$ near singular), however such a fractional inverse tends have *smaller* covariance values. This approach may provide an effective noise filter prior to relating two variables. Given its exponential form, scaling may provide a more objective means of damping unrelated variability disproportionately while retaining important covariability, compared with approaches like EOF truncation that completely neglect portions of variability.

After scaling *X* and *Y* using particular values of α and β, respectively, the cross-covariance of the scaled data X^*Y^{*T} gives the multivariate form of the sample scaled covariance (2.1) and is given as follows:

$$X^*Y^{*T} = T_x X Y^T T_y{}^T = (XX^T)^{(\alpha-1)/2} XY^T (YY^T)^{(\beta-1)/2} \qquad (2.4)$$

Again, by applying this transformation, multivariate forms of sample correlation, covariance and explained variance may be constructed when choosing α=β=0; α=β=1; and α=0, β=1, respectively. The strong benefit of scaled covariance here is that it continuously links standard measures of relation in a relatively straightforward manner, through partial whitening.

In PC-space, scaling preserves EOFs while acting to assign exponents to PC singular values (see Appendix A). For *X*, as α increases from zero to one, the weighting of the PCs changes from being all equal to one (the whitening transformation) to being weighted by the singular values (unchanged). From this it is clear that varying the degree of partial whitening acts to vary the disparity in the weighting of PCs, proportional to α and β for *X* and *Y*, respectively.

3. SCALED SVD ANALYSIS (SSVD)

In this section, Scaled Singular Value Decomposition Analysis (hereafter SSVD) is introduced as a new technique for linearly relating two variables by isolating sets of multivariate relationships. SSVD is developed as a continuum of methods that generates fixed patterns whose temporal variation maximizes the sample squared scaled covariance (2.1). First, CCA, MCA and RDA are derived in a single form as a constrained optimization problem. Second, the problem involved with SSVD is constructed as a particular generalized form. Third, SSVD solutions are formed and its basic properties are discussed. Lastly, some variants of SSVD are briefly discussed. Note that given computational constraints, SSVD and related methods are ideally computed in PC space, and SSVD is derived fully in PC space in Appendix A.

3.1 SSVD as a constrained optimization problem

The traditional methods of Canonical Correlation Analysis (CCA), Maximum Covariance Analysis (MCA) and Redundancy Analysis (RDA) may all be derived as optimization problems involving differing constraints whose solutions may be found using singular value decomposition (SVD). Maximization involving the basic form of linear relation (in terms of covariance XY^T) seeks pattern projection vectors q_x and q_y for the following optimization problem:

$$max_{q_x,q_y} \left({q_x}^T X Y^T q_y \right) \tag{3.1}$$

For CCA-1, MCA-1 and RDA-1 (leading modes of CCA, MCA and RDA), this optimization is subject to different constraints as follows:

$${q_x}^T X X^T q_x = 1 \text{ and } {q_y}^T Y Y^T q_y = 1 \text{ for CCA-1} \tag{3.2}$$

$${q_x}^T q_x = 1 \text{ and } {q_y}^T q_y = 1 \text{ for MCA-1} \tag{3.3}$$

$${q_x}^T X X^T q_x = 1 \text{ and } {q_y}^T q_y = 1 \text{ for RDA -1} \tag{3.4}$$

Note that only for MCA-1, the pattern projection vectors q_x and q_y must have unit variance (3.3), whereas the temporally-varying projections themselves ($X^T q_x$

and $Y^T q_y$) must have unit variance for CCA (3.2). These different constrained optimizations of CCA and MCA, as well as further properties and solutions, are examined in detail by Bretherton et al. (1992). Given that CCA and RDA are equivalent to MCA involving pre-whitened variables (T2008), the constrained optimization problems may be generalized to take on the form of MCA-1 as follows:

$$max_{\tilde{q}_x,\tilde{q}_y} \left(\tilde{q}_x{}^T \tilde{X}\tilde{Y}^T \tilde{q}_y\right) \tag{3.5}$$

$$\text{subject to } \tilde{q}_x{}^T \tilde{q}_x = 1 \text{ and } \tilde{q}_y{}^T \tilde{q}_y = 1 \tag{3.6}$$

For CCA-1, $\tilde{X}$ and $\tilde{Y}$ in Eq. 3.5 are whitened versions of X and Y, respectively, balanced by the modified solutions given by $\tilde{q}_x = (XX^T)^{1/2} q_x$ and $\tilde{q}_y = (YY^T)^{1/2} q_y$, through which substitution recovers (3.1) and (3.2). For RDA-1, whitening is only applied to X and not Y, whereas MCA-1 involves no transformations with Eq. 3.5 equivalent to Eq. 3.1 and Eq. 3.6 equivalent to Eq. 3.3.

With the scaling transformation (T_x and T_y of Eqs. 2.2 & 2.3, respectively) providing a particular generalization that encompasses transformations for CCA-1, MCA-1 and RDA-1, transformations are represented in a single form that yields the constrained optimization problem for Scaled SVD Analysis (SSVD) as follows:

$$\tilde{q}_x = T_x{}^{-1} q_x \quad \text{and } \tilde{q}_y = T_y{}^{-1} q_y \tag{3.7}$$

$$\tilde{X} = X^* = T_x X \quad \text{and } \tilde{Y} = Y^* = T_y Y \tag{3.8}$$

Note that inverse-scaling is given by $T_x^{-1}=(XX^T)^{(1-\alpha)/2}$ and $T_y^{-1}=(YY^T)^{(1-\beta)/2}$. Re-writing the SSVD problem with substitution into Eqs. 3.5 & 3.6 and replacing $\tilde{X}$ and $\tilde{Y}$ by those associated with scaling (2.2 & 2.3), the problem may be expressed in terms of q_x and q_y as a unified, regularized approach to the constrained optimization problem (as to compare with Equations 3.1-3.4) as follows:

$$max_{q_x,q_y} \left(q_x{}^T XY^T q_y\right) \tag{3.9}$$

$$\text{subject to } q_x{}^T (XX^T)^{1-\alpha} q_x = 1 \text{ and } q_y{}^T (YY^T)^{1-\beta} q_y = 1 \tag{3.10}$$

From Eqs. 3.9 & 3.10, CCA-1 constraints (3.2) are recovered for $\alpha=\beta=0$, MCA-1 constraints (3.3) are recovered for $\alpha=\beta=1$, and RDA-1 constraints (3.4) are recovered for $\alpha=0$, $\beta=1$.

3.2 SSVD maximization and basis sets

After substituting Eq. 3.8 into Eq. 3.5 and before computing SVD to obtain SSVD-1, it is clear that the SSVD-1 optimization problem yields $\max_{\tilde{q}_x, \tilde{q}_y}(\tilde{q}_x^T X^* Y^{*T} \tilde{q}_y)$ which maximizes the sample squared scaled covariance given in its univariate form by $(\hat{\sigma}_x{}^{\alpha}\hat{\rho}_{x,y}\hat{\sigma}_y{}^{\beta})^2$. This general form allows for the role of amplitude (treatment of $\hat{\sigma}_x$ and $\hat{\sigma}_y$) to be specified, and as shown previously, scaled covariance equates to correlation, covariance and explained variance for particular values of α and β. Equivalently SSVD recovers CCA, MCA and RDA, and the primary range of SSVD solutions may be visualized on the α-β plane in Fig. 3.1. The strong benefit of SSVD is that it provides a unified framework that continuously links traditional methods through a continuum of intermediate solutions arising from choosing $0 < \alpha, \beta < 1$.

The SSVD-1 solutions are computed by taking SVD of X^*Y^{*T} yielding orthonormal singular vectors (columns of U and V) and positive, sorted singular values (diagonal elements of S) as follows:

$$X^*Y^{*T} = T_x X Y^T T_y{}^T = USV^T \tag{3.11}$$

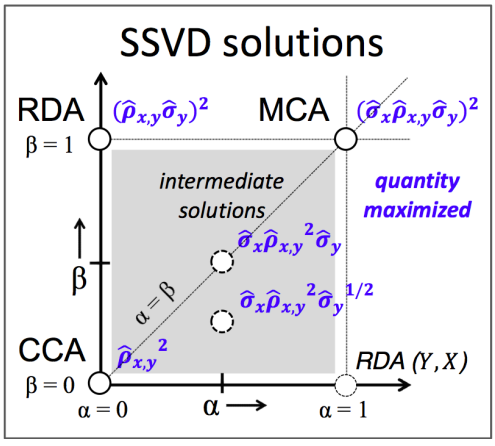

Figure 3.1 Diagram of Scaled SVD solutions as a function of α (x-axis) and β (y-axis). The quantity maximized for particular solutions is labeled in blue, and intermediate solutions ($0 < \alpha < 1$ and $0 < \beta < 1$) are represented by the gray-shaded region. Solutions for Canonical Correlation Analysis (CCA, α=β=0), Maximum Covariance Analysis (MCA, α=β=1) and Redundancy Analysis (RDA, α=0, β=1) are indicated.

From Eq. 3.11, it is clear that solutions $\tilde{q}_x$ and $\tilde{q}_y$ are given by the leading singular vectors with $\tilde{q}_x$=U_1 and $\tilde{q}_y$=V_1, which have unit variance satisfying Eq. 3.6 and yield the highest possible scaled covariance, or covariance between scaled variables X^* and Y^*, given by the leading singular value with $S_{1,1} = U_1^T X^* Y^{*T} V_1$. Due to the fact that singular vectors are orthonormal (U and V are unitary), it can be easily shown that the total squared scaled covariance is given by $\| X^* Y^{*T} \|_F^2 = \| S \|_F^2$. Apart from SSVD-1, the remaining sample squared scaled covariance, given by $\| S \|_F^2 - S_{1,1}{}^2$, is maximized by the second leading singular vectors U_2 and V_2 subject to the orthogonality constraint that $U_1^T U_2$=0 and $V_1^T V_2$=0. Further, the remaining $\| S \|_F^2 - S_{1,1}{}^2 - S_{2,2}{}^2$ is maximized by the third leading singular vectors and so on.

Singular vectors U and V provide the SSVD pattern projection vectors for the scaled variables that yield the SSVD temporal coefficients, or variates (R_x and R_y in Eqs. 3.13 & 3.15, respectively) with $R_x^T R_y$=S. For the associated SSVD patterns, U and V provide the associated pattern vectors in 'scaled space' that maximize the squared covariance between X^* and Y^* with X^*=UR_x^T and Y^*=VR_y^T, respectively. Thus, in line with the transformed solutions of 11, the SSVD patterns (P_x and P_y) are recovered by inverse-scaling U and V (3.12 & 3.14). Prior to typical post-normalization (shown later), the SSVD basis sets are given as follows:

$$P_x = T_x{}^{-1} U = (XX^T)^{(1-\alpha)/2} U \tag{3.12}$$

$$R_x = X^{*T} U = X^T T_x U = X^T (XX^T)^{(\alpha-1)/2} U \tag{3.13}$$

$$P_y = T_y{}^{-1} V = (YY^T)^{(1-\beta)/2} V \tag{3.14}$$

$$R_y = Y^{*T} V = Y^T T_y V = Y^T (YY^T)^{(\beta-1)/2} V \tag{3.15}$$

Note that the SSVD basis sets represent a complete decomposition with $P_x R_x^T$=$T_x^{-1} UU^T T_x X$=X, and similarly for Y. Scaled covariance may be represented in terms of the variates alone with $R_x^T R_y$=$U^T X^* Y^{*T} V$=S, from which it may be deduced that differing SSVD variates have zero cross-correlation as differing modes have zero scaled cross-covariance (S is diagonal).

In terms of varying α and β, symmetry is preserved when choosing α=β, whereas

$\alpha \neq \beta$ yields asymmetric relations in which the role of amplitude is treated differently for *X* than it is for *Y*. Again, choosing α=β=0 results in CCA patterns and variates for, while on the other hand choosing α=β=1 results in MCA patterns and variates. Also, choosing α=0 and β=1 results in RDA patterns and variates.

