RESEARCH REPORT 2015-04

APEC CLIMATE CENTER

Continuum Power CCA and Its Application to Statistical Downscaling

Erik T. Swenson
Climate Prediction Team

APCC
APEC CLIMATE CENTER

APEC CLIMATE CENTER
RESEARCH REPORT

Continuum Power CCA and Its Application to Statistical Downscaling

Erik T. Swenson
Climate Prediction Team

RESEARCH REPORT 2015-04

APCC
APEC CLIMATE CENTER

Preface

The development of multivariate statistical methods for isolating coupled relationships is useful for a range of climate applications. Applications include model evaluation, diagnosis of physical coupled processes (e.g. teleconnections), paleoclimate reconstruction, and in particular seasonal climate prediction. Regarding seasonal prediction, such approaches may be applied to statistical downscaling towards improving dynamical forecasts using empirical relationships derived from the observational record. Among the various existing methods, approaches using regularization to better estimate sample covariance matrices show promise in improving skill but have yet to be fully explored or applied to seasonal prediction.

As an operational climate prediction center, APCC is very interested in providing the most accurate and relevant seasonal forecast possible given the information available. Over the past decade, the multi-model approach has proven quite useful however further processing may be done to better incorporate observations, reduce common model bias, and improve skill at the regional scale. Along these lines, the development of new approaches for statistical downscaling is one key direction towards further improving the MME forecast. Thus, researchers at APCC explore various downscaling approaches.

In this study, a new regularization approach for isolating coupled relationships is introduced and applied to statistical downscaling in order to improve seasonal prediction of temperature and precipitation for various seasons and regions around the world. Results are contrasted with those resulting from using comparable pre-existing approaches for a range of cases in which the skill of the MME forecast is improved. In some cases, further improvement is found using the new approach.

The study not only works towards improving seasonal prediction in general but also provides specific guidance for when and where statistical downscaling is expected to improve dynamical prediction in a fairly objective framework. At APCC, the downscaling tool may be used to directly improve operation prediction at the regional scale, which has a clear socioeconomic benefit, particularly related to agricultural planning. Regarding this study, I acknowledge the dedicated work of Erik Swenson and the support from colleagues and researchers at APCC.

Dr. Chin-Seung Chung
Director / APEC Climate Center
March 2015

ABSTRACT

Various multivariate statistical methods exist for analyzing covariance and isolating linear relationships between datasets, most of which use Singular Value Decomposition (SVD) to maximize some quantity based on cross-covariance. In this study, Continuum Power CCA (CPCCA) is introduced as an extension of continuum power regression for isolating pairs of coupled patterns whose temporal variation maximizes the squared covariance between partially-whitened variables. Similar to the whitening transformation, the partial whitening transformation acts to decorrelate individual variables, but only to a partial degree with the added benefit of pre-conditioning sample covariance matrices prior to inversion, providing a more accurate estimate of the population covariance. CPCCA is a unified approach in the sense that the full range of solutions bridges Canonical Correlation Analysis (CCA), Maximum Covariance Analysis (MCA), Redundancy Analysis (RDA) as well as Principal Component Regression (PCR). Recommended CPCCA solutions include a regularization for CCA, a variance bias correction for MCA, and a regularization for RDA. An objective parameter choice is offered for regularization based on the covariance estimate of Ledoit and Wolf (2004). Following this, CPCCA is applied to statistical downscaling of seasonal surface temperature and precipitation based on sea surface temperature and sea level pressure predicted by the APCC MME hindcast for various seasons and regions. Along with other pattern-based downscaling approaches, CPCCA improves the skill over MME output and tends to yield higher skill than Regularized CCA. However, the use of prior EOF truncation (Barnett and Preisendorfer 1987) for CCA tends to yield higher skill than either form of regularization.

Contents

Continuum Power CCA and Its Application to Statistical Downscaling

1. INTRODUCTION

Various multivariate statistical methods have been developed and proven useful for analyzing covariance and estimating physically meaningful linear relationships between multiple coupled variables. Such approaches have many applications spanning a broad range of disciplines, although they are particularly useful for diagnosis and prediction of the physical climate system. Most linear methods use Singular Value Decomposition (SVD) to identify a related subspace with projection vectors that maximize some measure of relation. Two popular and well-understood symmetric methods are Canonical Correlation Analysis (CCA, Hotelling 1936) and Maximum Covariance Analysis (MCA, Wallace et al. 1992, Bretherton et al. 1992) that generate pairs of stationary spatial patterns whose temporal co-variation produces the highest possible squared cross-correlation and squared cross-covariance, respectively. In practice, CCA has been used to diagnose and predict land air temperature based on co-varying sea surface temperature (SST) patterns (Barnett and Preisendorfer 1987, DelSole and Shukla 2006, Barnston and Smith 1996). Similarly, MCA has been used to infer teleconnections with SST (Wallace et al. 1992 Czaja and Frankignoul 2002, Frankignoul and Sennechael 2007, Wu 2010) and for statistical downscaling (Tippett et al. 2008, Tung et al. 2013).

Alternatively, asymmetric techniques may be useful and include Redundancy Analysis (RDA, vonStorch and Zwiers 1999) and Partial Least Squares (PLS) regression (Wold 1966) which maximize the explained variance of one variable by another. These approaches have also been applied for diagnosis and prediction (e.g., Bakalian et al. 2010 and Smoliak et al. 2010, respectively).

CCA is generally the most flexible and theoretically attractive approach, however it is strongly affected by overfitting when applied to small samples that have a high degree of collinearity, typical characteristics of geophysical variables of the observed climate system. Consequently, CCA requires some form of predictor reduction or regularization in practice. Techniques that achieve this include Regularized CCA (RCCA, Vinod 1976) and Penalized CCA (Hastie et al. 1995), both of which use ridge regression to apply spatial smoothing, as well as Kernel CCA (Akaho 2001) and CCA using truncation according to Empirical Orthogonal Functions (EOFs) (hereafter BP-CCA, Barnett and Preisendorfer

1987). Unfortunately solutions of BP-CCA are often highly sensitive to the number of EOFs retained, which should be chosen objectively. Analogously, RCCA also requires objectively choosing a ridge parameter. Cross-validation is an effective means of justifying EOF truncation or a ridge parameter, however it may not be possible with small samples given that sample data generally cannot be used for both choosing the parameter and applying the approach. For BP-CCA, measures such as Akaike's information criterion (AIC, Akaike 1973, Hurvich and Tsai 1989) may be computed. An objective and optimal choice for the ridge parameter with respect to covariance estimation has also been developed (Ledoit and Wolf 2004, hereafter LW) for which application to RCCA has yielded robust results (Cruz-Cano and Lee 2014).

Regularization techniques are attractive because they can handle missing data and often yield more robust results than conventional techniques, particularly when the number of predictors (spatial points) exceeds the sample size. This is common in climatic data, and multiple climate studies have benefited from using regularization. To name a few, for paleoclimate reconstruction, Smerdon et al. (2010) evaluate BP-CCA alongside Regularized Expectation Maximization (Schneider 2001), and Tingley et al. (2012) summarize various approaches and the important issues involved. For climate change detection, Ribes et al. (2009) apply ridge regression to the optimal fingerprint method using the LW sample covariance estimate. For climate prediction, Lim et al. (2011) improve seasonal rainfall prediction using RCCA. Also, Fischer (2014) regularizes Principal Covariates Regression (deJong and Kiers 1992) yielding a continuum approach that links Principal Component Regression (PCR) with RDA and PLS regression, and further applies it to identify coupled patterns between seasonal rainfall and sea-level pressure.

An alternative regularization technique that has yet to be applied in this context is continuum power regression (Lorber et al. 1987, Stone and Brooks 1990), an approach that bridges ordinary least squares, PLS regression and PCR. Partial whitening, the transformation involved with continuum power regression, acts to decorrelate differing spatial points, but only to a partial degree specified by a parameter. Here it is applied as a regularization for sample covariance matrices prior to computing CCA, yielding a new approach for linearly relating variables that may be referred to as Continuum Power CCA (CPCCA). CPCCA maximizes the squared cross-covariance between partially-whitened

variables and provides a unique plane of solutions that bridges those of CCA, MCA and RDA, as well as PCR in its upper limit. CPCCA solutions include a regularization for CCA and RDA, as well as a variance bias correction for MCA.

This report is organized as follows. Section 2 presents a generalized SVD procedure for which fixed solutions include CCA, MCA, RDA as well as CCA regularization approaches. CPCCA is discussed alongside RCCA, and objective choices for parameter values are offered. Following this, section 3 presents the methodology for application to statistical downscaling along with a data description. Section 4 then presents the downscaling results for which CPCCA and related approaches are able to improve seasonal prediction of temperature and precipitation over that simulated by dynamical models. Lastly, section 5 provides a summary and conclusions.

2. CONTINUUM POWER CCA

In this section, Continuum Power CCA (hereafter CPCCA) is introduced as a new unified regularization approach for linearly relating two variables by isolating sets of multivariate relationships. First, a generalized procedure is presented for maximizing the cross-covariance between transformed variables. In this framework, the traditional methods of CCA, MCA and RDA are shown to be particular fixed cases. Then, CPCCA is derived alongside RCCA as a unique regularization technique that bridges traditional approaches by varying a set of parameters. Lastly, objective parameter choices are offered including one based on the sample covariance estimate of Ledoit and Wolf (2004).

2.1 A generalized SVD procedure

Provided with variables X and Y, matrices of anomalies (zero sample mean for each spatial point) with rows giving different spatial points and columns giving different samples (e.g., temporal variation), assume that spatial points have equal weighting (e.g., uniformly spaced grids) and have absorbed a factor of $1/\sqrt{\boldsymbol{n}\text{-}1}$, with $\boldsymbol{n}$ giving the number of samples. Then, individual covariances for X and Y are given simply by XX^T and YY^T, respectively, and the cross-covariance is given by XY^T.

