An Exploratory Study Of Seasonal Rainfall Variability In Australia Using Independent Component Analysis

Westra, S. & Sharma, S.

School of Civil and Environmental Engineering, University of New South Wales.
E-Mail: seth@civeng.unsw.edu.au

Keywords: *Hydrology, Australian rainfall, ICA, PCA, ENSO.*

EXTENDED ABSTRACT

Component extraction techniques have been used frequently by climate and water resources researchers to analyse high dimensional datasets such as global sea surface temperature (SST) and rainfall time series. The motivation for using these techniques is usually twofold; firstly to reduce the dimension of the dataset, by representing the data using a small number of components that are able to describe a significant proportion of the total variance, and secondly to enhance our understanding of the dynamics of the underlying system, by interpreting these components as representing physically significant 'modes' of climate variability.

In this study we explore the potential of a relatively new technique known as independent component analysis (ICA), which has been developed as a means to separate mixtures of signals when little is known about either the original signals, or the manner in which they have been mixed. This is a problem that occurs frequently in the climate field, where one wishes to understand the factors that contribute to the dynamical nature of a given set of observations.

The premise of ICA is based on central limit theorem, which asserts that if a set of independent random variables is mixed using a linear transformation, the result will be a set of variables that tend towards a Gaussian distribution. Reversing this logic, if one rotates a mixed dataset in a manner that maximises the divergence from a Gaussian distribution, then under certain conditions it is possible to retrieve the original independent variables. Therefore, ICA focuses on higher-order statistics that measure the divergence from a Gaussian distribution.

The ICA method is contrasted to the more widely used principal component analysis (PCA), which removes the correlation between the components while at the same time maximising the variance of successive principal components. This latter property in particular has proved to be useful to reduce the dimension of the datasets while retaining much of the information, and in the present study PCA is also used as a pre-processing step for ICA. The primary distinction between ICA and PCA is that while PCA uses only second order statistics to obtain uncorrelated components, ICA maximises the independence between components through the use of higher order statistics. Thus, while PCA may be well suited to variance maximisation and dimension reduction, ICA is fundamentally more suited to ensuring the statistical independence of the components and in certain cases is also capable of determining the underlying causes of this variability.

To demonstrate the potential of the ICA technique in highlighting physically 'interesting' modes of variability, we apply PCA and ICA to a set of seasonal rainfall time series from over 200 rainfall gauges located around the Australian continent. It is assumed, based on the results of a number of earlier studies, that the El Niño Southern Oscillation (ENSO) phenomenon is an important factor in influencing Australian rainfall. Furthermore, it has been shown that an inter-decadal phenomenon, particularly the Inter-decadal Pacific Oscillation (IPO), may influence the degree of correlation between ENSO and Australian rainfall, with an enhanced link when the IPO is negative, and a reduced link when the IPO is positive.

The results of this study consistently show that, for each season of the year, one independent component is significantly correlated with an index of the ENSO phenomenon known as the Southern Oscillation Index (SOI), with the highest correlation occurring during spring. Furthermore, during the IPO negative phase from 1946-1977, the correlation between one of the independent components and the SOI is further enhanced. This is contrasted with the PCA solution, in which the correlation coefficients for the majority of cases are not statistically significant. These results therefore indicate that ICA may have a significant potential to be applied to a number of alternative climate datasets to develop an improved understanding of the climate dynamics that govern those datasets.

1. INTRODUCTION

Component extraction techniques have been used in a variety of climate studies to (1) reduce the dimension of large datasets, and (2) aid in the identification and interpretation of significant 'modes' of climate variability. These two types of analysis are usually performed simultaneously, so that large climate datasets such as Australia-wide rainfall or the historical sea surface temperature (SST) reconstructions are represented by a relatively small set of 'components', with each component assigned a physical interpretation relating to the dynamics of the underlying system (Zwiers and Von Storch, 2004).