In general, traditional methods are unique cases of SSVD that satisfy orthogonality conditions unlike intermediate solutions (0<α<1 or 0<β<1). Provided *X* has nonzero spatial autocorrelation (XX^T is not diagonal), this is shown in that orthonormal pattern vectors are preserved (3.16) only when no transformation is made (α=1), and variates are restricted to be uncorrelated (3.17) only when whitening is performed to remove spatial autocorrelation (α=0). This is demonstrated for *X* as follows:

$$P_x{}^T P_x = U^T (XX^T)^{1-\alpha} U = I \quad iff \quad \alpha = 1 \tag{3.16}$$

$$R_x{}^T R_x = U^T (XX^T)^{\alpha} U = I \quad iff \quad \alpha = 0 \tag{3.17}$$

For *X*, variance may be cleanly partitioned when variates are uncorrelated with unit variance (all variance contained in patterns) such that the total variance $\| X \|_F^2$ may be represented as a sum of that from individual SSVD patterns with $\| X \|_F^2$ =*tr*[$(P_x{}^T P_x)(R_x{}^T R_x)$]= $\| P_x \|_F^2$, and similarly for *Y*. Otherwise for $\alpha \neq 0$, post-normalization of the weighted SSVD variates is desired so that patterns absorb full amplitude with physical units. For *X*, this is done using the normalization matrix S_x^R=(*Diag* $R_x{}^T R_x$)$^{1/2}$, with weighted SSVD patterns and normalized variates then given by $P_x S_x^R$ and $R_x S_x^{R^{-1}}$, respectively, and similarly for *Y*. This allows SSVD cross-correlation matrix ($\rho^{X,Y}$) to be represented by $\rho^{X,Y}$=$S_x^{R^{-1}}(R_x{}^T R_y) S_y^{R^{-1}}$=$S_x^{R^{-1}} S S_y^{R^{-1}}$, a diagonal matrix, and correlation between modes of *X* itself ($\rho^{X,X}$) to be represented by $\rho^{X,X}$ =$S_x^{R^{-1}}(R_x{}^T R_x) S_x^{R^{-1}}$, and similarly for $\rho^{Y,Y}$. Considering differing SSVD modes *j* and *k* (*j*≠*k*), constraints on cross-correlation ($\rho_{j,k}^{X,Y}$ and $\rho_{k,k}^{X,Y}$) act to limit the correlation values between differing modes of *X* itself ($\rho_{j,k}^{X,X}$, an off-diagonal element of $\rho^{X,X}$) such that $1+2\rho_{j,k}^{X,X}\rho_{j,k}^{X,Y}\rho_{k,k}^{X,Y} \geq \rho_{j,k}^{X,X^2}+\rho_{j,k}^{X,Y^2}+\rho_{k,k}^{X,Y^2}$ is satisfied. Given the requirement of zero cross-correlation between differing modes ($\rho_{j,k}^{X,Y}$=0), the inequality reduces to $\rho_{j,k}^{X,X^2} \leq (1-\rho_{k,k}^{X,Y^2})$, and assuming a high but typical value of $\rho_{k,k}^{X,Y}$=0.95, correlations between differing modes are restricted to be relatively small with $|\rho_{j,k}^{X,X}| \leq 0.3$. This means that for very high SSVD cross-correlation values, variates are generally *nearly* uncorrelated.

3.3 A CCA-like variant of SSVD

Recalling the SSVD singular vectors *U* and *V*, these pattern projection vectors may be used to generate alternative basis sets that are strictly temporally uncorrelated, as is the case for CCA. Such basis sets result from applying the corresponding *whitened* data to the projection vectors rather than the scaled data (as is for SSVD). For *X*, such basis sets are given by $P'_x=(XX^T)^{1/2}U$ and $R'_x=X^T(XX^T)^{-1/2}U$, respectively, and similarly for *Y* using *V*. The resultant modified patterns are related to the SSVD patterns through transformations from whitened space to scaled space, given for *X* by $P'_x=(XX^T)^{1/2}(XX^T)^{\alpha/2}P_x$, essentially an exponential inflation that disproportionately redistributes variance to spatial points that have higher variance already, given for *X* as follows:

$$P'_x = (XX^T)^{1/2}T_xP_x = (XX^T)^{\alpha/2}P_x \qquad (3.18)$$

Proportional to the nonzero magnitude of α, this acts to redirect pattern vectors towards the leading EOFs (see Appendix A), and provided a robust tendency of EOF patterns, this may be beneficial for increasing sub-sample pattern stability. As the maximization of scaled covariance is *not* invariant to such a transformation (provided $\alpha \neq 0$ or $\beta \neq 0$), the leading mode of this variant of SSVD necessarily yields a lower value of scaled covariance than that for SSVD-1 and does not seem to optimize any straightforward quantity. Despite this, the clear benefit of this variant to SSVD is that variates are always uncorrelated with $R'^T_x R'_x=I$ and $R'^T_y R'_y=I$. Also, this approach further alleviates other orthogonality constraints of MCA emphasized by Newman & Sardeshmukh (1995). This analysis was developed prior to SSVD and has been referred to as Covariance-weighted CCA (CW-CCA) or Pattern-Emphasizing MCA (PMCA) when no scaling is applied ($\alpha=\beta=1$). In practice, typical solutions of this variant to SSVD resemble SSVD solutions that use fairly small scaling values. This has not been rigorously tested and well-understood but may be related to the fact that such SSVD solutions produce variates that are nearly uncorrelated (similar to CCA), yet still retain some amplitude weighting (unlike CCA).

4. SYNTHETIC EXPERIMENTS

In this section, the behavior of SSVD is examined in an idealized context that is fully understood, specifically in terms of its expected performance in isolating an idealized coupled signal as a function of scaling values α and β. The primary motivation of examining SSVD with synthetic data is to provide guidance and understanding of its use that can be generalized to more complex linear relationships. The results of the leading SSVD mode (SSVD-1) are examined in terms of standard measures of relation alongside its skill in capturing a known coupled signal embedded in pairs of randomly-generated under-sampled synthetic datasets. Results are generated as a function of α and β so that a range of SSVD solutions may be contrasted with those of CCA, MCA and RDA. As is demonstrated, two primary findings arise from synthetic experiments. First, intermediate solutions of SSVD tend to yield the highest skill in isolating the coupled signal. Second, scaling values associated with the highest skill in isolating the coupled signal are proportional to the signal-to-noise ratio. Based on the synthetic experiments, two particular solutions of SSVD are proposed for common use. If the expected signal-to-noise ratio is similar for *X* and *Y*, SSVD using α=β=1/2 is recommended. On the other hand, if the expected signal-to-noise ratio for *X* is appreciably higher (lower) than that for *Y*, SSVD using α=1/2, β=1/4 (α=1/4, β=1/2) is recommended. A third alternative solution is also suggested that maintains the same symmetry (α/β=1) or asymmetry (α/β=1/2 or β/α=1/2), but uses very small (but nonzero) scaling values. In some situations, SSVD using EOF truncation may also be beneficial. Mathematical details for the generation of synthetic datasets, including the process of generating random data with an expected cross-correlation, are provided in Appendix C.

Synthetic datasets *X* and *Y* of equal spatial dimensions, specified as *m*=60 for all experiments, and an under-sampled number of time samples with *n*=30, consist of a common orthogonal set of population EOF patterns represented by spherical harmonics. Specifically, population EOF-1 and EOF-2 (EOF-3 and EOF-4) patterns are sine and cosine waves, respectively, with one half (three halves) of a period spanning the domain. By design, variance is evenly distributed across space. The associated variates consist of random white noise for which only one true coupled

signal is present with a population correlation of 0.9. An additional constraint for the signal is the fractional amplitude weighting of population EOF-1 which is specified differently for different experiments as a and b for X and Y, respectively. With a total expected variance of 1, population EOF-1 has a variance of 0.25, equivalent to the sum of variances of EOF-2, EOF-3 and EOF-4. In terms of a and b, signal variances are then given by $0.25(2a+1)/3$ and $0.25(2b+1)/3$, respectively, and range from $0.25/3 \approx 0.083$ to 0.25 (when changing a or b from 0 to 1) corresponding to a signal-to-noise ratio of roughly 0.3 to 0.6, respectively. Individual variances of the population EOFs decay exponentially with values that are not inconsistent with typical distributions for inter-annual variability of large-scale geophysical variables. These population EOFs are embedded in background noise that is random red noise in space, generated in a similar manner as that used by Bretherton et al. (1992) with a spatial autocorrelation of 0.9 (for neighboring spatial points). This background noise varies in time as uncorrelated white noise and is specified to have a variance of 0.5. Given that the red noise inevitably projects randomly onto the population EOFs, overfitting expectedly leads to sample EOFs that explain *more* variance than the population EOFs.

For different amplitude signals specified by a and b, SSVD is computed with 100 synthetic realizations across a range of exponential scaling factors α and β varying evenly from $0 \leq \alpha \leq 1.5$ and $0 \leq \beta \leq 1.5$ in increments of 0.05 spanning the solutions of CCA, MCA and RDA. By design, these traditional methods maximize correlation (*COR*), squared covariance fraction (*SCF*), and the fraction of variance of Y explained by X (*FVE*), respectively, and these standard measures of relation are computed in order to assess the strength of the dominant relationship represented by the leading SSVD mode (SSVD-1). With SSVD-1 given by $X_1=P_{x,1}R_{x,1}^{\mathrm{T}}$ and $Y_1=P_{y,1}R_{y,1}^{\mathrm{T}}$, *SCF* and *FVE* are computed using a weighted mean-squared error approach with $\| A \|_F$ giving the Frobenius matrix norm of A as follows:

$$SCF = 1 - \|(X - X_1)(Y - Y_1)^T\|_F^2 / \|XY^T\|_F^2 \quad (4.1)$$

$$FVE = 1 - \left\|(XX^T)^{-1/2}(X - X_1)(Y - Y_1)^T\right\|_F^2 / \left\|(XX^T)^{-1/2}XY^T\right\|_F^2 \quad (4.2)$$

Note that the square of the Frobenius matrix norm of A is equivalent to the

sum of the diagonal elements (or trace) of AA^T, given by $\| A \|_F^2 = tr(AA^T)$. As the true signal is known, in addition to the measures of relation, primary focus is given to the skill of isolating the signal using SSVD-1. This is also computed in terms of a weighted mean-squared error, and is represented by the fraction of signal variance explained for individual variables X and Y ($FSVE_x$ and $FSVE_y$, respectively), given as follows:

$$FSVE_x = 1 - \left\| X^{signal} - X_1 \right\|_F^2 / \left\| X^{signal} \right\|_F^2 \quad (4.3)$$

$$FSVE_y = 1 - \left\| Y^{signal} - Y_1 \right\|_F^2 / \left\| Y^{signal} \right\|_F^2 \quad (4.4)$$

The skill of isolating the coupled signal as a whole is given simply by averaging the skill with the individual variables, with $FSVE=(FSVE_x+FSVE_y)/2$. Finally, overfitting is accounted for by computing leave-one-out cross-validation for *FSVE* and all measures of relation.

By design, the traditional methods produce modes that maximize standard measures of relation in data used for the analysis (training data), and a null hypothesis is that the methods also yield the highest values for the corresponding measures of relation in independent data, estimated with cross-validation. On the other hand, the practical goal of isolating the coupled signal depends on how characteristic the measures of relation are for the coupled signal, and mainly the characteristics of unrelated noise. In the absence of background noise, CCA-1 uniquely recovers the signal exactly. To some degree, this may be considered a bias in some circumstances such that CCA may be expectedly more suitable at recovering the signal in the presence of background noise (as it is without noise). In light of this, the signal is also alternatively considered as MCA-1 of the data in the absence of background noise.

Note that for under-sampled data ($n<m$), without BP's EOF pre-filtering, CCA expectedly yields an artificially high correlation due to overfitting, and such results are generally incomparable to MCA and that from other methods. Also due to overfitting, RDA converges to EOF-1 for Y regardless of any information contained in X, as EOF-1 explains the most variance. For these reasons, SSVD methods are also contrasted using prior EOF truncation.

By definition, CCA, MCA and RDA always maximize *COR*, *SCF* and *FVE*, respectively, in the training data. However, the coupled signal is best isolated with SSVD-1 using intermediate scaling values. This is verified in Fig. 4.1 which shows the α-β coordinates for the maximum average *FSVE*, *COR*, *SCF* and *FVE* alongside the average *FSVE* (contours), computed for a = b = 1, 0.75, 0.5. For smaller amplitude signals (a = b = 0.75, 0.5), intermediate solutions of SSVD-1 using α=β=0.25 are most skillful and stand out over MCA-1.