In order to identify dominant linear relationships between X and Y, one may seek fixed pattern vectors that maximize cross-covariance while being bounded by a set of norms. Such pattern vectors results as solutions to the optimization problem

$$max_{\mathbf{q}_x,\mathbf{q}_y} \left(\mathbf{q}_x{}^T XY^T \mathbf{q}_y\right) \tag{2.1}$$

subject to $$\mathbf{q}_x{}^T T_x{}^{-2} \mathbf{q}_x = 1 \quad \text{and} \quad \mathbf{q}_y{}^T T_y{}^{-2} \mathbf{q}_y = 1, \tag{2.2}$$

where T_x and T_y may represent any set of symmetric positive definite linear transformations for X and Y, respectively. After defining transformed variables $X^*=T_xX$ and $Y^*=T_yY$ along with inverse-transformed solutions $q_x^* = T_x{}^{-1} q_x$ and $q_y^* = T_y{}^{-1} q_y$, the optimization problem of Eqs. (2.1) and (2.2) may be re-expressed as

$$max_{\mathbf{q}_x^*,\mathbf{q}_y^*} \left(\mathbf{q}_x^{*T} X^* Y^{*T} \mathbf{q}_y^*\right) \tag{2.3}$$

subject to $$\mathbf{q}_x^{*T} \mathbf{q}_x^* = 1 \text{ and } \mathbf{q}_y^{*T} \mathbf{q}_y^* = 1. \tag{2.4}$$

In the above form, solutions q_x^* and q_y^* must yield the highest squared cross-covariance between transformed variables while being bounded to have unit variance. They may be found by computing an SVD of $X^* Y^{*T}$, decomposing it in terms of orthonormal singular vectors (columns of $\boldsymbol{U}$ and $\boldsymbol{V}$) and positive, sorted singular values (diagonal elements of S), with

$$X^* Y^{*T} = T_x XY^T T_y{}^T = USV^T. \tag{2.5}$$

Solutions are then given by the leading singular vectors with $q_x^* = u_1$ and $q_y^* = v_1$ (first column of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively), which necessarily satisfy Eqs. (2.3) and (2.4), with cross- covariance given by the leading singular value $s_1 = u_1{}^T X^* Y^{*T} v_1$ (first diagonal element of S). As singular vectors are orthonormal, it may be easily shown that the total squared cross- covariance is given by $\|X^* Y^{*T}\|_F^2 = \|S\|_F^2 = s_1^2 + \cdots + s_n{}^2$ (with F indicating the Frobenius norm such that its square yields the sum of the squares of diagonal elements). Apart from what is explained by the leading mode, the remaining squared cross-covariance, given by $\|S\|_F^2 - s_1^2$, is maximized by the second leading singular vectors u_2 and v_2 subject to the orthogonality constraints $u_1{}^T u_2 = 0$ and $v_1{}^T v_2 = 0$. Further, the remaining $\|S\|_F^2 - s_1^2 - s_2^2$ is maximized by the third leading singular vectors that are orthogonal to the first two and so on.

Singular vectors U and V may be used to construct full basis sets for the original untransformed variables X and Y that consecutively maximize cross-covariance between the transformed variables. Such pattern vectors result from inverse-transforming the singular vectors, and are given by the columns of

$$P_x = T_x{}^{-1}U \qquad \text{and} \qquad P_y = T_y{}^{-1}V, \tag{2.6}$$

with the associated variates (or temporal coefficients) resulting from projecting X^* and Y^* onto the singular vectors, given by the columns of

$$R_x = X^T T_x U \qquad \text{and} \qquad R_y = Y^T T_y V. \tag{2.7}$$

Eqs. (2.6) and (2.7) provide a complete decomposition with $P_x R_x{}^T = T_x{}^{-1} U U^T T_x X = X$, and similarly for Y. Note that in terms of the variates alone, the cross-covariance between transformed variables may be represented by $R_x{}^T R_y = U^T X^* Y^{*T} V = S$, from which it is clear that differing variates are uncorrelated (S is diagonal).

For the weighted variates of X, normalization is typically desired so that patterns absorb variance and may be represented with the same units as X. This is done through the normalization matrix $S_x^R = (Diag\ R_x{}^T R_x)^{1/2}$, where $Diag\ B$ denotes a diagonal matrix that contains the diagonal elements of B. Then the weighted patterns and normalized variates are given by the columns of $P_x S_x{}^R$ and $R_x S_x{}^{R^{-1}}$, respectively, and similarly for Y.

2.2 Fix solutions and regularization

Basis sets that maximize various quantities based on cross-covariance may be found according to Eqs. (2.3) - (2.7), and they differ only in terms of the set of transformations that are used, T_x and T_y. One particularly useful transformation, the whitening transformation (DelSole and Tippett 2007), is given by the inverse square root of the sample covariance matrix and represents the multivariate form of normalization, given for X by $(XX^T)^{-1/2}$. Multiplication with the original variable acts to completely decorrelate (and normalize) differing spatial points (or predictors), given that the sample covariance matrix of a whitened variable is equivalent to the identity matrix, with $(XX^T)^{-1/2} XX^T (XX^T)^{-1/2} = \mathrm{I}$. This equivalently removes the disparity in EOF weighting (not shown).

Noted by Tippett et al. (2008), the cross-covariance of whitened variables equates to cross-correlation, for which Eqs. (2.3) - (2.7) yield CCA. On the other hand, transforming neither X nor Y retains cross-covariance, and yields MCA. Further, whitening only X results in the variance of Y explained by X, and yields RDA. To summarize, one may apply the transformations

$$T_x = (XX^T)^{-1/2} \quad \text{and} \quad T_y = (YY^T)^{-1/2} \quad \text{for CCA,} \tag{2.8}$$
$$T_x = I \quad \text{and} \quad T_y = I \quad \text{for MCA,} \tag{2.9}$$
$$T_x = (XX^T)^{-1/2} \quad \text{and} \quad T_y = I \quad \text{for RDA.} \tag{2.10}$$

CCA is inherently an ideal approach in how modes are uncorrelated (it may be shown that $R_x^T R_x = I$ and $R_y^T R_y = I$), thus providing a least-squares estimate as well as an optimal noise filter for autoregressive models (DelSole and Chang 2003). However, with the process of whitening, noise that is present in sample covariance matrices tends to be inflated during inversion, particularly in the presence of high collinearity (i.e. large spatial autocorrelation). Further amplified by SVD, this can lead to dramatic overfitting and strongly affect the leading modes if it by chance increases the squared correlation. Note that the sample covariance matrix also tends to overestimate covariation between differing spatial points, and some sort of shrinking or regularization can not only condition the matrix but provide a covariance estimate that is consistently closer to the population (Ledoit and Wolf 2004). This is especially necessary when the number of spatial points exceeds the sample size $\boldsymbol{n}$, in which case the sample covariance matrix is actually invertible and a Moore-Penrose pseudo-inverse may be taken instead.

Ridge regression, or Tikhonov regularization, provides an effective approach that may be applied to the whitening transformation with Regularized CCA (RCCA, Vinod 1976). It involves adding a small positive value to the diagonal of the sample covariance matrices, thereby shrinking and conditioning them prior to taking the inverse square root, given by

$$\begin{aligned} T_x &= [(1-\rho_x)XX^T + \rho_x \mu_x I]^{-1/2} \\ \text{and} \quad T_y &= \left[(1-\rho_y)YY^T + \rho_y \mu_y I\right]^{-1/2} \end{aligned} \quad \text{for RCCA,} \tag{2.11}$$

with constants $\mu_x = \|X\|_F^2 / m_x$ and $\mu_y = \|Y\|_F^2 / m_y$ such that total variance is maintained. An alternative approach that has yet to applied to CCA is continuum power regression (Lorber

et al. 1987, Stone and Brooks 1990), which is analogous to ridge regression but has an exponential form (rather than linear). Applying this to whitening results in a transformation that may be referred to as partial whitening, given by

$$T_x = (XX^T)^{(\alpha-1)/2} \quad \text{and} \quad T_y = (YY^T)^{(\beta-1)/2} \quad \text{for CPCCA,} \tag{2.12}$$

This operation acts to only partially decorrelate and normalize the data to a specified degree. Note that parameter notation differs from that of Stone and Brooks (1990), and parameters may be re-expressed as any positive function. For X, partial whitening replaces the sample covariance matrix XX^T with $(XX^T)^{1-\alpha}$, which can often be closer to the population covariance (depending on the value of α). Also, supposing that the sample covariance matrix has a condition number of K, $(XX^T)^{1-\alpha}$ has a condition number of $K^{|1-\alpha|}$ which is always lower for $0 < \alpha < 2$ assuming that $K > 1$. Applying this to Eqs. (2.3) - (2.7) yields a new approach, introduced here as Continuum Power CCA (CPCCA). For X, inverting $(XX^T)^{(1-\alpha)/2}$ in Eq. 2.12 is not always possible, and it is always invertible when the number of spatial points exceeds the sample size. As noted prior, in this case the inverse is replaced with the Moore-Penrose pseudo-inverse (as done for statistical downscaling results presented in the next section).

Both RCCA and CPCCA continuously bridge the fixed solutions of CCA, MCA and RDA. This is evident for RCCA by varying parameters ρ_x and ρ_y in Eq. (2.11), as well as for CPCCA by varying α and ß in Eq. (2.12). These transformations equate to the whitening for parameter values of zero and to the identity matrix (within a constant) for parameter values of one, with fractional values corresponding to regularization. From the point of view of CCA, one may then consider MCA ($\rho_x = \rho_y = 1$ and $\alpha = ß = 1$) to be an extreme regularization. MCA basically replaces the sample covariance with the perfectly conditioned identity matrix, conveniently avoiding inversion altogether. This can effectively stabilize solutions, however it also constrains relationships to be represented by orthogonal transformations (Newman and Sardeshmukh 1995), which may be unrealistic. RCCA and CPCCA simply offer different paths between solutions, either of which may be preferable under different conditions.