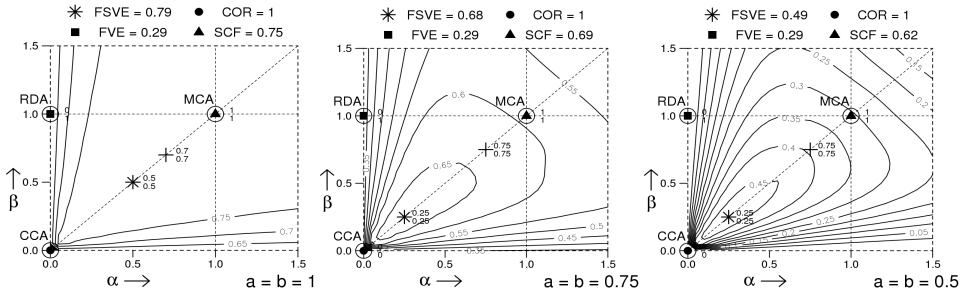

Figure 4.1 Fraction of signal variance explained (*FSVE*, contours at intervals of 0.05) for SSVD-1 plotted as a function of scaling parameters α and β averaged across 100 synthetic realizations for which the fraction of EOF-1 for the signal in *X* and *Y* (*a* and *b*, respectively) is specified as 1 (left), 0.75 (middle) and 0.5 (right). Plotted are the α-β coordinates for maximum *FSVE*, *FSVE* of the alternate signal (MCA-1 of data without noise), correlation (*COR*), squared covariance fraction (*SCF*), and variance of Y explained by X (*FVE*) as indicated by an asterisk, plus, solid dot, solid triangle and solid square, respectively. The corresponding values are given above the plot.

This contrast is enhanced as amplitude of the signal is decreased, despite all SSVD solutions generally showing less skill. If the signal is represented by MCA-1 in the absence of noise, MCA-1 (α=β=1) is still sub-optimal in isolating it, and SSVD-1 with α=β=0.75 generally yields the highest *FSVE*. When choosing very small α or very small β, SSVD-1 starts to diverge from the coupled signal, especially as the signal weakens, a prior expectation of CCA and RDA without the use of EOF pre-filtering. Skill for CCA tends to be incomparable to other solutions. However, small symmetric scaling values fairly close to zero (CCA) still interestingly give high skill, and generally produce the highest median skill (not shown). After cross-validation, results for average *FSVE* are generally preserved with an even decrease for all values of α and β (Fig. 4.2). *COR* and *SCF* however are consistently higher for intermediate solutions of SSVD-1, and they have better correspondence to *FSVE* after cross-validation. As a and b are decreased from 1 to 0.5, the location

of maximum *COR* shifts from α=β=0.8 to α=β=0.2. Qualitatively, as the amplitude of the signal decreases, scaling down the data helps produce modes that yield higher correlation and better resemble the actual signal. This makes intuitive sense that when the amplitude of the signal is smaller relative to the amplitude of the background noise, amplitude is consequently a less efficient filter for distinguishing the signal from noise, and loosening its role as a constraint produces better results. On the other hand, *SCF* has much less sensitivity and values of α and β around 0.5-0.6 consistently produce the highest *SCF*. As it does in training data, RDA also produces the highest cross-validated *FVE* for *a*=*b*=0.5, however intermediate solutions can produce slightly higher values for larger amplitudes. With *a*=*b*, the signal itself is symmetric and therefore symmetric solutions produce the highest values of *FSVE*, *COR* and *SCF*.

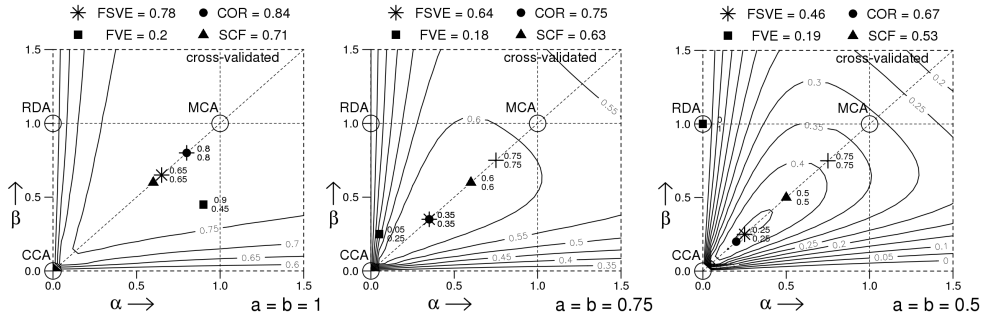

Figure 4.2 Same as Fig. 4.1, but computed after leave-one-out cross-validation.

In order to assess the role of sample size, SSVD-1 is alternatively computed with well-sampled independent datasets of length *n*=300 (rather than *n*=30), again for *a* = *b* = 1, 0.75, 0.5, and results are averaged across 10 independent realizations. As expected with longer samples, *FSVE*, *COR*, *SCF* and *FVE* are consistently much higher, and there is less contrast between solutions for higher amplitude signals (Fig. 4.3). The representation of the signal is more important, and unsurprisingly MCA-1 generally has the highest skill in capturing the signal if it is represented by MCA-1 in the absence of noise. MCA-1 also tends to yield the highest *SCF* for lower amplitude signals. However, consistent with smaller sample sizes (Fig. 4.2), similar intermediate solutions interestingly yield the highest *FSVE* and *COR*, somewhat higher than that for CCA-1. In general, some benefit of amplitude scaling used by

SSVD is independent to the sample size, and thus SSVD may provide an advantage for isolating well-sampled coupled signals as well.

If one considers taking the approach of Barnett & Preisendorfer (1987) with EOF truncation prior to computing SSVD, it is no surprise that retaining the 4 leading EOFs (roughly 60% of the total variance) yields the best results (Fig. 4.4) considering that the true signal involves the first 4 population EOFs. As expected, solutions including CCA and RDA that involve the whitening of one or both variables (α=0 or β=0) are much more comparable. Pre-filtering yields some improvement in isolating the signal, especially for *a*=*b*=0.5 for which CCA yields the highest skill and highest correlation. On the other hand, for *a*=*b*=1,0.75, intermediate solutions still produce the highest values for *FSVE*, *COR* and *SCF*, although the coordinates of maxima are closer to those of CCA and MCA, especially for *a*=*b*=0.75 for which only a minimal degree of amplitude gives slightly better results than CCA. In these cases, all SSVD-1 solutions generally give very similar results.

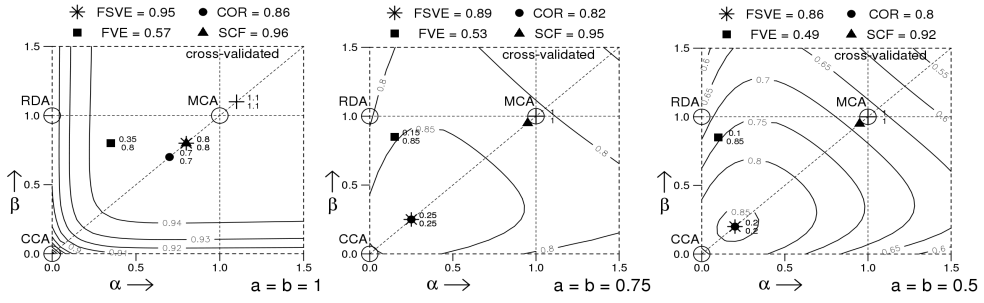

Figure 4.3 Same as Fig. 4.2, but averaged across 10 well-sampled synthetic realizations of length *n*=300. The FSVE contour interval for *a*=*b*=1 (left) is decreased from 0.05 to 0.01.

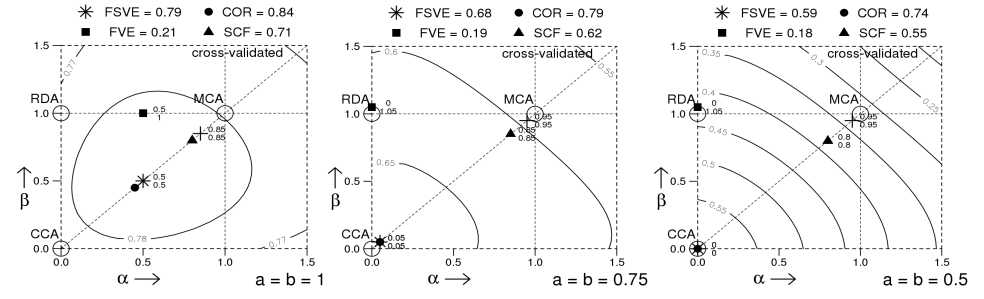

Figure 4.4 Same as Fig. 4.2, although datasets are truncated to their first 4 leading EOFs prior to computing SSVD. The *FSVE* contour interval for *a*=*b*=1 (left) is decreased from 0.05 to 0.01.

For many practical scenarios, it is not uncommon for the signal-to-noise ratio of *X* to be appreciably different than that for *Y* considering that the nature of background noise may be fundamentally different. In order to address this, a few sets of synthetic SSVD experiments are repeated for which $a \neq b$. With an asymmetry in amplitude, a hypothesis is that SSVD-1 solutions with asymmetric scaling ($\alpha \neq \beta$) will produce better results. Drawing from the previous results, skill may be presumably higher for SSVD-1 using scaling values directly proportionate to signal amplitude for individual variables as well. The cases considered are for a=1, b=0.5; a=0.75, b=0.25 and a=0, b=1 (Fig. 4.5), and results generally support the hypothesis. For $a > b$ ($a < b$), the highest values for *FSVE*, *COR* and *SCF* are produced by choosing scaling values such that $\alpha > \beta$ ($\alpha < \beta$). SSVD-1 results for a=0.75, b=0.25 and a=0, b=1 are in strong contrast to each other. For scaling values of α=0.4, β=0.25, SSVD-1 yields a maximum *FSVE* of 0.45 when a=0.75, b=0.25, however it tends to diverge and produce no skill when a=0, b=1. Conversely for scaling values of α=0.35, β=1.05, SSVD-1 yields a maximum *FSVE* of 0.44 when a=0, b=1 yet no skill when a=0.75, b=0.25. In consideration of this systematic sensitivity, SSVD-1 is able to capture at least twice as much signal variance than MCA-1 for solutions with significant asymmetry in signal amplitude. This sensitivity holds for symmetric signals, but to a much lesser degree. Considering a single appropriate set of solutions that fall within the range of highest skill, given an appreciably higher signal-to-noise ratio for *X* compared with that in *Y*, an appropriate choice is α=1/2, β=1/4 (and vice-versa).

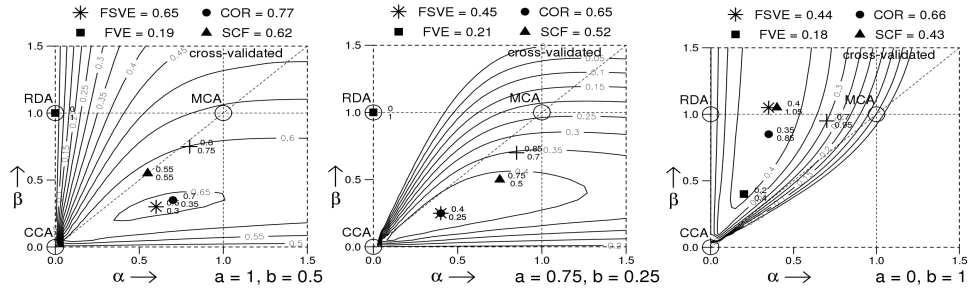

Figure 4.5 Fraction of signal variance explained (*FSVE*, contours at intervals of 0.05) for SSVD-1 plotted as a function of scaling parameters α and β averaged across 100 synthetic realizations for which the fraction of EOF-1 for the signal in *X* and *Y* is specified as a = 1, b = 0.5 (left column), a = 0.75, b = 0.25 (middle column) and a = 0, b = 1 (right column), respectively. Plotted are the α-β coordinates for maximum *FSVE*, *FSVE* of the alternate signal (MCA-1 of data without noise), correlation (*COR*), squared covariance fraction (*SCF*), and variance of *Y* explained by *X* (*FVE*) as indicated by an asterisk, plus, solid dot, solid triangle and solid square, respectively. The corresponding values are given above the plot.

Also, the considering when a=0.75, b=0.25, a fairly representative situation with differential signal-to-noise ratios, intermediate solutions generally following β/α=1/2 produce the highest skill, particularly solutions with very low scaling values. Note that in general, the median location of highest skill on the α-β plane is lower than the mean location. Also consistent with previous results, it is evident that RDA produces the highest cross-validated *FVE* unless the amplitude of the signal in *Y* is appreciable, for which SSVD-1 using intermediate scaling values (with nonzero amplitude given to *X*) produces higher *FVE*. Although as for CCA, RDA is more usefully applied to pre-filtered datasets.

Considering SSVD-1 of data truncated to retain only the leading four EOFs, results are investigated for the same cases (Fig. 4.6). It is clear that when the amplitude of one variable is appreciably less than the other, the highest skill (*FSVE*) is found when the amplitude of the weaker amplitude variable is neglected entirely (with prior whitening), while the stronger amplitude variable is scaled down to some degree. Again, with EOF pre-filtering the skill for the range of SSVD solutions is not much better than that for CCA-1, and BP's approach to CCA is a useful and consistent option. This is also true for RDA, however some systematic scaling is found to increase *FVE*.

Various other cases are tested (not shown) and generally verify the findings presented here. Of course, one must keep in mind that like MCA and CCA, SSVD is a numerical algorithm that *always* produces a result, even when no true coupled signal is present or the true coupled signal has such small amplitude that it is undetectable. However, in these situations, measures of relation for SSVD-1 dramatically decrease to insignificant values after cross-validation.

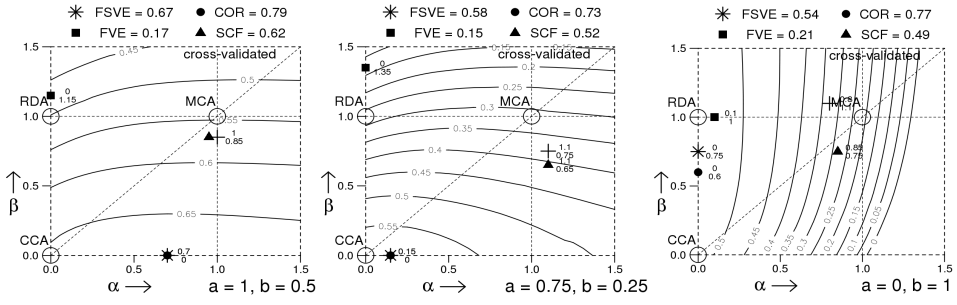

Figure 4.6 Same as Fig. 4.5, although SSVD is computed on data truncated to retain only the first 4 leading EOFs.

There are two possible caveats for the representation of the coupled signals here. One is that they only contain information from the leading 4 population EOFs, such that EOF truncation is a particularly useful prior constraint and BP's approach using CCA has high skill in converging to the coupled signal. In real applications, it is likely that details of a coupled signal are present that project on the trailing EOFs neglected with truncation. For such situations, SSVD is more suitable and closer convergence to a true signal is possible considering SSVD involves the entire dataset.

5. APPLICATION TO TROPICAL-EXTRATROPICAL TELECONNECTIONS

With the goal of better representing important physical modes of tropical-extratropical interaction on inter-annual time scales, SSVD is used here to capture the dominant linear relationships between Northern Hemispheric (NH) 500 hPa geopotential height (Z_{500}) and tropical Pacific precipitation (PREC) during boreal winter. SSVD solutions are contrasted with those of MCA and CCA. Z_{500} is chosen to best represent fluctuations in the extratropical tropospheric circulation as it has a direct and clearly understood association with both jet-level dynamics and lower-level thermodynamics. Out of the paradigm of tropical forcing, the significant atmospheric driver of year-to-year variability is the redistribution of latent heat release associated with seasonal changes in tropical precipitation (PREC) coupled with the slowly-evolving underlying sea surface temperature (SST), particularly across the tropical Pacific. In the upper-troposphere, convectively-driven temperature changes are translated into vertical motion and sources of anomalous vorticity that excite both equatorial waves and off-equatorial Rossby waves that can emanate poleward and substantially alter the extratropical circulation during boreal wintertime. On the other hand, the extratropics can also plausibly drive changes in tropical Pacific convection, the balancing circulation, and the underlying SST.

In this study, the focus on PREC, rather SST as is typically considered, is physically justified to be more closely linked with the extratropics, and any potentially complicating thermodynamic aspects of the local SST-PREC relationship may not be considered.

The leading relationships that are examined isolate the dominant ENSO signal as well as the extratropical circulation changes associated with ENSO Modoki and/or flavors of ENSO. The former relationship involves AO/NAO-like variability and is verified by a handful of past studies with some physical justification that is still not yet fully established. This mode likely represents two-way coupling of a mixture of processes. SSVD solutions are demonstrated to represent these modes with much (slightly) higher cross-validated correlation (covariance), particularly asymmetric solutions with fairly small scaling values.

5.1 Datasets & methodology

The observed datasets used in this study are NH (20°N - 90°N) Z_{500} from the NCEP-DOE AMIP-II Reanalysis (Kanamitsu et al. 2002) and PREC from the CPC Merged Analysis of Precipitation (CMAP, Xie & Arkin 1997) spanning the tropical Pacific including the Maritime Continent and eastern Indian Ocean (20°S - 20°N, 90°E - 90°W). Both datasets span only the recent satellite-era (1979-2011). Note that results are generally insensitive to the choice of precipitation product and are reproducible using precipitation estimated from the Global Precipitation Climatology Project (GPCP, Adler et al. 2003), despite GPCP having comparatively reduced variability over tropical oceans (Yin et al. 2004). Z_{500} and PREC co-variability generally peaks in mid-winter (January-February, JF), with non-ENSO co-variability actually peaking when PREC lags Z_{500}. With a focus beyond ENSO, here a focus is on co-variability between FM PREC and JF Z_{500}, however relationships are generally robust for a small range of lead/lags across boreal winter. In addition to the fixed climatological mean, trends and decadal variability across the 33-year period are removed prior by subtracting a least squares fit to the first three Legendre polynomials. Note that PREC is first linearly interpolated, or offset rather, to a 2.5° grid (precisely intersecting dateline and equator) in order to relate with output from multiple coupled general circulation models, as discussed later. The spatial patterns and associated variates are examined alongside the traditional measures of relation: correlation (*COR*), squared covariance fraction (*SCF*), and the fraction of variance of *Y* explained by *X* (*FVE*). Results are examined using a leave-3-out cross-validation procedure for which SSVD is computed

on all but 3 years. Anomalies of the middle independent year are then projected onto the leading SSVD modes, and the procedure is iterated across the entire sample through which independent projections are related with each other across the whole sample. SSVD solutions are contrasted with those of MCA and CCA. Both the proposed symmetric (α=β=1/2) and asymmetric (α=1/2, β=1/4) solutions of SSVD are considered. Based on the prior knowledge that the extratropical circulation by nature contains much more unrelated noise in contrast to the tropics, an asymmetric method that scales down Z_{500} disproportionately more than PREC is used, and it is hypothesized to give stronger results.

5.2 Results

First examining the seasonality of leading SSVD modes, it is clear that co-variability peaks in JF for PREC and Z_{500} simultaneously (not shown). It remains high when PREC is lagged from 1 to 3 months, particularly for the second mode involving ENSO Modoki. This range spanning boreal mid-winter to spring is generally characterized by the same statistically robust pairs of patterns involving ENSO and ENSO Modoki teleconnections. This is best characterized by relationships between JF Z_{500} and FM PREC, which are investigated here in more detail. The ENSO signal, reflected by MCA-1, the two SSVD-1 solutions, and CCA-1 using a prior EOF truncation number of 5 (Fig. 5.1), all pairs of patterns are highly similar and link the eastward shift in precipitation across the Pacific (about a third of total PREC variance) to the classic wave train pattern in Z_{500} that characterizes a stationary Rossby wave response dominated by low heights in the Northeast Pacific (15-20% of total Z_{500} variance). Considering the magnitude of the primary high and low height centers associated with the wave train, SSVD acts to decrease the magnitude compared with MCA. Asymmetric scaling further decreases the amplitude of the height signal (±10 m relative to MCA-1), although keep in mind this is not a broad linear change in amplitude. CCA-1 is fairly similar with some downstream discrepancy over Eurasia. In tropical PREC, MCA-1 and SSVD-1 solutions have some contrast with CCA-1 with a stronger positive anomalous rainfall pattern in the central Pacific and more suppression over the southward shift in the ITCZ east of the dateline. SSVD-1 also has a stronger

and more spatially smoother region of suppression in the West Pacific. Although differences in ENSO spatial patterns are fairly subtle, co-variablity in terms of *COR* is much higher for SSVD-1 compared with MCA-1 and CCA-1, and in terms of *SCF* it is comparable with all around 0.5 although slightly higher for the asymmetric SSVD solution, which also has the highest correlation of 0.75 (Table 5.1).

The ENSO Modoki relationship with the extratropics involves a tripole structure in PREC for which changes in the west/central Pacific is out of phase with weaker changes in western and eastern Pacific (Fig. 5.2). This PREC structure co-varies with a Z_{500} pattern projecting onto the AO and NAO yet with more action over the Atlantic and Pacific fringes of the Arctic. This appears to be a mixture of multiple phenomena and has less clear physical interpretation.

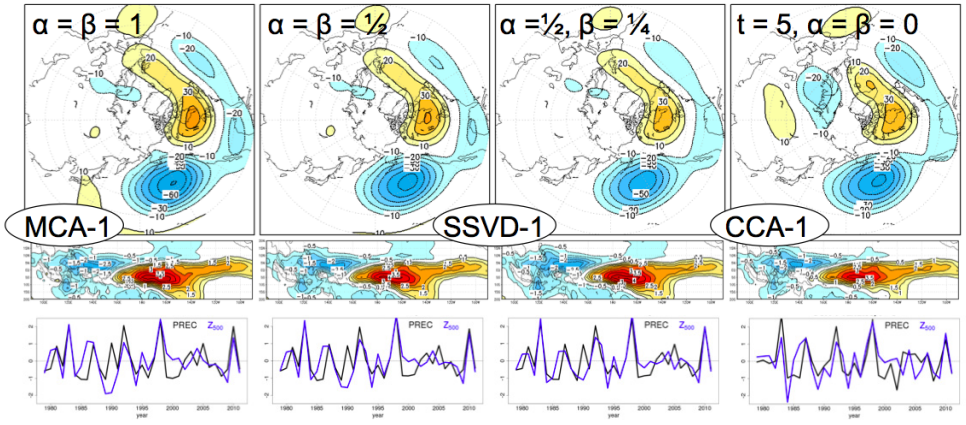

Figure 5.1 SSVD-1 JF Z_{500} patterns (top row), FM PREC patterns (middle row), and the associated variates (Z_{500} in black, PREC in blue). Columns correspond to MCA-1 (left), intermediate solutions of SSVD-1 using α=β=1/2 (middle left) and α=1/2, β=1/4 (middle right), and CCA-1 using a prior EOF truncation number of 5 for both PREC and Z_{500} (right).

Table 5.1 Cross-validated values for correlation (*COR*) and squared covariance fraction (*SCF*) rounded to the 2nd decimal place for SSVD-1 and related solutions of Fig. 5.1.

measure of relation	MCA-1	SSVD-1 α=β=1/2	SSVD-1 α=1/2, β=1/4	CCA-1 t = 5
COR	0.60	0.70	0.75	0.63
SCF	0.48	0.49	0.50	0.49

ENSO Modoki itself represents warm pool El Niño events, but also is strongly involved with significant La Niña events as it may be considered to absorb some asymmetry of ENSO. The AO-like/NAO pattern is fairly robust, yet projects much more on the AO in early winter (not shown). Wu (2010) discusses this observed association with ENSO Modoki similarly using MCA but examining co-variability with SST rather across a much longer time period. He argues that changes in the extratropical circulation lead and apparently having an important role as forcing on the tropics during the winter season, although much of the persistence is likely attributable to tropical forcing such that this is a two-way coupled phenomenon. Li et al. (2006) confirm that SST warming associated with a similar structure in precipitation produces an AO-like/NAO response very similar to the pattern show here with MCA-2 and SSVD-2 solutions. It is obvious the CCA-2 diverges from MCA and SSVD solutions, and the Z_{500} north-south gradient for CCA-2 is displaced about 150 equatorward with a pattern with much less resemblance to the AO at high latitudes. Also, there is stronger variability over Siberia also. Modes in Z_{500} all explain about 15% of the total variance, whereas CCA-2 for PREC is much weaker explaining about half of the 20% of total variance explained by MCA-2 and SSVD-2 solutions. Again, SSVD-2 has more subtle differences compared with MCA-2, mainly less changes in subtropical precipitation associated with a weakened North Pacific Z_{500} gradient and a slightly weakened NAO signal, differences that are enhanced when using asymmetric scaling. Consistent with the explained variance, MCA-2 and SSVD-2 solutions have roughly the same *SCF*, almost twice as much as that for CCA-2 (Table 5.2). However, again SSVD-2 solutions have much higher *COR*, particularly for the asymmetric solution (value of 0.76 compared with < 0.6 for MCA-2 and CCA-2). This proves the hypothesis stated earlier and is consistent with the results from synthetic data. It is plausible that as the ENSO and ENSO Modoki signals in the extratropics are much smaller relative to unrelated internal variability, amplitude is less of a distinguishing factor for their teleconnection signals and the linear relationships can be more strongly represented by scaling down extratropical variability when relating with tropical precipitation.