The full range of CPCCA solutions may be better understood in terms of the measure of relation that is maximized, the squared cross-covariance between partially-whitened

variables. This quantity is essentially the multivariate form of the univariate measure

$$VAR_x{}^{\alpha} \times COR^2 \times VAR_y{}^{\beta} \tag{2.13}$$

where $\boldsymbol{VAR_x}$ and $\boldsymbol{VAR_y}$ are variances corresponding to $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, and $\boldsymbol{COR}$ is cross-correlation. Given that Eq. (2.13) recovers cross-correlation, cross-covariance and the variance of one variable explained by the other, CPCCA conveniently bridges CCA (α = ß = 0), MCA (α = ß = 1) and RDA (α = 0 and ß = 1, or vice-versa), respectively. Additionally, apart from special circumstances, CPCCA converges to PCR in its upper limit (α→∞ or ß→∞) yielding the EOFs of individual variables (not shown). One way this may be seen is that Eq. (2.13) is completely dominated by $\boldsymbol{VAR_x}^{\alpha}$ or $\boldsymbol{VAR_y}^{ß}$ as either α or ß get large, respectively. Given the exponential form, CPCCA solutions tend to resemble PCR for parameter values not much greater than one, particularly if the value of the other parameter is much less than one.

Given computational constraints, CPCCA and related methods are ideally computed in PC space. See Swenson (2014) for further details of CPCCA including computation in PC space.

2.3 Objective parameter choices for regularization

Without prior knowledge, an objective basis is required for choosing α and ß before applying CPCCA as a regularization. Regarding the estimation of an unknown population covariance matrix, Ledoit and Wolf (2004) elegantly derive a choice for the ridge parameter (denoted by LW) that is asymptotically optimal (*as both* the sample size n and number of spatial points m_x or m_y increase together) and objective in that it is determined from the sample data alone. For X, this is given by

$$\rho_{x.LW} = \frac{\sum_{i=1}^{n} \left\| (n-1)\mathbf{x_i}\mathbf{x_i}^T - XX^T \right\|_F^2}{\left\| XX^T - \mu_x I \right\|_F^2} . \tag{2.14}$$

This recently has been applied to RCCA by Cruz-Cano and Lee (2014). However, it is apparent that no optimal and objective parameter choice has been similarly derived for continuum power regression (and consequently CPCCA), and a solution analogous to that

of Ledoit and Wolf (2004) for the ridge parameter is either highly complex or doesn't exist. Alternatively, here a solution is chosen that estimates the sample covariance matrix closest to the LW estimate. This is found iteratively by varying α to minimize mean-squared error, given by

$$\min \ \left\| v_x (XX^T)^{1-\alpha} - \left(1 - \rho_{x,LW}\right) XX^T - \rho_{x,LW} \mu_x I \right\|_F^2, \tag{2.15}$$

with $\boldsymbol{v_x} = \|\boldsymbol{X}\|_F^2 / \|(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})^{(1-\alpha)/2}\|_F^2$ so that the total variance is maintained.

A Monte Carlo procedure is conducted involving X of dimensions $m_x = 60$ and $n = 30$ drawn randomly given a known population covariance matrix distributed as red noise, or an AR(1) process, in space (or predictors) and Gaussian white noise in time (or samples). It is found that this LW-based covariance estimate that uses partial whitening tends to be closer to the population covariance matrix than the LW estimate itself for spatial autocorrelations of 0.85 or greater (not shown). This is a significant result considering that the estimate is presumably sub-optimal, and it may be quite relevant given that such data is (to first order) quite characteristic of inter-annually-varying planetary-scale geophysical variables of the observed climate system. As for a fixed set of parameter values, application of CPCCA to similar data with an embedded coupled mode reveals that $\alpha = ß \approx 0.1$ tends to be a skillful choice for isolating the coupled mode (not shown).

Additionally, CPCCA using $\alpha = ß = 0.5$ is proposed as a variance bias correction for MCA. Given that MCA maximizes squared cross-covariance, large-magnitude and/or large-scale (spatially coherent) variability is weighted disproportionately more than weaker variability, and using parameter values of 0.5 corrects the variance contribution to the bias. After substituting these values into Eq. (2.13), this is evident in how standard deviation (square root of variance) is used.

3. DATA DESCRIPTION AND DOWNSCALING METHODOLOGY

CPCCA and related methods are applied to statistical downscaling in order to improve seasonal prediction skill of temperature and precipitation for various regions and seasons. The Model Output Statistics (MOS, Wilks 1995) approach is used in which three month averaged GCM-predicted tropical (20°S – 20°N) SST and sea level pressure (SLP) is regressed with the simultaneously observed land CRU 0.5° temperature and precipitation (Harris et al. 2012) over various regions. By design, this approach is capable of eliminating a degree of systematic model bias. The predicted SST and SLP is taken as the multi-model ensemble (MME) mean of seasonal anomalies (with respect to individual GCM climatology) predicted by 6 – 9 coupled GCMs of the APCC hindcast dataset, namely the APCC model (CAM3 coupled with POP, e=10), the JMA model (e=5), the MSC CanCM3 and CanCM4 models (e=10 for each), the NASA GMAO GEOS-5 model (e=9), the NCEP CFS (e=10), the BOM POAMA model (e=33), the PNU model (e=20) and the SUT1 model (e=6), with e giving the ensemble size. Information for individual models may be found at http://www.apcc21.org/eng/html/hapcc030001.html. The GCMs are chosen based on the fact that they span a common 26-year period that is used for the study, 1983-2008. Other models are neglected because their hindcast periods span significantly less years (e.g., ending in 2005). Note also that one model is neglected based on low skill in SST and SLP (not shown). The seasons examined are December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON), and the specific models used for each season are shown in Table 1. Hindcast simulations are initialized at a one-month lead (on the 1st of the previous month).

Table 1. GCMs used for each season.

Season	number of models	models
DJF	7	APCC, MSC_CanCM3, MSC_CanCM4, NASA, NCEP, POAMA, PNU
MAM	9	APCC, JMA, MSC_CanCM3, MSC_CanCM4, NASA, NCEP, POAMA, PNU, SUT1
JJA	8	APCC, MSC_CanCM3, MSC_CanCM4, NASA, NCEP, POAMA, PNU, SUT1
SON	6	MSC_CanCM3, MSC_CanCM4, NASA, NCEP, PNU, SUT1

The benchmark for statistical downscaling results is the MME-simulated precipitation (direct model output, DMO) 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), Principal Component Regression (PCR), Maximum Covariance Analysis (MCA), Canonical Correlation Analysis (CCA) using prior EOF truncation (Barnett and Preisendorfer 1987, BP-CCA), Regularized CCA (RCCA) using the LW covariance estimate (Ledoit and Wolf 2004, Cruz-Cano and Lee 2014), as well as Continuum Power CCA (CPCCA) using a covariance estimate closest to the LW estimate and fixed parameters of 0.1 and 0.5. All pattern-based methods are used to identify the subset of predictors and predictands through which regression is computed. The number of mode pairs used is chosen based on which gives the largest area with significant correlation. Note that for BP-CCA, the EOF truncation number is also based on the highest skill, a degree of subjectivity that could provide additional skill for BP-CCA. The skill metrics and this approach for comparing methods mimic the format of Tippett et al. (2008), and are stated at the end of this section.

For the predictor variables SST and SLP, fixed longitude domains are used in the tropics (20°S – 20°N), namely the Indian Ocean and West Pacific (30°E – 180°), the Indo-Pacific (30°E – 90°W), the Pacific (120°E – 90°W), and the Pacific-Atlantic (120°E – 0°). The

appropriate domain and variable (either SST or SLP) is based on the highest skill (average for all methods). The specific predictand regions are first identified using physical insight and further adjusted to enclose the primary large-scale region of skill resulting from the various approaches (the various approaches tended to give qualitatively similar results). The domains for predictor and predictand variables for all cases may be viewed in Fig. 1. Sensitivity to the domains is also examined to which skill is fairly insensitive. For instance, when increasing the size of the predictand domain, regions with skill tend to maintain skill despite the inclusion of surrounding areas that have no skill, which then degrades the total skill averaged over the larger domain (not shown).

This skill of the results is examined in terms of correlation and root mean squared error (RMSE) with respect to the observed temperature and precipitation, and values are cross-validated using a leave-1-out approach. Correlation is examined graphically and quantified for each case in terms of the number of grid points with a significant positive correlation, indicated by sum(corr).

4. RESULTS

4.1 December – February (DJF)

During boreal winter (DJF), the El Niño Southern Oscillation (ENSO) tends to peak in amplitude and thus dominates prediction skill, skill that originates in the tropical Pacific Ocean. Secondarily, ENSO Modoki (Ashok 2007), or flavors of ENSO (Kao and Yu 2009, Kug 2009) can have a significant effect as well. This is known to contribute to seasonal prediction skill of surface temperature globally, particularly in the mid-latitude Northern Hemisphere. Relevant statistical downscaling studies indicate prediction skill over East Asia (e.g., over Korea by Sohn et al. 2013) as well as over North America (Barnett and Preisendorfer 1987, DelSole and Shukla 2006, Barnston and Smith 1996). On the other hand, the associated prediction skill of precipitation tends to be confined to the tropics and subtropics. Consistent with past work, enhanced prediction skill using statistical downscaling is identified for DJF surface temperature over East Asia and North America, and additionally for DJF precipitation over the La Plata Basin of southern South America.

4.1.1 East Asian temperature based on Pacific SST

Over East Asia, the APCC MME (direct model output, DMO) has reasonable skill in simulating wintertime surface temperature, which is found to be significant in terms of correlation over the Korean Peninsula, Japan and an isolated area of central China (Fig. 2). With exception to MCA and PCR, and despite all approaches increasing the RMSE, the overall skill in correlation is improved with all forms of statistical downscaling based on SST in the tropical Pacific, particularly in eastern China. Interestingly, grid point regression is the most skillful approach in terms of correlated area (however the least skillful in terms of RMSE). Among the pattern-based approaches, CPCCA using $\alpha = ß = 0.1$ or LW-based parameters is the most skillful technique in terms of correlated area, whereas BP-CCA clearly yields the lowest RMSE (consistent with the use of far fewer predictors). The variance bias correction for MCA is more skillful than MCA. In this case, it is clear that DMO is more skillful over Japan and the Korean Peninsula, and thus downscaling may be more relevant over China alone, which yields similar results among different approaches (not shown).

4.1.2 North American temperature based on Pacific SLP

Over central-to-northeastern North America, DMO is skillful in simulating wintertime surface temperature only across northern Canada, particular east of the Hudson Bay (Fig. 3). This skill is generally retained for most of the statistical downscaling approaches that use tropical Pacific SLP as a predictor, whereas downscaling dramatically improves skill over the Great Lakes and Midwestern U. S. Among the different approaches, grid point regression is the least skillful (RMSE nearly 3 times more than other approaches) and BP-CCA is clearly the most skillful (largest correlated area and only approach with lower RMSE than DMO). Again, CPCCA using $\alpha = ß = 0.1$ or LW-based parameters is quite skillful in terms of correlated area, although RMSE tends to be higher. BP-CCA and CPCCA using LW-based parameters have the most additional skill in the eastern U. S. and northeastern Canada.

4.1.3 La Plata Basin precipitation based on Indo-Pacific SLP

During austral summer over the lower La Plata Basin in South America, DMO is

skillful in simulating precipitation over a broad region most south of Paraguay and north of Uruguay spanning from the Argentinian Andes to the Atlantic Coast of southern Brazil (Fig. 4). Using Indo-Pacific SLP as a predictor, grid point regression leads to minor improvement near the Andes and a reduction in skill elsewhere, leading to a very large RMSE. However, the region of skill is broadened to the north and south using a single pair of SLP and rainfall patterns for BP-CCA and RCCA, as well as for CPCCA using $\alpha = ß = 0.1$ or LW-based parameters. These methods link rainfall anomalies of generally the same sign with a dipole in SLP over the Pacific largely associated with the Southern Oscillation involved with ENSO (not shown). CCA regularization tends to have higher skill than other approaches, and CPCCA using LW-based parameters has the largest correlated area as well as the lowest RMSE.

4.2 March – May (MAM)

During boreal spring (MAM), ENSO and/or flavors of ENSO have an effect on prediction skill to a lesser extent than during the winter. Other factors become more relevant such as snow cover, forcing from the Atlantic, etc. During MAM, enhanced prediction skill using statistical downscaling is identified for precipitation over Northeast Brazil.

4.2.1 Precipitation in NE Brazil based on Pacific-Atlantic SLP

Over northeastern Brazil, DMO is quite skillful in simulating rainfall (Fig. 5). For most of the statistical downscaling approaches based on Pacific-Atlantic SLP, skill remains high but tends to be slightly to moderately lower than DMO. CPCCA using $\alpha = ß = 0.5$ does yield higher skill than MCA as well as a lower RMSE than DMO. However, BP-CCA is clearly the most skillful approach having a lower RMSE than DMO with a larger correlated area. Similar to results for surface temperature in East Asia, BP-CCA uses a smaller number of modes for prediction.

4.3 June-July-August (JJA)

During boreal summer (JJA), the Asian Monsoon is active for which large inter-annual fluctuations are often forced by ENSO and/or flavors of ENSO, as well as persistent oceanic anomalies in the Indian Ocean. Various studies have investigated the impacts on the Indian Monsoon, although prediction has remained fairly unskillful. On the other hand, prediction

skill for precipitation has been improved over South China by Tung et al. (2013) who conduct statistical downscaling using MCA. Following these results, enhanced prediction skill using statistical downscaling is identified for JJA precipitation over South China and over the western Maritime Continent, as well as for JJA surface temperature over the Indian Subcontinent.

4.3.1 South China precipitation based on Indo-W. Pacific SLP

Over a region spanning South China, Taiwan, the northern Philippines and part of Vietnam, DMO has little to no skill in simulating summer precipitation (Fig. 6). In terms of correlation, precipitation prediction is significantly improved by all of the statistical downscaling approaches based on SLP over the tropical Indian Ocean and West Pacific. This is related to not only ENSO or flavors of ENSO, but also an east-west pressure gradient over the Indian Ocean as well. Again, BP-CCA is quite skillful and has the lowest RMSE with much improvement in southeastern China. CPCCA using ($\alpha = ß = 0.1$) is the most skillful approach in terms of correlated area. All CPCCA solutions tend to improve skill over MCA and RCCA, which encourages further use given that MCA has proven skillful for downscaling with station rainfall (Tung et al. 2013).

4.3.2 W. Maritime Continent precipitation based on Indo-W. Pacific SLP

Over Sumatra and western Malaysia, DMO has moderate skill in simulating boreal summer rainfall in southern Sumatra and isolated parts of Malaysia and northern Sumatra (Fig. 7). This is disappointing given the proximity to slowly-evolving oceanic forcing that gives rise to predictability and prediction skill. Grid point regression based on SLP is found to degrade the skill, whereas all pattern-based approaches improve the skill in terms of correlated area, particularly in central Sumatra. Interestingly, CPCCA solutions are generally less skillful than comparable approaches, but again BP-CCA produces the most skill.

4.3.3 Temperature over the Indian Subcontinent based on Indo-Pacific SLP

During the monsoon season over the Indian Subcontinent, unfortunately there is very little skill in predicting precipitation (not shown), however there is skill in predicting surface temperature (Fig. 8). DMO is skillful largely in South India along the coasts of the

Arabian Sea and Bay of Bengal. Skill is improved in inland India and extended northward by using statistical downscaling with SLP over the tropical Indo-Pacific as a predictor. All pattern-based approaches (with exception to RCCA) have lower RMSE and a larger correlated area than DMO. CPCCA using $\alpha = ß = 0.1$ or LW-based parameters has relatively higher skill than other approaches, however BP-CCA is again the most skillful approach, using less predictors.

4.4 September-October-November (SON)

During boreal autumn (SON), not only ENSO and/or flavors of ENSO can impact global surface temperature and precipitation, but the Indian Ocean Dipole (Saji et al. 1999) tends to reach its peak amplitude, which can impact surrounding South Asia, East Africa and Australia. For SON, enhanced prediction skill using statistical downscaling is identified for both surface temperature and precipitation over Australia. Prediction skill is also identified for surface temperature in East Africa, however the direct MME simulated surface temperature is more skillful (not shown).

4.4.1 Australian temperature based on Indo-W. Pacific SST

Over Australia during Austral spring, DMO is skillful in simulating surface temperature primarily in southeastern Australia, and to a lesser extent in the southwest and northern fringes (Fig. 9). All downscaling approaches based on SST spanning the Indian Ocean and West Pacific significantly improve correlations broadly across central and eastern Australia. However, some minor degradation in skill is found in the southwest. Interestingly, PCR is among the most skillful approaches, and CPCCA does not seem to provide an advantage over other approaches. Consistent with other results, BP-CCA stands out as the most skillful approach with more improvement in the southwest as well.

4.4.2 Australian precipitation based on Indo-W. Pacific SST

For Australian precipitation during Austral spring, DMO has a much smaller area of skill, again confined mainly to parts of northern and southeastern Australia (Fig. 10). Pattern-based statistical downscaling using SST spanning the Indian Ocean and West Pacific does appear to improve prediction skill appreciably, particularly in northwestern and southeastern Australia. CPCCA using various parameters is generally more skillful

than RCCA, however again, BP-CCA is the most skillful approach. BP-CCA and other approaches use fewer coupled modes that are related to a mixture of ENSO and the IOD (not shown).

A synthesis of the downscaling results for all cases is shown in Tables 2 and 3 in terms of RMSE and sum(corr). Similarities in comparisons are discussed further in the next section.

Table 2. The total root mean squared error (RMSE) between predicted and observed variables for direct MME output (DMO), grid point regression (Grid point), PCR, MCA, CPCCA using α = ß = 0.5, CPCCA using α = ß = 0.1, CPCCA using LW-based parameters (LW-CPCCA), RCCA using LW parameters (LW-RCCA) and CCA using EOF truncation (BP-CCA). Values are given for DJF surface temperature (T_{sfc}, × 10^3 K) over East Asia using tropical Pacific SST (1), DJF T_{sfc} over N. America using Pacific SLP (2), DJF precipitation (PREC, × 10^3mm/day) over the lower La Plata Basin using Indo-Pacific SLP (3), MAM PREC over NE Brazil using Pacific-Atlantic SLP (4), JJA PREC over South China using Indo-W. Pacific SLP (5), JJA PREC over the W. Maritime Continent using Indo-Pacific SLP (6), JJA T_{sfc} over India using Indo-Pacific SLP (7), as well as SON T_{sfc} (8) and PREC (9) over Australia using Indo-W.Pacific SST.

	DMO	Grid point	PCR	MCA	CPCCA α=ß=0.5	CPCCA α=ß=0.1	LW-CPCCA	LW-RCCA	BP-CCA
1	61.2	80.4	78.4	80.1	76.6	74.1	73.6	74.6	66.8
2	203.0	701.1	205.9	258.1	283.9	316.0	270.0	212.1	191.5
3	24.5	65.4	24.8	24.7	24.7	22.9	22.6	22.6	23.5
4	42.6	72.6	47.6	43.6	42.5	46.9	44.7	45.5	39.1
5	62.5	147.2	62.2	60.3	60.1	61.3	60.5	60.2	57.2
6	12.4	28.3	12.3	13.3	13.2	13.6	12.4	11.7	10.9
7	8.1	11.6	7.2	7.1	7.4	7.5	7.4	7.2	6.8
8	33.0	57.1	29.6	29.3	31.3	37.1	35.6	30.1	28.8
9	11.6	19.8	10.9	10.2	10.3	10.7	11.4	10.7	10.1

Table 3. Same as in Table 2 but for the total number of grid points that are significantly positively correlated (indicated in Figs. 2 – 10 by sum(corr), 5% significance using two-sided t-test associated with a value of about 0.38).

	DMO	Grid point	PCR	MCA	CPCCA α=ß=0.5	CPCCA α=ß=0.1	LW-CPCCA	LW-RCCA	BP-CCA
1	730	1384	691	715	952	1303	1275	1049	933
2	1527	2249	2659	2497	2547	2803	2795	2638	2876
3	258	260	236	257	263	318	335	318	289
4	749	572	706	708	714	712	722	706	790
5	29	237	220	229	247	273	266	239	266
6	130	57	178	162	157	155	156	175	210
7	813	760	827	828	832	866	850	808	894
8	1473	1431	1831	1806	1822	1757	1774	1826	1987
9	695	653	755	810	819	812	798	760	918

5. SUMMARY AND CONCLUSIONS

Given the various approaches for analyzing covariance and isolating linear relationships between multiple datasets, regularization and unified (or continuum) approaches are beneficial for reducing the effect of sample noise. Ridge regression has been used to regularize CCA (RCCA, Vinod 1976); although up to this point the comparable technique of continuum power regression (Lorber et al. 1987, Stone and Brooks 1990) has not yet been similarly applied. This is done in this study for which Continuum Power CCA (CPCCA) is introduced as a new and effective approach. As opposed to applying the full whitening

transformation to individual variables during the computation of CCA, CPCCA applies the partial whitening transformation. The partial whitening transformation acts to decorrelate variables, but only to a partial degree that is specified according parameters α and ß. This acts to condition the sample covariance through exponential shrinking and can provide an alternative estimate that is closer to the population covariance.