Considering the range of SSVD solutions, *COR*, *SCF* and *FVE* are examined in a similar manner as with the solutions from synthetic data as a function of α and

β, with a focus on *COR* (Fig. 5.3). In the training data, CCA, MCA and RDA are again optimal in yielding the highest respective measures of relation. Intermediate SSVD solutions have intermediate values of *COR*. In general, α<1/2 or β<1/2 produces training *COR* values of 0.9 or greater, however after cross-validation, values remain relatively high, still reaching 0.78 and 0.85 for SSVD-1 and SSVD-2, respectively, much higher the MCA solutions that are 0.6 or less. Consistent with synthetic results of Fig. 4.5, small scaling values along β/α=1/2 yield the highest values. Also consistent, larger values of scaling produce the highest *SCF*. The highest cross-validated *FVE* is not found for RDA, however the SSVD-1 solution with the highest *FVE* alone is inconsistent with that for the sum of SSVD-1 and SSVD-2. Keep in mind that such an analysis is useful for contrasting SSVD particular solutions, although for a single analysis, scaling values that produce the highest values of relation may be sensitive to sampling. Again, the choice of method must be independent from the results. However, confirmation of the relationships determined in synthetic data in this real observed relationship in the climate system provides further support and certainly reinforces the practical use of SSVD for relating two variables and isolating coupled signals in data. In the next section, a practical application of SSVD is tested.

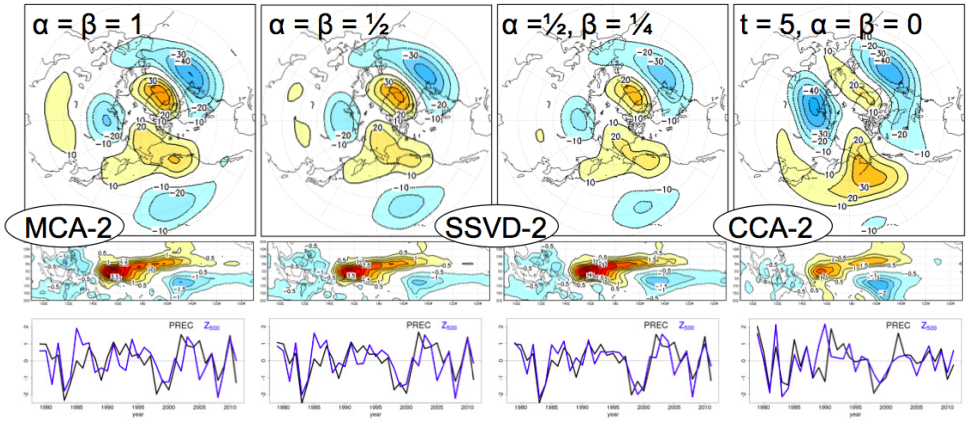

Figure 5.2 Same as Fig. 5.1 but for SSVD-2 solutions.

Table 5.2 Cross-validated values for correlation (*COR*) and squared covariance fraction (*SCF*) rounded to the 2nd decimal place for SSVD-2 and related solutions of Fig. 5.2.

measure of relation	MCA-2	SSVD-1 $\alpha=\beta=1/2$	SSVD-1 $\alpha=1/2,\ \beta=1/4$	CCA-2 t = 5
COR	0.51	0.65	0.76	0.59
SCF	0.24	0.24	0.24	0.13

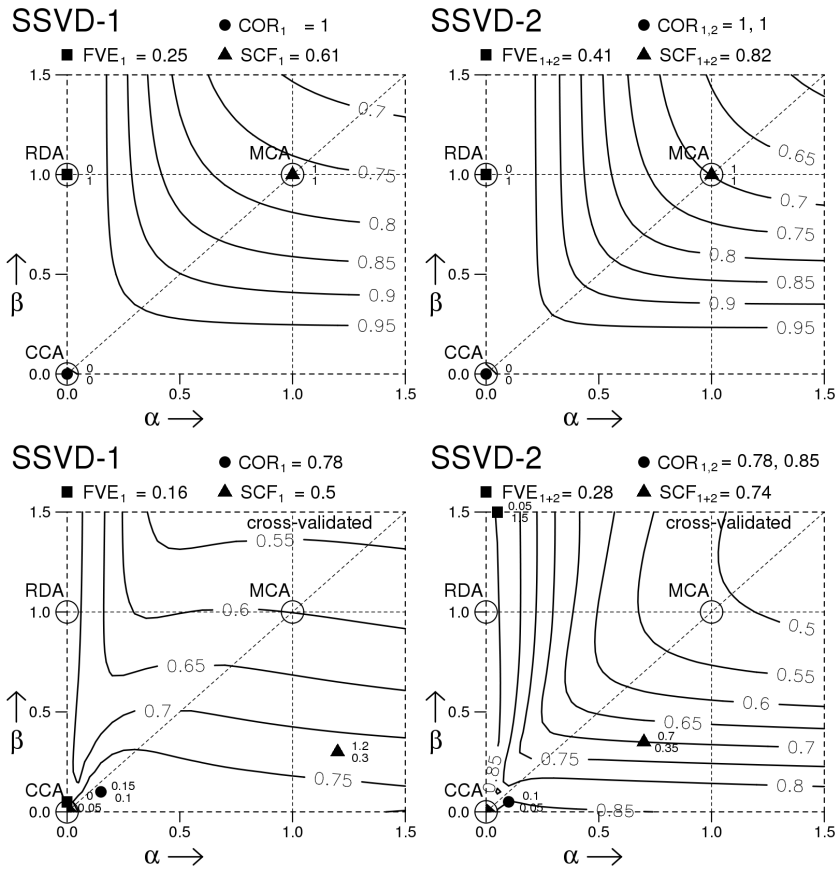

Figure 5.3 Correlation (*COR*, contours) for SSVD-1 (left) and SSVD-2 (right) plotted as a function of scaling parameters α and β. Values computed from training data are plotted in the upper row whereas cross-validated values are plotted in the lower row. α-β coordinates for maximum *COR*, *SCF*, and variance of JF Z_{500} explained by FM PREC (*FVE*) are indicated by a solid dot, solid triangle and solid square, and the corresponding values are given above the plots, rounded to the 2nd decimal place. Values alongside SSVD-2 are maximum values in terms of the combined contribution of both SSVD-1 and SSVD-2.

6. PRELIMINARY RESULTS OF STATISTICAL DOWNSCALING

In this section, SSVD and related methods are applied to statistical downscaling in effort to improve seasonal prediction skill of precipitation over the Maritime Continent during June-August (JJA). In this region, seasonal predictability is relatively high due to the direct impact of tropical SST, largely associated with ENSO, and general circulation models (GCMs) generally have high skill. Still, model bias including that simply related to resolution or lack of local topography may degrade skill. Such degradation may sometimes be corrected using statistical downscaling.

Here the Model Output Statistics (MOS, Wilks 1995) approach is used in which GCM-predicted tropical Indo-Pacific (60°E - 120°W, 15°S - 15°N) SST during JJA is regressed with the simultaneously observed land CRU 0.5° precipitation (Mitchell et al. 2003) over the Maritime Continent (roughly enclosed by 90°E - 130°E, 10°S - 20°N). By design, this approach is capable of eliminating a degree of systematic model bias. The predicted SST is taken as the multi-model ensemble (MME) mean of seasonal anomalies (with respect to individual GCM climatology) predicted by 4 coupled GCMs of the APCC hindcast dataset, namely the APCC model (CAM3 coupled with POP using 5 ensemble members), the CanCM4 model (Meteorological Service of Canada, 10 ensemble members), the NASA GMAO GEOS-5 model (9 ensemble members), and the NCEP CFS (20 ensemble members). The GCMs are chosen based on the fact that they have higher skill and span a longer 26-year common period, 1983-2008, the time period used for this study. Hindcast simulations are initialized at a one month lead on May 1st. Results are examined in terms of correlation and root mean squared error (*RMSE*) with respect to observed precipitation, and values are cross-validated using leave-3-out approach as discussed previously.

The benchmark for statistical downscaling results is the MME-simulated precipitation (direct model output) interpolated to the 0.5° resolution for comparison with CRU observations. The methods used for statistical downscaling include regression at individual grid points (Grid point) using the leading principal components of SST, principal component regression (PCR, involving a subset of PCs for both SST and precipitation), CCA, MCA, RDA, as well as SSVD using two previously-proposed sets

of scaling values (α=β=1/2 and α=1/4, β=1/2) and a set of solutions involving very small scaling values (α=10^{-5}, β=2×10^{-5}) while maintaining the same scaling ratio of α/β=1/2. Such small values are consistent with results of the previous section - that retaining only a tiny fraction of amplitude can act as a sufficient constraint to filter noise. However, values of α=10^{-5}, β=2×10^{-5} are chosen here with a degree of subjectivity because they give appreciably different results compared with the other asymmetric approach. All pattern-based methods are applied in an identical fashion, as different approaches to choose a *subset* of predictors and predictands through which to compute regression. The subset of predictors, or number of mode pairs used, is chosen based on that which gives the highest skill (in terms of area with significant correlation). For all methods except MCA and SSVD, EOF truncation is used with the truncation number also based on the highest skill. In general, these skill metrics and this approach for comparing methods mimic the format of T2008.

By initial examination of JJA precipitation prediction over the Maritime Continent (Fig. 6.1), it is clear that the direct model output performs better than statistical downscaling using the GCM-predicted SST. Unlike downscaling methods, the MME has scattered regions of skill over parts of Malaysia, northern Sumatra, Java, and eastern Papua New Guinea. Also, higher skill is found in the most skillful region of West Kalimantan. Downscaling methods are all generally comparable to each other. It is also surprising that the simplest approach, grid point regression, slightly outperforms MCA on average. The best downscaling approach is a solution for which CCA is equivalent to RDA, which does improve correlations in parts of eastern Malaysia and northern Sumatra. In terms of area of significant correlation, asymmetric solutions of SSVD also perform better than other methods, however this is not true for the total *RMSE*. In general, it appears that statistical downscaling does not add much value over the Maritime Continent for JJA. One reason for this is demonstrated next.

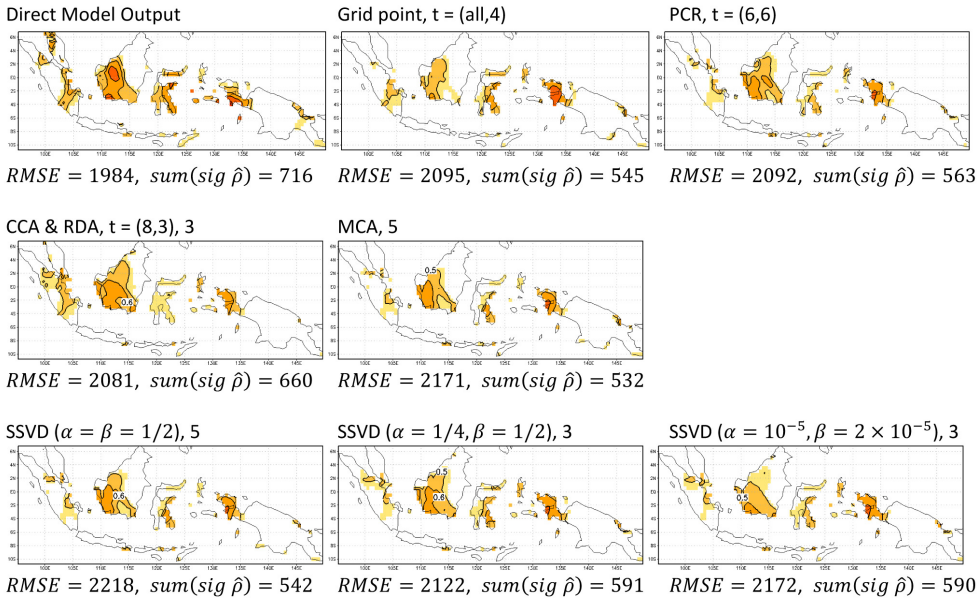

Figure 6.1 Cross-validated correlation (*COR*) between predicted and observed precipitation plotted for MME direct model output (upper left), regression at each grid point (upper middle), PCR (upper right), CCA & RDA (in this case equivalent, middle left), MCA (middle, middle), SSVD using α=β=1/2 (lower left), SSVD using α=1/4, β=1/2 (lower middle), and SSVD using α=10^{-5}, β=2×10^{-5} (lower right). Only statistically significant correlation values (> 0.36) are plotted. If EOF truncation is used, it is indicated above the plots by t = (truncation for *X*, truncation for *Y*). Number of mode pairs used for regression is also indicated above plots. Below each plot are two measures of skill: *RMSE* and the sum of the number grid points with significant correlation.