CPCCA maximizes the squared cross-covariance between partially-whitened variables. As a unified approach, the range of CPCCA solutions bridge those of CCA (α = ß = 0), MCA (α = ß = 1) and RDA (α = 0 and ß = 1, or vice-versa) with PCR as an upper limit (α→∞ or ß→∞). Within this range, useful CPCCA solutions are a regularization for CCA, a variance bias correction for MCA (α = ß = 0.5), and a regularization for RDA (fractional α with fixed , ß = 1 or vice-versa). For regularization, previous results applying CPCCA to synthetic data indicate that a parameter value near 0.1 tends to be a good choice for a small sample of highly autocorrelated data. In addition to this, a more objective, sample dependent choice is proposed based on the LW covariance estimate for ridge regression (Ledoit and Wolf 2004). This involves choosing partial whitening parameters that estimate the sample covariance matrix closest to the LW estimate, which appears to be quite effective and closer to the population covariance for high autocorrelations.

Applied to statistical downscaling of surface temperature and precipitation for various seasons and regions, CPCCA is found to yield quite skillful prediction and improve direct model output (DMO), with results comparable to other methods. CPCCA using LW-based parameters yielded higher skill than RCCA (using the LW covariance estimate) in terms of correlated area for 7 out of the 9 downscaling experiments (Table 3). This appears consistent with prior Monte Carlo results in which shifting the LW covariance estimate to an estimate constructed with partial whitening yielded a sample covariance consistently closer to the population covariance (for high autocorrelations). However, interestingly CPCCA yields a higher RMSE than RCCA for two-thirds of the cases (Table 2), an apparent contradiction. There is some indication that including additional predictors that results in modest increases the area of significant correlation consequently can increase RMSE in the surrounding unskillful areas. Part of this may be an artifact of choosing the number of modes based on the highest correlated area. Nevertheless, in many cases the plane of solutions spanned by CPCCA is preferable to that spanned by RCCA, however comparison

is sample dependent.

Interestingly, CCA using a prior EOF truncation (BP-CCA, Barnett and Preisendorfer 1987) tends to yield the most skillful prediction and results often stand out with the least number of predictors, which is an attractive quality. Specifically, among the other downscaling approaches, BP-CCA has the largest correlated area in 6 out of 9 experiments, and the lowest RMSE in 8 out of 9 of the experiments. However, in this study BP-CCA likely has inflated skill over other approaches in that the exact EOF truncation used is based on skill from cross-validation, whereas the parameters used for RCCA and CPCCA are not. It is unclear if using another basis for determining EOF truncation would yield similar results, although it is found that BP-CCA results are quite sensitive to the number of EOFs retained and the number of modes used (consistent with past results).

Overall, statistical downscaling using SVD-based approaches that identify coupled patterns is found to be quite useful and can improve skill significantly in seasonal surface temperature and precipitation over that from dynamical models in an MME framework. As only average skill is considered in the study, more can be done to quantify forecast skill conditional on the magnitude of downscaling predictors (e.g., strength of ENSO). For the various regions examined, this report may provide a guide for improving seasonal prediction. In these regions, the use of CPCCA and BP-CCA is recommended for application to operational seasonal forecasting by producing an additional downscaled forecast to act as guidance and/or a replacement of the dynamical MME forecast.

REFERENCES

Akaho, S., 2001: A kernel method for canonical correlation analysis. Int'l Meeting on Psychometric Society, Osaka, Japan.

Akaike, H., 1973: Information theory and an extension of the maximum likelihood principle. Second International Symposium for Information Theory, 267–281.

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, 1147–1167.

Bakalian, F., H. Ritchie, K. Thompson, and W. Merryfield, 2010: Exploring atmosphere-oceancoupling using Principal Component and Redundancy Analysis. J. Climate, 23, 4926–4943.

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 T. M. Smith, 1996: Specification and prediction of global surface temperature and precipitation from global SST using CCA. J. Climate, 9, 2660-2697.

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

Cruz-Cano, R. and M.-L. T. Lee, 2014: Fast regularized canonical correlation analysis. Comput. Stat. Data An., 70, 88–100.

Czaja, A. and C. Frankignoul, 2002: Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation. J. Climate, 15, 606-623.

de Jong, S. and H. A. L. Kiers, 1992: Principal Covariates Regression part 1. theory. Chemom. Intell. Lab. Syst., 14, 155–164, doi:10.1016/0169-7439(92)80100-i.

DelSole, T. and P. Chang, 2003: Predictable component analysis, canonical correlation analysis, and autoregressive models. J. Atmos. Sci., 60, 409–416.

DelSole, T. and J. Shukla, 2006: Specification of wintertime North American surface temperature. J. Climate, 19, 2691–2716.

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

Fischer, M. J., 2014: Regularized principal covariates regression and its application to finding coupled patterns in climate fields. J. Geophys. Res., 119, 1266–1276.

Frankignoul, C. and N. Sennechael, 2007: Observed influence of North Pacific SST anomalies on the atmospheric circulation. J. Climate, 20, 592–606.

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

Hotelling, H., 1936: Relations between two sets of variates. Biometrika, 28, 321–377.

Hurvich, C. M. and C.-L. Tsai, 1989: Regression and time series model selection in small samples. Biometrika, 76, 297-307.

Kao, H.-Y. and J.-Y. Yu, 2009: Contrasting eastern-Pacific and central-Pacific types of ENSO. J. Clim., 22, 1499–1515.

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.

Ledoit, O. and M. Wolf, 2004: A well-conditioned estimator for large-dimensional covariance matrices. J. Multivar. Anal., 88, 365–411.

Lim, Y., S. Jo, J. Lee, H.-S. Oh, and H.-S. Kang, 2011: An improvement of seasonal climate prediction by regularized canonical correlation analysis. Int. J. Climatol., 32, doi:10.1002/ joc.2368.

Lorber, A., L. E. Wangen, and B. R. Kowalski, 1987: A theoretical foundation for the PLS algorithm. J. Chemometrics, 1, 19–31.

Harris, I., P. D. Jones, T. J. Osborn and D. H. Lister, 2012: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 dataset. Submitted to International

Journal of Climatology.

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

Ribes, A., J.-M. Azais, and S. Planton, 2009: Adaptation of the optimal fingerprint method for climate change detection using a well-conditioned covariance matrix estimate. Clim. Dyn., 33, 707–722.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. Nature, 401, 360–363.

Schneider, T., 2001: Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. J. Climate, 14, 853–871.

Smerdon, J. E., A. Kaplan, D. Chang, and M. N. Evans, 2010: A pseudoproxy evaluation of the CCA and RegEM methods for reconstructing climate fields of the last millennium. J. Climate, 23, 4856–4880.

Smoliak, B. V., J. M. Wallace, M. T. Stoelinga, and T. P. Mitchell, 2010: Application of partial least squares regression to the diagnosis of year-to-year variations in Pacific Northwest snowpack and Atlantic hurricanes. Geophys. Res. Lett., 37, doi:10.1029/2009GL041478.

Sohn, S.-J., J.-B. Ahn, and C.-Y. Tam, 2013: Six month–lead downscaling prediction of winter to spring drought in South Korea based on a multimodel ensemble. Geophys. Res. Lett., 40, doi: 10.1002/grl.50133.

Stone, M. and R. J. Brooks, 1990: Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression. J. Roy. Statist. Soc. Ser. B, 52, 237–269.

Swenson, E. T., 2014: Continuum Power CCA: a unified approach for isolating coupled modes. Accepted to J. Climate.

Tingley, M. P., P. F. Craigmile, M. Haran, B. Li, E. Mannshardt, and B. Rajaratnam, 2012: Piecing together the past: statistical insights into paleoclimatic reconstructions.

Quat. Sci. Rev., 35, 1–22.

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

Tung, Y. L., C.-Y. T. S.-J. Sohn, and J. L. Chu, 2013: Improving the seasonal forecast for summertime South China rainfall using statistical downscaling. J. Geophys. Res., 118, 5147-5159.

Vinod, H. D., 1976: Canonical ridge and econometrics of joint production. J. Econometrics, 4, 176–166.

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. Climate, 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, P. R. Krishnaiah, Ed., Academic Press, 391–420.

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

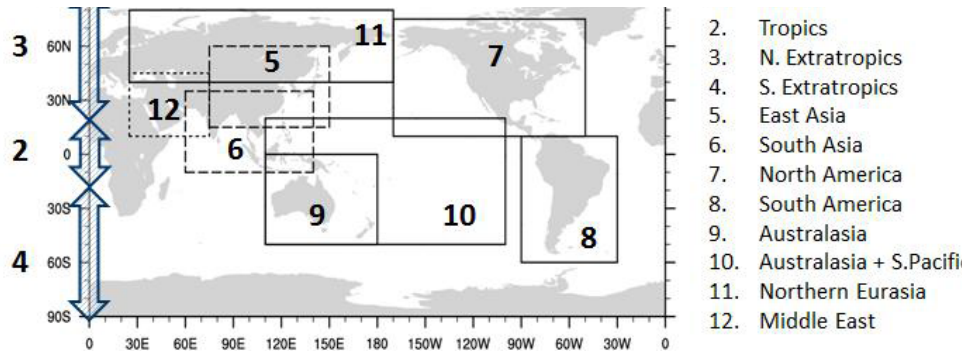

Figure 1. Domains used for statistical downscaling. Blue boxes indicated different predictor domains used (for either SST or SLP), and red boxes indicate different predictand domains used (for either surface temperature or precipitation). The exact coordinates of domains are provided in the captions of Figs. 2 – 10.

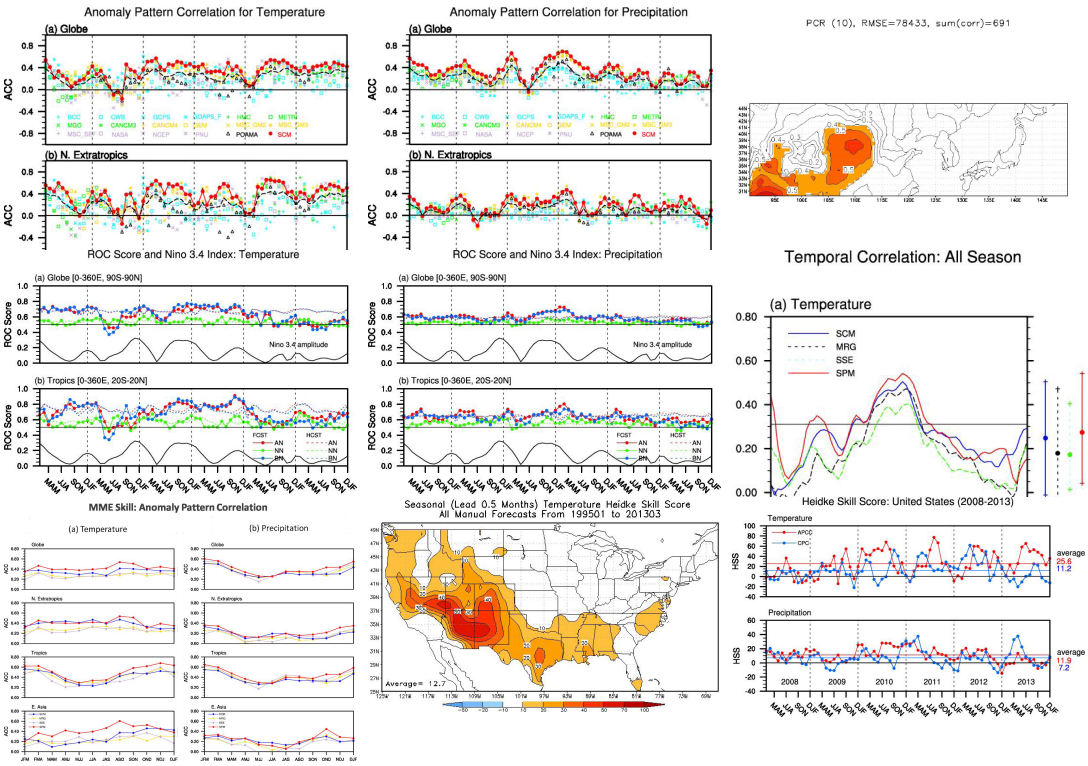

Figure 2. Correlation between predicted and observed surface temperature (contours, 5% significance shaded using two-sided t-test) over East Asia (90°E – 150°E, 30°N – 45°N) during DJF for direct MME output (DMO, upper left) and from statistical downscaling using MME SST over the tropical Pacific Ocean (120°E – 90°W, 20°S – 20°N) as the predictor. Downscaling approaches apply a leave-one-out cross-validation and include linear regression at each grid point (Grid point, upper center), PCR, (upper right), MCA (middle left), BP-CCA (using 10, 4 EOFs, middle center), RCCA (using LW parameters, middle right) as well as CPCCA using $\alpha = ß = 0.5$ (lower left), $\alpha = ß = 0.1$ (lower center) and LW-based parameters (lower right). The parameter values and/or number of leading modes used are indicated in parentheses. The total root mean squared error (RMSE) and the number of grid points that are significantly positively correlated, indicated by sum(corr), are also given.

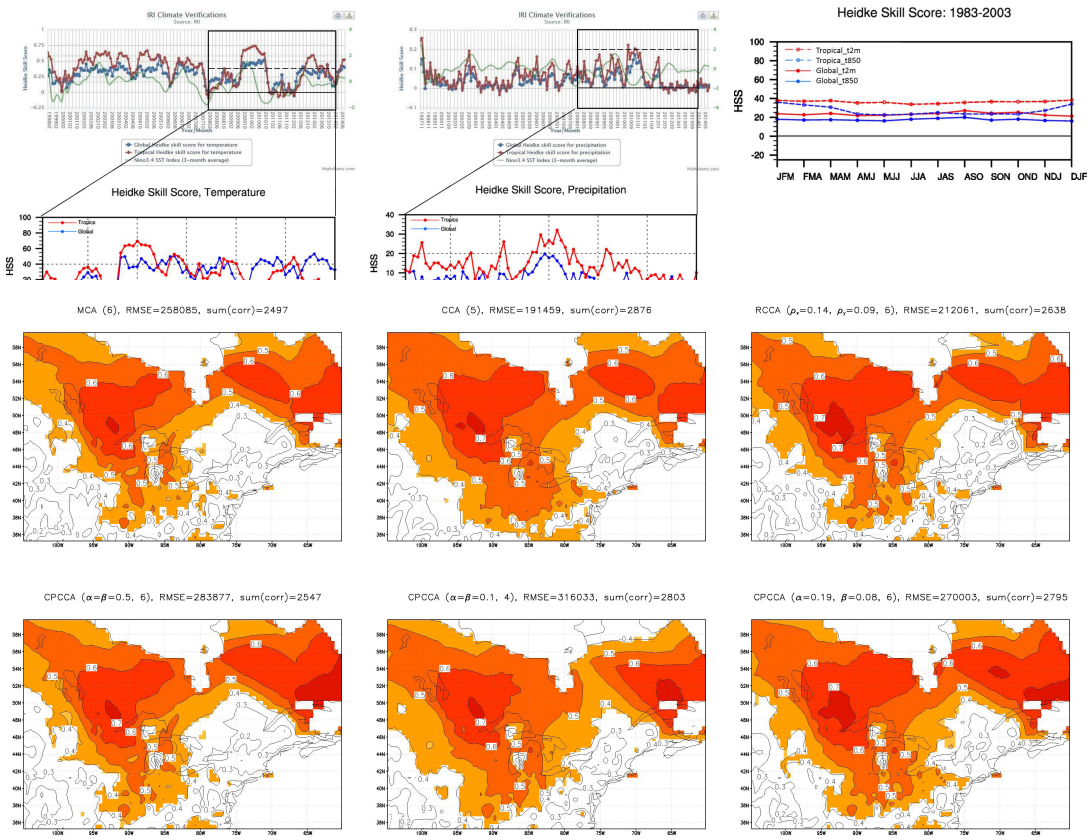

Figure 3. Same as Fig. 2 except for DJF surface temperature over North America (centered on the Great Lakes region, 105°W – 60°W, 35°N – 60°N) for which statistical downscaling approaches use MME SLP over the tropical Pacific Ocean (120°E – 90°W, 20°S – 20°N) as the predictor. Also, BP-CCA uses 6 EOFs for SLP and 7 EOFs for temperature (middle center).

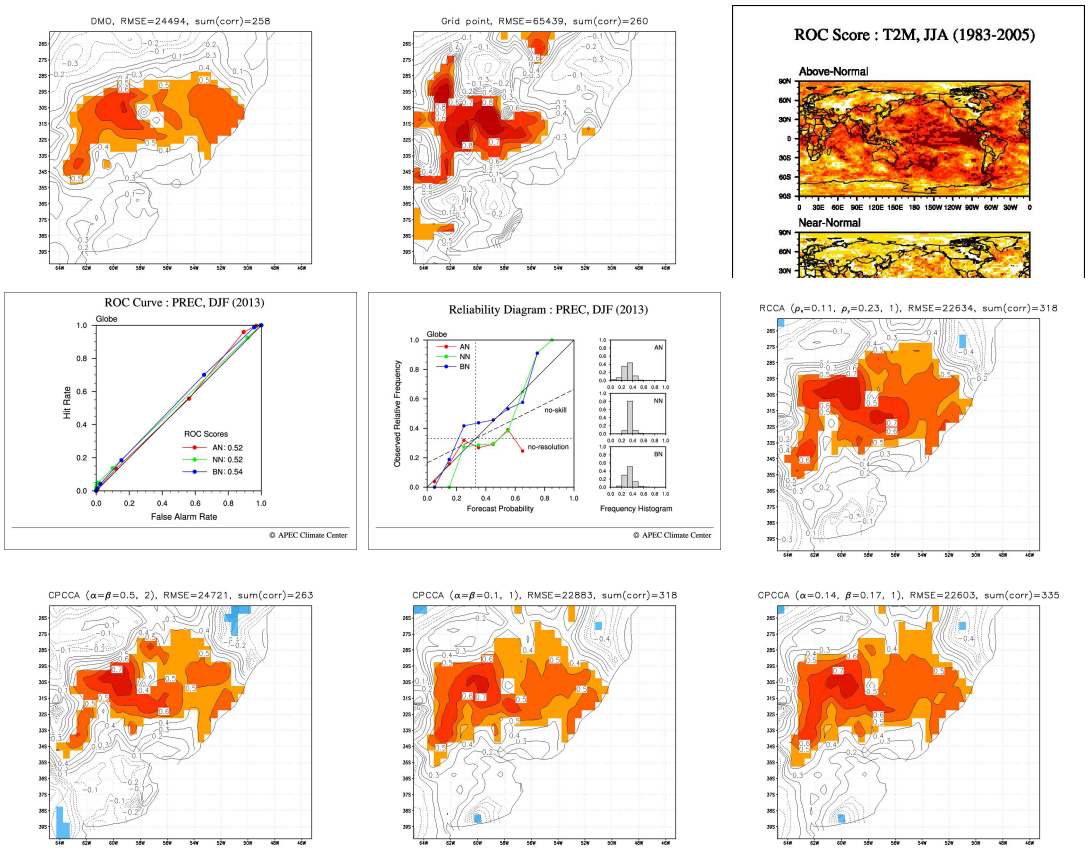

Figure 4. Same as Fig. 2 except for DJF precipitation over the lower La Plata Basin in South America (65°W – 45°W, 40°S – 25°S) for which statistical downscaling approaches use MME SLP over the tropical Indo-Pacific (30°E – 90°W, 20°S – 20°N) as the predictor. Also, BP-CCA uses 2 EOFs for SLP and 7 EOFs for precipitation (middle center).