It is plausible that low frequency variability across the time period impacts rainfall yet is not captured by inter-annual co-variation with SST such as ENSO. In order to investigate low frequency variability of precipitation over the Maritime Continent, the precipitation trend and decadal variation is examined alongside that of SST (Fig. 6.2), computed by regressing each grid point with the first two Legendre polynomials (linear trend + quadratic term). After removing this variability, RDA is computed and the first three modes are examined (Fig. 6.3). As is clear, RDA-2 captures the ENSO-forced response associated with dry (wet) conditions during El Niño (La Niña). In SST, ENSO appears to project onto decadal variation involving central Pacific cooling (warming) predicted in the early 1980s and late 2000s (1990s) coincident with strong wet (dry) conditions over much of the Maritime Continent. However, the decadal signal in SST has more action in the central Pacific consistent with different flavors

of ENSO (Kau & Yu 2009, Kug et al. 2009). The ENSO-forced response projects onto the decadal precipitation signal although is much weaker than the decadal signal and does not include some variability of the opposite sign over Papua New Guinea. Confirmed with RDA in the prior analysis (not shown), the inter-annual ENSO relationship cannot account for the strong decadal fluctuation in precipitation, which is not represented in additional modes likely due to orthogonality constraints. As for the other RDA modes, it appears that RDA-1 does not contribute much to skill and may not represent anything physical, although RDA-3 may link Indian Ocean basin-wide warming to enhanced precipitation over Kalimantan. In general, the other pattern-based methods produce qualitatively similar pairs of modes (not shown).

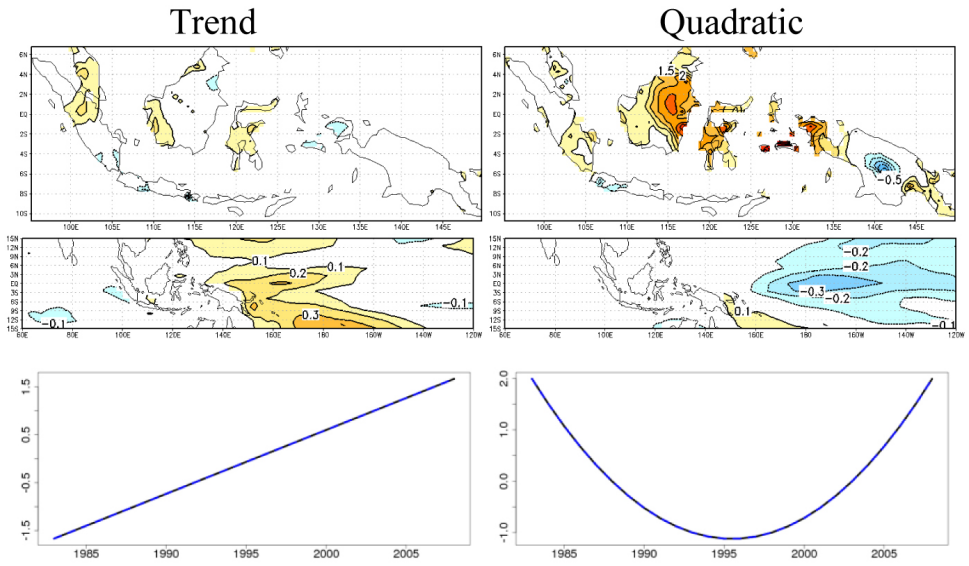

Figure 6.2 Observed precipitation (top row) and model-predicted SST (middle row) regressed to linear trend (left column) and quadratic term (right column) with temporal variation (bottom row).

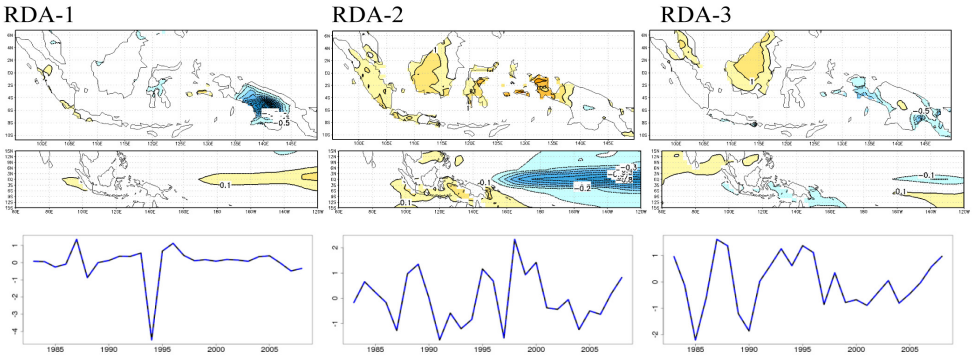

Figure 6.3 RDA-1 (left column), RDA-2 (middle column) and RDA-3 (right column) patterns for observed precipitation (top row) and model-predicted SST (middle row) alongside RDA variates (from training data, bottom row). Prior to analysis, linear trend and decadal variability shown in Fig. 6.2 are removed.

If one desires to predict only the interannual component of JJA precipitation variability over the Maritime Continent, statistical downscaling may provide some benefit over dynamical MME prediction. Downscaling is repeated using the various approaches after removing the trend and decadal variability, and downscaling methods are indeed more comparable to direct model output (Fig. 6.4). The MME prediction alone is fairly skillful, especially over West Kalimantan, although there is generally no skill in northeastern Kalimantan and central New Guinea.

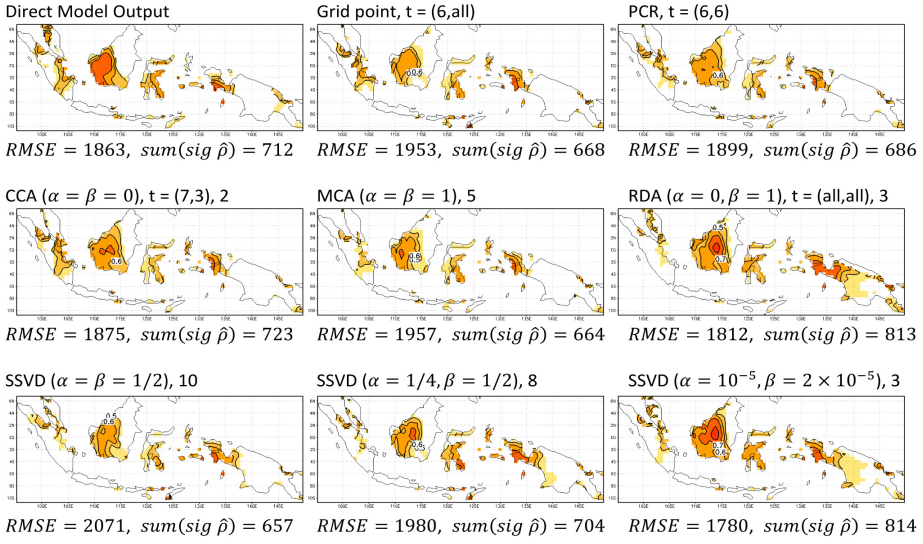

Figure 6.4 Same as Fig. 6.1, but skill is computed with the trend and decadal variability (Fig. 6.2) removed prior to the analysis. Also, the best solutions for CCA (middle left) and RDA (middle right) are distinct.

Again, grid point regression produces very similar skill with MCA, skill slightly trailing that of PCR, something that indicates a benefit of EOF-filtering the MME-predicted SST. Still, these methods do not improve the MME prediction of precipitation. On the other hand, CCA produces comparable results, and RDA appreciably enhances skill over the MME prediction. Comparison of methods is similar to the results of T2008 who use statistical downscaling to improve MAM precipitation prediction over NE Brazil, yet RDA provides more benefit in this case. Still the net result is promising, particularly due to the fact that the direct MME precipitation prediction is already skillful. Much of this improvement using RDA is found in areas of the strong decadal signal, namely part of northeastern Kalimantan and Papua New Guinea. Again, keep in mind that skill is computed with respect to the observed precipitation *without* its decadal component, and generally this improvement from downscaling is masked by its lack of representation of the decadal variation. For SSVD, the symmetric solution (α=β=1/2) is the least skillful for downscaling, consistent with MCA, whereas the asymmetric solution with the amplitude of SST scaled further (α=1/4, β=1/2) acts to improve the skill a bit, more comparable with CCA and direct model output. By using very small scaling values (α=10^{-5}, β=2×10^{-5}) while maintaining α/β=1/2, skill of the MME-predicted precipitation is also improved appreciably. This improvement is comparable to RDA in terms of correlated area, however the most skillful region over Kalimantan is improved further, consistent with a lower *RMSE*. Keep in mind that direct model output is slightly degraded by statistical downscaling in some regions (e.g. southwestern Kalimantan), where local dynamical processes possibly provide more skill. Also, results here are only beneficial if the decadal component is predictable, which may be estimated using other approaches, although none have been tested here. Still, these results certainly encourage the application of SSVD and related methods to statistical downscaling, which is proposed for further future research particularly to regions/seasons that have much less skill for MME prediction.

7. SUMMARY AND CONCLUSIONS

Considering that common univariate measures of relation vary only in terms of the role of amplitude, they may be generalized into one generic form introduced here as scaled covariance. As demonstrated, scaled covariance is equivalent to covariance between scaled variables, with the scaling transformation acting to apply a specified degree of normalization and decorrelation. For intermediate scaling values, this is referred to as partial whitening, a transformation that bridges full whitening and the identity matrix (no transformation). Depending on scaling values, this transformation is suggested to be an effective noise filter for examining covariance between two variables. Given its exponential form, scaling may provide a more objective means of disproportionately damping noise, and is suggested as a useful alternative to EOF truncation (Barnett & Preisendorfer 1987) that completely neglect portions of variability. Scaled covariance is maximized by Scaled Singular Value Decomposition Analysis (SSVD), an approach that provides essentially a continuum of methods linking CCA, MCA and RDA. Scaled covariance varies in terms of its variable specification of the role of amplitude according to scaling parameters α and β. SSVD basis sets maintain strict orthogonality properties only for α=0,1 or β=0,1 solutions for which CCA, MCA and RDA are unique solutions. Also, scaling values of α=1/2 or β=1/2 remove amplitude bias due to the fact the *squared* covariance is maximized with SVD (as in MCA) rather than absolute covariance.

In summary, two primary findings arise from synthetic experiments. First, intermediate solutions of SSVD tend to yield the highest skill in isolating the coupled signal. Such SSVD solutions also produce the highest cross-validated values of correlation and covariance, interestingly appreciably higher than those for CCA and MCA, methods that are optimal in training data. This strongly suggests that such SSVD solutions filter unrelated noise much more effectively when representing sample relationships. This is consistent with the notion that disproportionate amplitude damping through partial whitening can help distinguish signal from noise. Second, scaling values associated with the highest skill in isolating the coupled signal are proportional to the signal-to-noise ratio. In other words, SSVD converges closer to a coupled signal with low (high) amplitude with respect to background noise if a low (high) value of scaling is used within the intermediate range. If the signal-to-noise

ratio for *X* is appreciably different than that for *Y*, asymmetric solutions ($\alpha \neq \beta$) are preferable using values of either 1/2 or 1/4. In general, solutions along α/β=1/2 or β/α=1/2 are most skillful, and the median α-β locations of highest skill and correlation are generally associated with smaller scaling values. Ideally, a rough unbiased prior expectation of the signal-to-noise ratio for each variable is necessary for specifying the scaling values, and it is generally not advisable to simply find the scaling values that produce SSVD solutions with the highest cross-validated values of measures of relation in a particular set of sample datasets. As screening in this way is not random nor are differing SSVD solutions independent, this procedure itself violates statistical assumptions of estimating sample statistics, and can plausibly inflate values artificially. An objective basis for choosing scaling parameters is important.