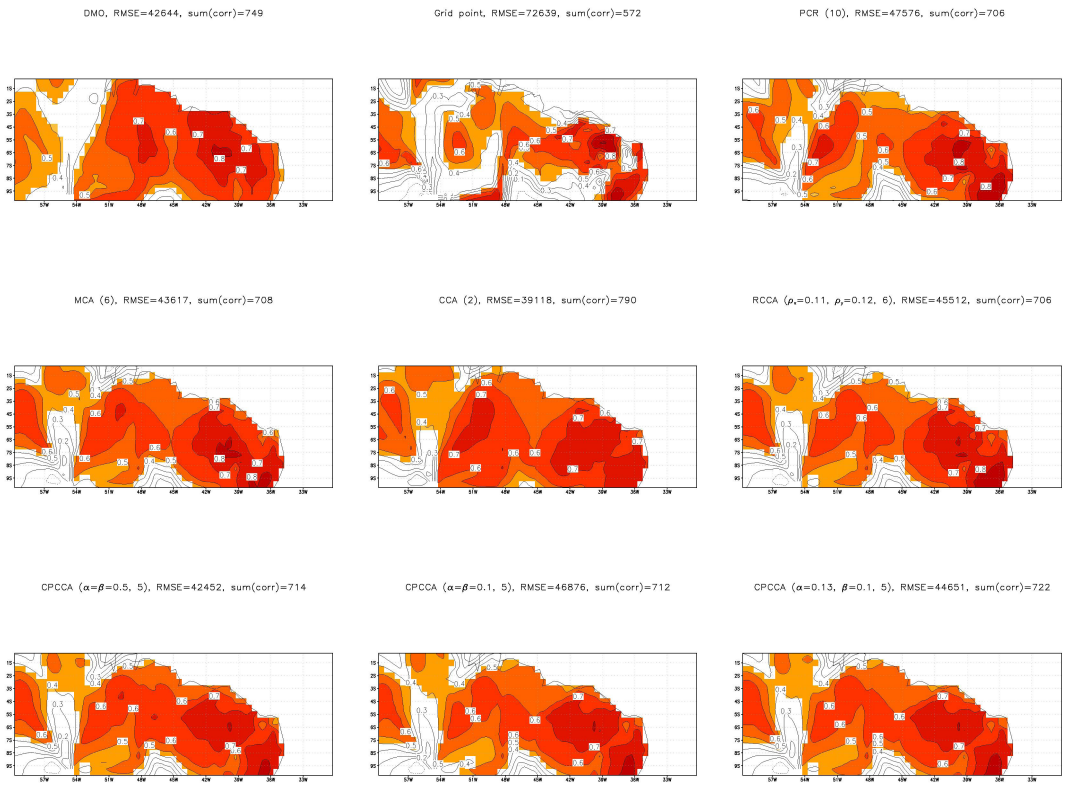

Figure 5. Same as Fig. 2 except for MAM precipitation over Northeast Brazil (60°W – 30°W, 10°S – 0°) for which statistical downscaling approaches use MME SLP spanning the tropical Pacific and Atlantic Oceans (120°E – 0°, 20°S – 20°N) as the predictor. Also, BP-CCA uses 10 EOFs for SLP and 5 EOFs for precipitation (middle center).

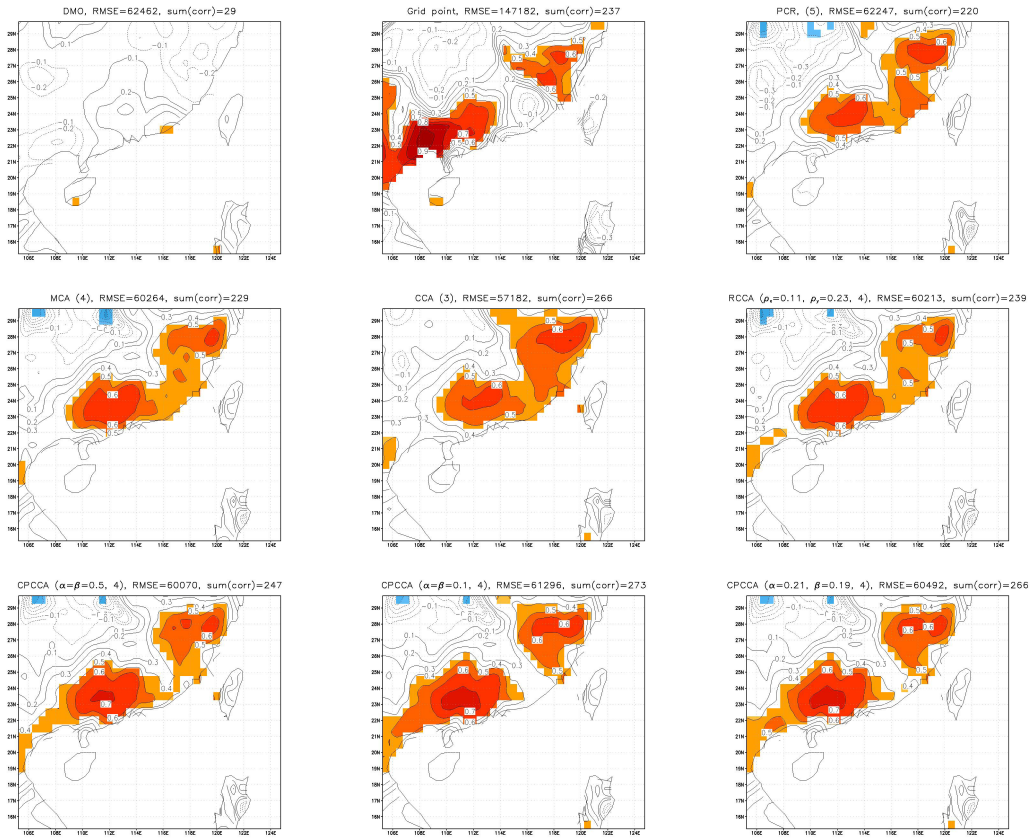

Figure 6. Same as Fig. 2 except for JJA precipitation over South China (105°E – 125°E, 15°N – 30°N) for which statistical downscaling approaches use MME SLP over the tropical Indian Ocean and West Pacific (30°E – 180°, 20°S – 20°N) as the predictor. Also, BP-CCA uses 5 EOFs for SLP and 3 EOFs for precipitation (middle center).

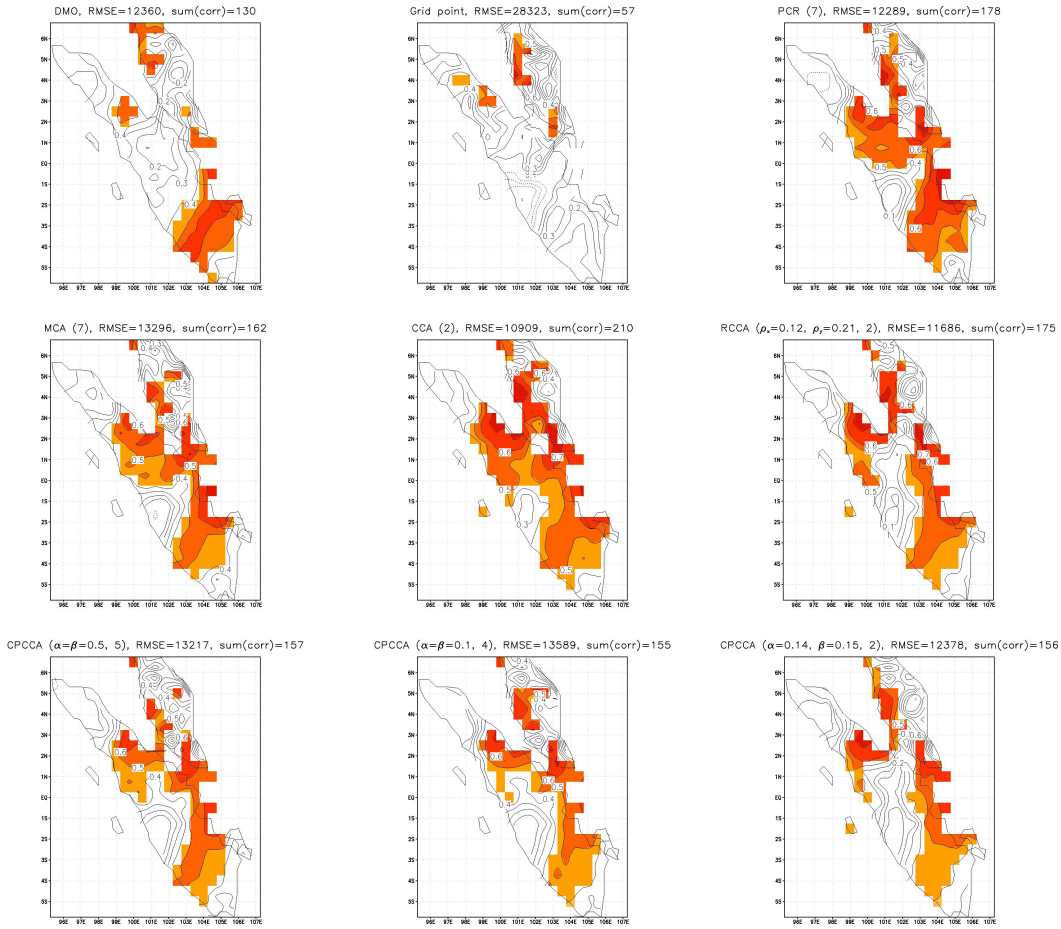

Figure 7. Same as Fig. 2 except for JJA precipitation over the western Maritime Continent (95°E – 107.5°E, 6°S – 7°N) for which statistical downscaling approaches use MME SLP over the tropical Indo-Pacific (30°E – 90°W, 20°S – 20°N) as the predictor. Also, BP-CCA uses 7 EOFs for SLP and 2 EOFs for precipitation (middle center).

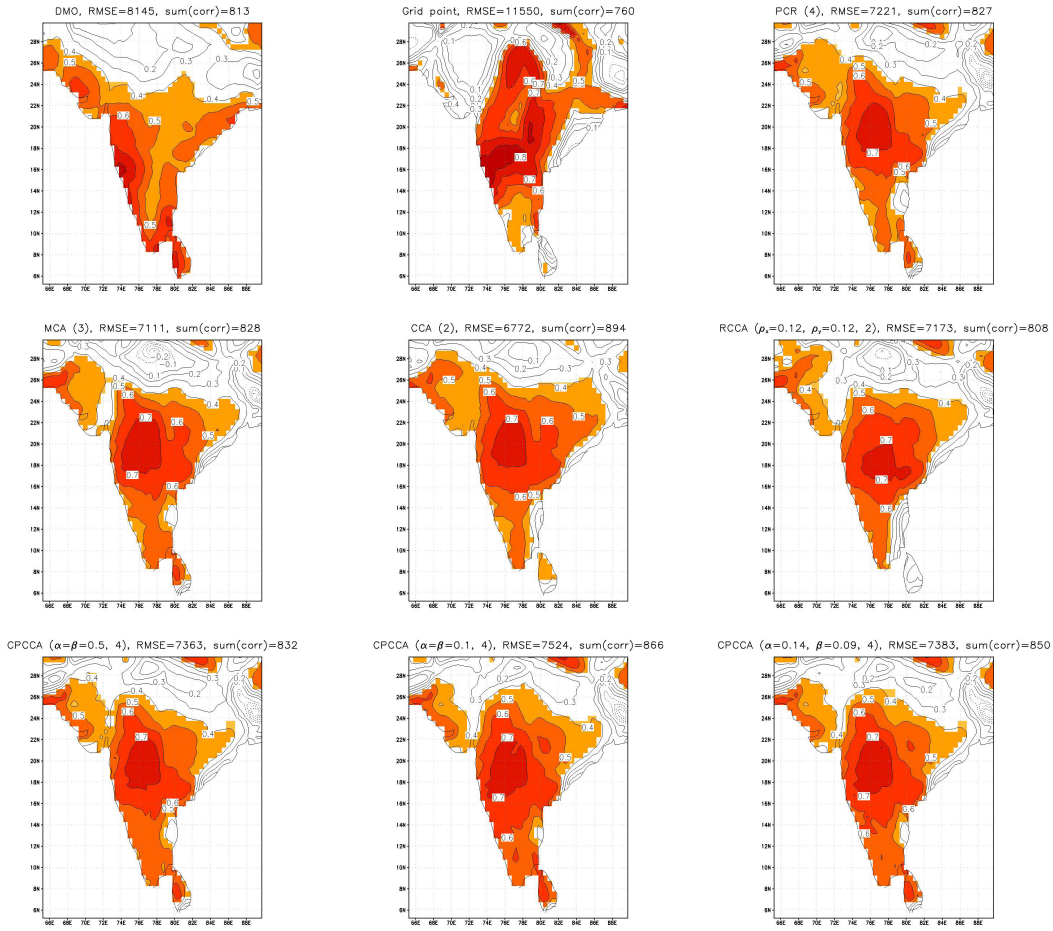

Figure 8. Same as Fig. 2 except for JJA surface temperature over the Indian Subcontinent (65°E – 90°E, 5°N – 30°N) for which statistical downscaling approaches use MME SLP over the tropical Indo-Pacific (30°E – 90°W, 20°S – 20°N) as the predictor. Also, BP-CCA uses 3 EOFs for SLP and 2 EOFs for temperature (middle center).

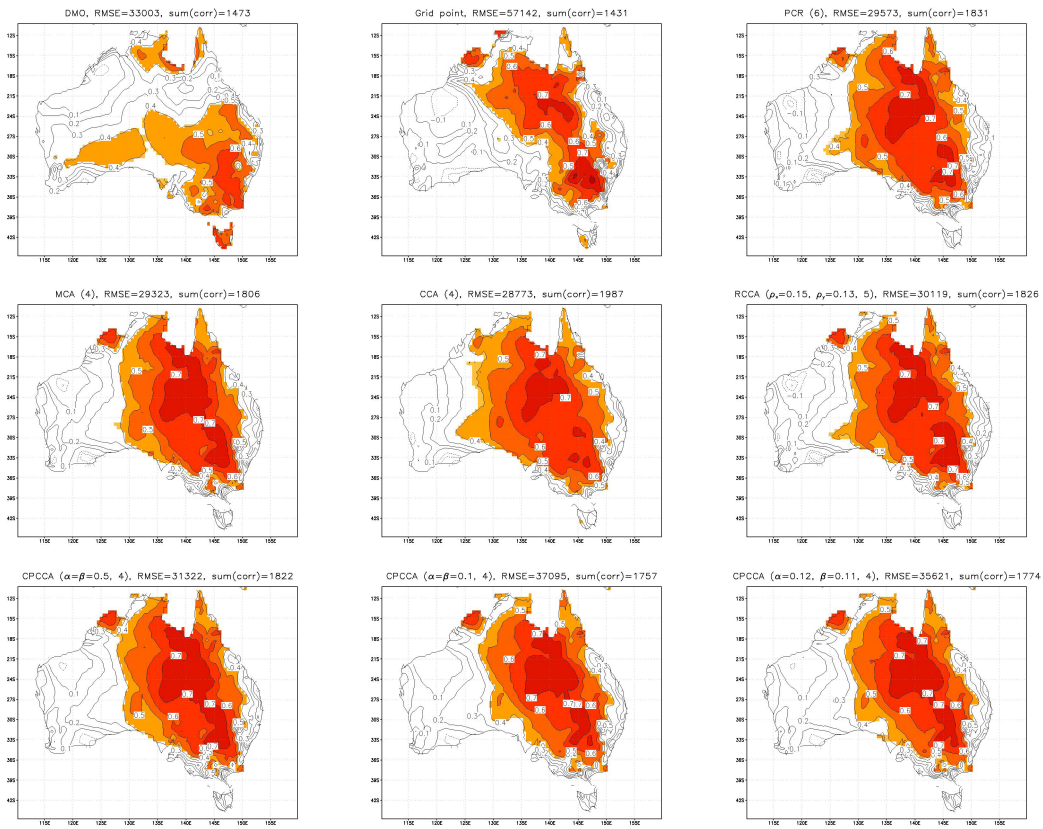

Figure 9. Same as Fig. 2 except for SON surface temperature over Australia (110°E – 160°E, 45°S – 10°S) for which statistical downscaling approaches use MME SST over the tropical Indian Ocean and West Pacific (30°E – 180°, 20°S – 20°N) as the predictor. Also, BP-CCA uses 6 EOFs for SST and 4 EOFs for temperature (middle center).

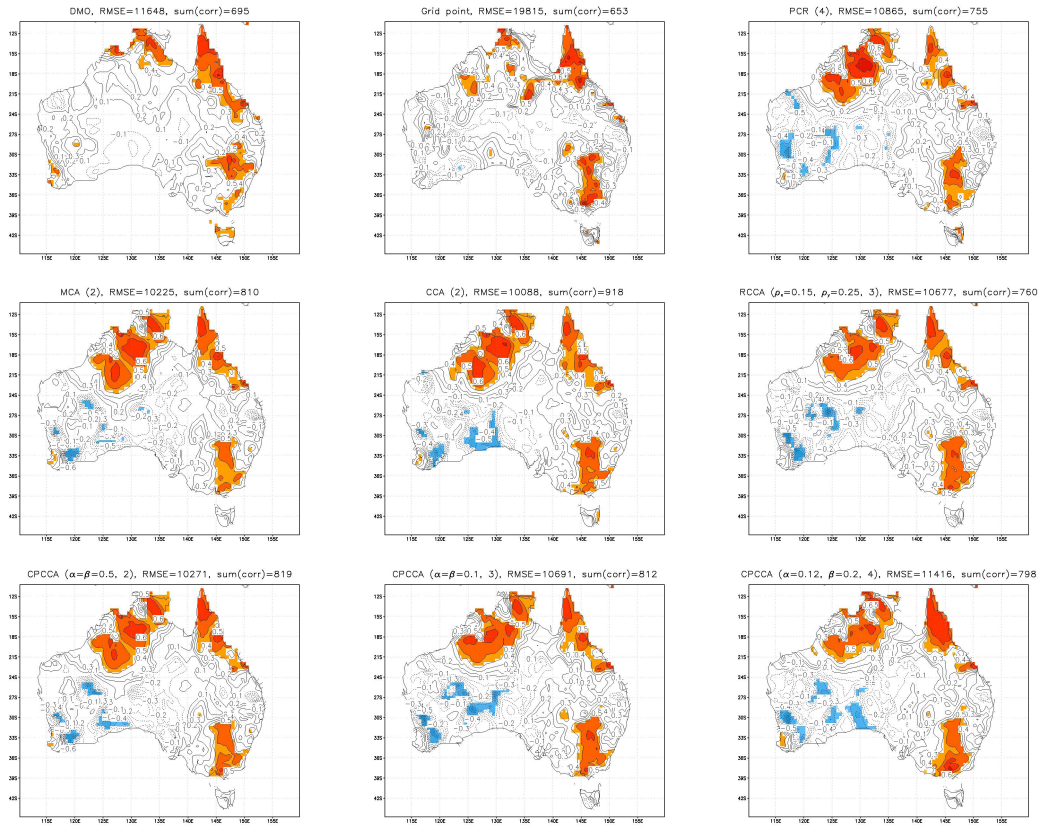

Figure 10. Same as Fig. 2 except for SON precipitation over Australia (110°E – 160°E, 45°S – 10°S) for which statistical downscaling approaches use MME SST over the tropical Indian Ocean and West Pacific (30°E – 180°, 20°S – 20°N) as the predictor. Also, BP-CCA uses 4 EOFs for SST and 7 EOFs for precipitation (middle center).

RESEARCH REPORT 2015-04

Continuum Power CCA and Its Application to Statistical Downscaling

Erik T. Swenson Climate Prediction Team

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

www.facebook.com/apcc21 www.twitter.com/apcc21 www.flickr.com/apcc21
www.youtube.com/APECClimateCenter21 www.plus.google.com/+APECClimateCenter21