SSVD is applied alongside MCA and CCA to isolate leading modes of co-variability between the tropics and extratropics during boreal winter, namely ENSO and ENSO Modoki teleconnections. Two of the primary symmetric and asymmetric solutions of SSVD that are justified from results of the synthetic experiments are used. Based on the prior knowledge that the extratropical circulation is more dominated by internal noise unrelated to the tropics, the asymmetric method is hypothesized to provide a stronger representation of the observed relationships. SSVD is shown to produce stronger tropical-extratropical teleconnection relationships involving ENSO and ENSO Modoki with significantly higher cross-validation correlations. Consistent with the hypothesis, the asymmetric SSVD solution (α=1/2, β=1/4) produces the highest correlation for both modes. The amount of covariance explained is very similar for all methods (with exception to CCA-2 which deviates from MCA-2 and SSVD-2 solutions), although the asymmetric SSVD solution also has slightly higher *SCF*. SSVD has a slightly modified representation of the relationships with less amplitude given to extratropical circulation patterns associated with a stronger, more coherent accompanying PREC signal in the West Pacific during ENSO, and less variation over the eastern Pacific ITCZ during ENSO Modoki. One may speculate to whether or not these differences reveal details of overfitting across the sample, given that SSVD patterns are much more robust with cross-validation. In light of constraints for SSVD, multiple co-existing linear relationships cannot be captured exactly. Based on cross-validation, if one assumes adequate convergence to the real physical relationships involving

ENSO and ENSO Modoki/flavors of ENSO, then one may presume based on appreciable increases in correlation that SSVD patterns give a more realistic representation of these relationships. By examining the range of solutions, results are generally consistent with the synthetic experiments. For very small scaling values along β/α=1/2, SSVD produces the strongest relationships and yield correlation values as high as 0.78 and 0.85 for ENSO and ENSO Modoki teleconnections, respectively. This is quite promising and reinforces the use of small scaling values as well. Further research is needed to add confidence to the nature of the association between ENSO Modoki and AO-like/NAO circulation changes, as this mode may be very relevant to seasonal predictability apart from the dominant ENSO signal. Also, implications of misrepresenting these teleconnections in GCMs used for seasonal prediction need to be investigated and better estimation of these observed relationships would aide phenomena-based diagnostic studies. Such diagnostic studies are a future topic of research for applying SSVD.

Statistical downscaling experiments are conducted over the Maritime Continent during JJA towards improving the seasonal prediction of precipitation of the MME. From this, it is apparent that the direct MME prediction is fairly skillful over this region, presumably due to its direct proximity to the sources of predictability (e.g., ENSO), and unfortunately downscaling does not improve skill partly due to its inability to capture a strong decadal signal in rainfall. However, in terms of the inter-annual variability, statistical downscaling can improve the prediction (if the trend and decadal variability are removed prior). RDA adds skill to the MME, and SSVD using small asymmetric scaling values can improve skill even further given that SSVD produces the lowest *RMSE*. This result illustrates the benefit of SSVD and encourages further application to statistical downscaling, for large-scale seasonal prediction as well as alternative regional contexts including finer-resolution hydrological prediction. Hypothetically speaking, more benefit of statistical downscaling with SSVD may be found if applied to regions/seasons for which MME seasonal prediction has very little skill, such as South Asia during the summer monsoon (Tung et al. 2013). This is left for further research, and a goal is to implement SSVD and/or related methods into operational use at APCC. The prospect of improving seasonal prediction skill with statistical downscaling as an aide to the raw MME output motivates future work as it is clearly in the best interest of those that rely on accurate seasonal forecasts.

APPENDIX A

Scaled SVD Analysis in PC-space

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a useful technique for data compression that decomposes a variable into a unique orthogonal basis set with leading modes that maximize variance in a consecutive manner. PCA is traditionally computed through an eigenvector decomposition of the covariance matrix, although PCA may also be computed using singular value decomposition (SVD) as shown by Wilks (2005). First one may consider datasets X and Y, matrices of anomalies with dimensions of space by time ($m_x \times n$ and $m_y \times n$, respectively). Again, uniform spatial weighting is assumed and fields are multiplied by a factor of $1/\sqrt{n}$. Also, the number of spatial grid points is assumed to be greater than the number of time samples with $m_x > n$ and $m_y > n$. One way of computing individual PCA for X and Y involves taking SVD resulting in $X=U_xS_xV_x^T$ and $Y=U_yS_yV_y^T$, from which the weighted pattern vectors (EOF patterns) are given simply by U_xS_x and U_yS_y with associated variates (PCs) given by V_x and V_y, respectively. For most practical situations, the number of spatial grid points is greater than the number of time samples ($m_x > n$ and $m_y > n$), and in this case an *economy* SVD is computed such that U_x is *not* square with dimensions of $m_x \times n$ such that $U_x^TU_x=I$ but $U_x^TU_x \neq I$. For this, S_x and V_x are square matrices with dimensions of $n \times n$. V_x is unitary and S_x is a diagonal matrix with diagonal elements giving the positive sorted PC singular values, values that give the individual standard deviations of the corresponding EOFs. These properties hold for U_y, S_y and V_y generated from PCA computed on Y.

Scaling and scaled covariance in terms of EOFs and PCs

After computing PCA with X and Y with $X=U_xS_xV_x^T$ and $Y=U_yS_yV_y^T$, respectively, the individual covariance for X and Y may be expressed in terms of EOFs as $XX^T=U_xS_x^2U_x^T$ and $YY^T=U_yS_y^2U_y^T$, respectively. From this it is clear that the whitening transformation

simplifies to $(XX^T)^{-1/2}=U_xS_x^{-1}U_x^T$, and similarly for Y. It follows that whitening simply acts to weight EOFs equally (all weighted by one) by replacing S_x with the identity matrix, with $U_xS_x^{-1}U_x^TX = U_xV_x^T$. In terms of the PC singular values, scaling transformations of Eqs. 2.2 and 2.3 may be written similarly as the whitening transformation as follows:

$$T_x = U_xS_x^{\alpha-1}U_x^T \quad \text{(A1)}$$

$$T_Y = U_yS_y^{\beta-1}U_y^T \quad \text{(A2)}$$

Applying the transformation reveals that scaling preserve EOF patterns and variates acting to apply only PC singular values to the powers of α and β yielding $X^*=T_xX=U_xS_x^{\alpha}V_x^T$ and $Y^*=T_yY=U_yS_y^{\beta}V_y^T$, respectively. Despite the exponential form, scaling is linear for fixed values of α and β. Also, as α and/or β are increased from zero, the disparity in PC singular values of the scaled variables (S_x^{α} and/or S_y^{β}) is increased from zero. The transformed weighted PCs are recovered from projecting the scaled variables onto the unweighted orthogonal EOF patterns yielding $X^{*T}U_x = V_xS_x^{\alpha}$ and $Y^{*T}U_y = V_yS_y^{\beta}$. Then the cross-covariance of the scaled data (2.4) may be represented in PC space in terms of PC singular values and PC cross-correlation $V_x^TV_y$, given as follows:

$$U_x^TX^*Y^{*T}U_y = S_x^{\alpha}(V_x^TV_y)S_y^{\beta} \quad \text{(A3)}$$

Scaled SVD analysis in PC-space

In PC space, covariance may be represented in terms of only the PCs with $U_x^TXY^TU_y = S_xF_x^TF_yS_y$. It follows that the SSVD constrained optimization problem (3.9 & 3.10) may be expressed in PC space as follows:

$$max_{q_x,q_y}\ (q_x^T(S_xV_x^TV_yS_y)q_y) \quad \text{(A4)}$$

$$\text{subject to}\ \ q_x^TS_x^{2(1-\alpha)}q_x = 1\ \text{ and }\ q_y^TS_y^{2(1-\beta)}q_y = 1 \quad \text{(A5)}$$

If one substitutes appropriate values for α and β, it is clear that the only distinction between CCA-1 (α=β=0), MCA-1 (α=β=1) and RDA-1 (α=0, β=1) is the role of

amplitude in the constraints (A5) represented as the particular inclusion or exclusion of S_x and/or S_y. In order to satisfy Eq. A5, it is clear that q_x and q_y must take on the form of ${S_x}^{\alpha-1}$ and ${S_y}^{\beta-1}$ multiplied by normalized vectors. Combined with Eq. A4, such vectors must provide linear combinations of PCs that maximize scaled covariance (A3), and thus are found by decomposing scaled covariance with SVD as follows:

$$ {S_x}^{\alpha}\left({V_x}^{T}V_y\right){S_y}^{\beta} = U'S'V'^{T} \tag{A6} $$

From Eq. A6, unitary singular vectors U' and V' provide linear combinations of the scaled PCs corresponding to individual SSVD modes that consecutively maximize squared scaled covariance given by diagonal elements of S'. SSVD-1 solutions may then be given as $q_x = {S_x}^{\alpha-1}U_1'$ and $q_y = {S_y}^{\beta-1}V_1'$ which yield the highest possible scaled covariance of $S'_{1,1}$.

Computing SVD in PC space is generally highly desirable if spatial dimensions greatly exceed the number of time samples. This is because the computational time of computing SVD of the $n \times n$ matrix ${U_x}^{T}X^{*}Y^{*T}U_y$, in addition to the time taken for computing individual PCA, is typically *much* less than the time taken for computing SVD on $X^{*}Y^{*T}$ itself (3.11), an $m_x \times m_y$ matrix.

The physical patterns for SSVD-1 are recovered by transforming solutions of Eqs. A4 & A5 (q_x and q_y) back to physical space, given by U_xq_x and U_yq_y. For the ordered set of SSVD patterns (P_x and P_y of Eqs. 3.12 & 3.14, respectively), this is consistent with inverse-scaling U' and V' after transforming to physical space, given that $U=U_xU'$ and $V=U_yV'$. The associated SSVD variates result directly from taking linear combinations U' and V' of the scaled PCs (A8 & A10), and SSVD basis sets may be expressed in terms of EOFs and PCs as follows:

$$ P_x = U_x{S_x}^{1-\alpha}U' \tag{A7} $$

$$ R_x = V_x{S_x}^{\alpha}U' \tag{A8} $$

$$ P_y = U_y{S_y}^{1-\beta}V' \tag{A9} $$

$$ R_y = V_y{S_y}^{\beta}V' \tag{A10} $$

For *X*, the orthogonality constraints of Eqs. 3.16 & 3.17 may be simplified in terms of EOFs and PCs as follows:

$$P_x{}^T P_x = U'^T S_x{}^{2(1-\alpha)} U' = I \quad iff \quad \alpha = 1 \tag{A11}$$

$$R_x{}^T R_x = U'^T S_x{}^{2\alpha} U' = I \quad iff \quad \alpha = 0 \tag{A12}$$

A variant of SSVD in PC-space

An alternative CCA-like form of SSVD is discussed briefly in Section 3.3. In terms of EOFs and PCs, the associated basis sets of this variant of SSVD are given more simply by taking linear combinations U' and V' of the fully-weighted EOFs and un-weighted (whitened) PCs. For *X*, these basis sets are given by $P'_x = U_x S_x U'$ and $R'_x = V_x U'$, respectively, and similarly for the EOFs and PCs of *Y* using V'. Considering Eq. A7, the modified patterns are related to the SSVD patterns through a transformation that disproportionately inflates EOFs (provided $\alpha \neq 0$), given for *X* as follows:

$$P'_x = U_x S_x{}^{\alpha} U_x{}^T P_x \tag{A13}$$

By shifting more amplitude to leading EOFs, the net effect is to redirect pattern vectors towards the leading EOFs. Again, given that variates are represented as a product of two unitary matrices ($R'_x = V_x U'$ and $R'_y = V_y V'$), they are uncorrelated with $R'_x{}^T R'_x = I$ and $R'_y{}^T R'_y = I$.

APPENDIX B

Scaled Covariance as linear regression coefficient

From univariate linear regression, with x as the predictor and y as the predictand, the least squares solution $\hat{y}$ is given by $\hat{y}=mx$ with the regression coefficient m given by $m=\langle xy\rangle/\langle x^2\rangle=\sigma_x^{-1}\rho_{xy}\sigma_y$. T2008 demonstrate that different linear transformations may be taken to yield a modified regression coefficient in such a way as to represent correlation, covariance and square root of explained variance. Analogously but in a different form, linear transformations may be represented in terms α and β such that the regression coefficient yields scaled covariance (1), with the modified solution $\hat{y}'=m'x'$. These transformed variables are given by $x'=x^*/\langle x^{*2}\rangle=\sigma_x^{-\alpha-1}x$ and $y'=y^*=\sigma_y^{\beta-1}y$.

In multiple dimensions, the sample scaled covariance (2.4 & A3) may also be formed as a matrix of regression coefficients between transformed data. The least squares solution between predictor $\boldsymbol{x}$ and predictand $\boldsymbol{y}$ (time-varying column vectors of X and Y, respectively) is given by $\hat{y}=M^Tx$, with the matrix of regression coefficients M given by $M=(XX^T)^{-1}XY^T$. A transpose is taken in this form as to reverse the order of X and Y in M so that it has clearer correspondence to the univariate form given by $m=\sigma_x^{-2}\langle xy\rangle=\sigma_x^{-1}\rho_{xy}\sigma_y$. In terms of EOFs and PCs, M is given by $M = U_xS_x^{-1}V_x^TV_yS_yU_y^T$, with the cross-correlation $V_x^TV_y$ and standard deviations given by diagonal elements of S_x and S_y. Transformed variables X' and Y' may then be formed to yield a modified solution $\hat{y}'=M'^Tx'$ such that regression coefficient M' produces scaled covariance (2.4 & A3) given by $M' = (X'X'^T)^{-1}X'Y'^T = X^*Y^{*T}$. In terms of α and β, such transformed variables are given as follows:

$$X' = \left(X^*X^{*T}\right)^{-1}X^* = (XX^T)^{(-\alpha-1)/2}X = \left(U_xS_x{}^{-\alpha-1}U_x{}^T\right)U_xS_xV_x{}^T = U_xS_x{}^{-\alpha}V_x{}^T \quad \text{(B1)}$$

$$Y' = Y^* = (YY^T)^{(\beta-1)/2}Y = \left(U_yS_y{}^{\beta-1}U_y{}^T\right)U_yS_yV_y{}^T = U_yS_y{}^{\beta}V_y{}^T \quad \text{(B2)}$$

APPENDIX C

Synthetic datasets

The population EOF patterns (E) for both X and Y (both of spatial dimension m=60) for all experiments are given as follows:

$$E_{i1} = \sin\left(\pi \frac{i-1}{m}\right)\sqrt{\frac{2}{m}}S_{11} \quad \text{(C1)}$$

$$E_{i2} = \cos\left(\pi \frac{i-1}{m}\right)\sqrt{\frac{2}{m}}S_{22} \quad \text{(C2)}$$

$$E_{i3} = \sin\left(3\pi \frac{i-1}{m}\right)\sqrt{\frac{2}{m}}S_{33} \quad \text{(C3)}$$

$$E_{i4} = \cos\left(3\pi \frac{i-1}{m}\right)\sqrt{\frac{2}{m}}S_{44} \quad \text{(C4)}$$

The associated variances are given by diagonal elements of $E^TE = S^2$, a diagonal 4×4 matrix.

The associated variates, or columns of F_x and F_y, are generated such that specified linear combinations U^{signal} and V^{signa} yield coupled signal variates R_x^{signal} and R_x^{signal} that are expected to have a specified cross-correlation matrix $\rho^{X,Y}$:

$$\langle U^{signal^T} {F_x}^T F_y V^{signal}\rangle = \langle R_x^{signal^T} R_y^{signal}\rangle = \rho^{X,Y} \quad \text{(C5)}$$

For all experiments, diagonal elements of $\rho^{X,Y}$ are specified as $\rho_{ii}^{X,Y} = \begin{cases} 0.9, & \text{if } i = 1 \\ 0, & \text{if } i > 1 \end{cases}$, such that only one true coupled signal is present with a correlation of 0.9. Coupled signal variates are generated randomly by first computing a Cholesky decomposition of the 8×8 correlation matrix, given by $\rho = \begin{bmatrix} I & \rho^{X,Y} \\ \rho^{Y,X} & I \end{bmatrix} = U_\rho U_\rho^T$, and followed by multiplying a randomly-generated n×8 white noise matrix W by U_ρ^T. With the white noise such that $\langle W^TW\rangle = I$ and $\langle U_\rho W^T W U_\rho{}^T\rangle = U_\rho U_\rho{}^T = \rho$, random coupled signal variates satisfying Eq. C5 may be generated from the columns of $WU_\rho{}^T = [R_x^{signal} \;\; R_y^{signal}]$. Given unitary matrices U^{signal} and V^{signal}, PCs are produced from

$F_x = R_x^{signal}\ U^{signal^T}$ and $F_y = R_Y^{signal}\ V^{signal^T}$, with $<F_x^T F_x> = I$ and $<F_y^T F_y> = I$.

The linear combinations U^{signal} and V^{signal} are also random with exception to the weight of population EOF-1 which is specified as $U_{11}^{signal} = \sqrt{a}$ and $V_{11}^{signal} = \sqrt{b}$ with $0 \leq a \leq 1$ and $0 \leq b \leq 1$, respectively, with the fractional amplitude weighting of population EOF-1 for X and Y given as a and b, respectively. Note that the a used here should be distinguished from the univariate regression coefficient presented earlier.

With signal patterns given by $P_x^{signal} = EU^{signal}$ and $P_y^{signal} = EV^{signal}$, again note that only the first pattern vector contributes to the time-varying signal, given by $X^{signal} = P_{x,1}^{signal}\ R_{x,1}^{signal^T}$ and $Y^{signal} = P_{y,1}^{signal}\ R_{y,1}^{signal^T}$, with the subscript 1 referring to the first column. In the absence of background noise, this is perfectly represented by CCA-1. The coupled signal is also alternatively represented as MCA-1, given by $\widetilde{X}^{signal} = ES^{-1}U_{MCA,1}(F_x S U_{MCA,1})^T$ and $\widetilde{Y}^{signal} = ES^{-1}V_{MCA,1}(F_y S V_{MCA,1})^T$, using the leading MCA singular vectors $U_{MCA,1}$ and $V_{MCA,1}$.

Given a and b, the expectations of the true signal variances depend only the variance of population EOF-1 (S_{11}^2), which is weighted by a and b, and the sum of variance of the remaining population EOFs, which is weighted by $1-a$ and $1-b$, respectively. For all experiments, $S_{11}^2 = S_{22}^2 + S_{33}^2 + S_{44}^2 = 0.25$, and the total expected variance is specified as one such that variance is equivalent to the fraction of explained variance. The expected signal variance may then be given as follows:

$$\langle \left\| X^{signal} \right\|_F^2 \rangle = \langle P_{x,1}^{signal^T} P_{x,1}^{signal} \rangle = 0.25(2a+1)/3 \qquad \text{(C6)}$$

$$\langle \left\| Y^{signal} \right\|_F^2 \rangle = \langle P_{y,1}^{signal^T} P_{y,1}^{signal} \rangle = 0.25(2b+1)/3 \qquad \text{(C7)}$$

Increasing a or b from 0 to 1 increases the expected true signal variances from $0.25/3 \approx 0.083$ to 0.25 corresponding to a signal-to-noise ratio of roughly $\sqrt{0.083/(1-0.083)} \approx 0.3$ to $\sqrt{0.25/(1-0.25)} \approx 0.6$, respectively. Also, the disparity of the individual variances follows a function $S_{ii}^2 = 0.25\tau^{i-1}$ with a value of $\tau \approx 0.544$ chosen to satisfy the constraint of a sum of 0.5.

The time-varying population EOFs are embedded in randomly-generated background noise N with spatial (temporal) variation of red (white) noise and with expected

variance specified to be $1-tr(S^2)$ = 0.5 in order to maintain a total expected variance of one. Similarly as Bretherton et al. (1992), red noise is generated by a Markov process as follows:

$$N_i = rN_{i-1} + \sqrt{1-r^2}W_i \quad \text{(C8)}$$

Note in this case, individual subscripts of Eq. C8 represent rows. The *ith* row of *N*, N_i, gives the *ith* spatial point varying in time and is equivalent to a weighted average between values at the neighboring spatial point N_{i-1} and random white noise W_i with population unit variance, again with $<W^TW>$ = *I*, but with *W* here having dimensions of space by time (*m*×*n*). The neighboring grid point weighting is given by the autocorrelation *r* which is specified in all experiments as *r*=0.9. By design, the unweighted added noise is expected to have unit variance with $<N^TN>$ = *I*, and the noise generated for *X* (N_x) is independent of that for *Y* (N_y). Finally, the total randomly-generated datasets are given as follows:

$$X = EF_x^T + \sqrt{0.5}N_x \quad \text{(C9)}$$

$$Y = EF_y^T + \sqrt{0.5}N_y \quad \text{(C10)}$$

As the red noise randomly projects onto the population EOFs, keep in mind that the added noise $\sqrt{0.5}\,N_x$ generally has nonzero covariance with EF_x^T, as is true for the noise of *Y*.

REFERENCES

Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, **4**, 1147-1167.

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007.

Barnett, T.P and R. Preisendorfer, 1987: Origins of monthly and seasonal forecast skill for United States Surface Air temperatures determined by Canonical Correlation Analysis. *Mon. Wea. Rev.,* **115**, 1825-1850.

Barnston, A. G., and C. F. Ropelewski, 1992: Prediction of ENSO episodes using Canonical Correlation Analysis. *J. Clim.*, **5**, 1316-1345.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Clim.*, **5**, 541-560.

Cherry, S., 1996: Singular Value Decomposition Analysis and Canonical Correlation Analysis. *J. Clim.*, **9**, 2003-2009.

DelSole, T., J. Shukla, 2006: Specification of wintertime North American surface temperature. *J. Clim.*, **19**, 2691-2716.

DelSole, T. and M. K. Tippett, 2007: Predictability: recent insights from information theory. *Rev. Geophys.,* **45**, RG4002, doi:10.1029/2006RG000202.

Donohue, K. D., A. Agrinsoni, and J. Hannemann, 2007: Audio signal delay estimation using partial whitening. *SoutheastCon, 2007. proceedings*, Richmond, VA, March 22-25, 2007.

Hastie, T., A. Buja, and R. Tibshirani, 1995: Penalized discriminant analysis. *Ann. Stat.,* **23**, 73-102.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631-1643.

Kao, H.-Y., and J.-Y. Yu, 2009: Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J. Clim.*, **22**, 615-632.

Kug, J. S., F. F. Jin, and S.-I. An, 2009: Two types of El Niño events: cold tongue El Niño and warm pool El Niño. *J. Clim.*, **22**, 1499-1515.

Li, S., M. P. Hoerling, S. Peng, and K. M. Weickmann, 2006: The annular response to tropical Pacific SST forcing. *J. Clim.,* **19**, 1802-1819.

Mitchell, T. D., T. R. Carter, P. Jones, and M. Hulme, 2003: A comprehensive set of climate scenarios for Europe and the globe. *Tyndall Centre Working Paper,* **55**.

Newman, M., and P. D. Sardeshmukh, 1995: A caveat concerning singular value decomposition. *J. Clim.,* 8, 352-360.

Swenson, E.T., 2013: Scaled SVD Analysis and its application to tropical-extratropical teleconnections. *In preparation.*

Tippett, M. K., T. DelSole, S. J. Mason, A. G. Barnston, 2008: Regression-based methods for finding coupled patterns. *J. Clim.*, **21**, 4384-4398.

Tung, Y. L., C.-Y. Tam, S.-J. Sohn, and J.L. Chu, 2013: Improving the seasonal forecast for summertime South China rainfall using statistical downscaling. *J. Geophys. Res.*, **118**, 5 147-5159.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research.* Cambridge University Press, 494 pp.

Wallace, J. M., C. Smith, and C. S. Bretherton, 1992: Singular Value Decomposition of wintertime sea surface temperature and 500-mb height anomalies. *J. Clim.*, **5**, 561-576.

Wilks, D. S., 2005: Statistical Methods in the Atmospheric Sciences, 2nd ed., *Academic*, San Diego, Calif.

Wold, H. 1966: Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, edited by Krishnaiah, P. R., pp. 391-420, Academic, New York.

Wu, Q., 2010: Forcing of tropical SST anomalies by wintertime AO-like variability. *J. Clim.*, **23**, 2465-2472.

Xie, P., and P.A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539 - 2558.

Yin, X., A. Gruber, and P. Arkin, 2004: Comparison of the GPCP and CMAP merged gauge-satellite monthly precipitation products for the period 1979-2001. *J. Hydrometeorol.*, **5**, 1207-1222.

APCC RESEARCH REPORT 2013-02

- Construction of BSISO Forecast System and Application to Summer Monsoon Prediction
- Revision of Climate Change by Dynamic Downscaling over the Maritime Continents
- Long-Lead Station-Scaled Prediction of a Hydrological Drought in South Korea Based on Bivariate Downscaling
- Development of Scaled SVD Analysis and Related Methods with Focus on Application to Tropical-Extratropical Teleconnections

APEC Climate Center

12, Centum 7-ro, Haeundae-gu, Busan 612-020,
Republic of Korea
Tel: +82-51-745-3900 Fax: +82-51-745-3949
www.apcc21.org

비매품
94550
9 788997 333943
ISBN 978-89-97333-94-3
ISBN 978-89-97333-92-9 (세트)