The dominant component extraction technique currently in use is known as principal component analysis (PCA), or the closely related empirical orthogonal function analysis (EOF; Wilks, 1995). This technique focuses on second-order statistics, by reducing the correlation of the extracted components while at the same time maximizing the variance (in a least squares sense) of successive principal components. There are many examples in the literature where this technique has been applied to climate datasets, and this has assisted in the identification of a wide range of climate phenomena, such as the Interdecadal Pacific Oscillation (IPO; Zhang et al, 1997), the Indian Ocean Dipole (IOD; Saji et al, 1999), and the Artic Oscillation (AO; Thompson and Wallace, 1998).

Despite the popularity of PCA, the technique suffers from a number of limitations, such as the restriction that principal components must be mutually orthogonal, thereby limiting the interpretability of components after the first component, and that successive components must explain the maximum remaining variance, thus potentially resulting in the mixing of several independent physical phenomena into one principal component. To address these concerns, a variety of rotation techniques have been developed (Richman, 1986), which involve the application of a linear transform to the PCs, so that the PCs are no longer constrained to be orthogonal and so the interpretations of the rotated PCs can be simplified. This is usually achieved by minimizing some form of objective function, which aims to measure certain properties such as the simplicity of the geometrical patterns identified by the rotated PCs. The wide variety of available objective functions illustrates one of the fundamental limitations of rotational PCA methods, which is that the criteria used for rotation will always be somewhat arbitrary, so that the meaning of the rotated PCs will always be open to interpretation.

A recently developed alternative technique is known as independent component analysis (ICA), which seeks to separate mixtures of signals when little is known about either the nature of the original signals, or the manner in which they have been mixed. ICA seeks to achieve this aim by maximizing the independence of the extracted components. Thus, in contrast to second-order methods such as PCA, ICA not only removes the correlation between the signals but also reduces higher-order statistical dependencies, attempting to make the signals as independent from each other as possible (Hyvarinen et al, 2001).

The objective of this paper is to highlight the differences between PCA and ICA, and to demonstrate that ICA has some important advantages with respect to the interpretability of the extracted components. We will commence with a review of the theory of ICA, and a simple synthetic example to illustrate the difference between ICA and PCA. These two methods will then be applied on a dataset of Australian rainfall, and differences between the results will be discussed.

2. INDEPENDENT COMPONENT ANALYSIS

2.1. Overview

The ICA method was first introduced by Herault and Jutten [1986], and has been applied successfully in a wide range of fields, including blind source separation and feature extraction (see Hyvarinen et al, 2001, and references therein). The simplest and most commonly used form of ICA involves the mixing of an n-dimensional source vector, $\mathbf{s} = (s_1,\ldots, s_n)^T$, referred to as the independent components (ICs), resulting in an m-dimensional observation vector, $\mathbf{x} = (x_1,\ldots, x_m)^T$ (Comon, 1994) These ICs are assumed to be non-Gaussian (with the possible exception of at most one IC; Hyvarinen *et al*, 2001), mutually statistically independent and zero-mean. In addition, it is assumed in this paper that $n \leq m$. Put into vector-matrix notation, and assuming that the mixing is both linear and stationary, yields:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (1)$$

where $\mathbf{A}$ is known as the mixing matrix of dimension $m \times n$. The objective of ICA is to estimate the mixing matrix, $\mathbf{A}$, as well as the independent components, $\mathbf{s}$, knowing only the observations $\mathbf{x}$. This can be achieved up to some scalar multiple of $\mathbf{s}$, since any constant multiplying an independent component in Equation 1 can be cancelled by dividing the corresponding column of the mixing matrix **A** by the same constant.

Central to the identification of the ICs from the data **x** is the assumption that all except at-most one IC will be "maximally non-Gaussian" (Hyvarinen *et al*, 2001). This follows from the logic outlined in the central limit theorem, which is that if one mixes independent random variables through a linear transformation, the result will be a set of variables that tend to be Gaussian. If one reverses this logic, it can be presumed that the original independent components must have a distribution that has minimal similarity to a Gaussian distribution. Consequently, the approach adopted to extract ICs from data containing mixed signals amounts to finding a transformation that results in variables that exhibit maximal divergence from a Gaussian distribution as defined through an appropriately specified statistic. This focus on higher order statistics explains why ICA is often successful at finding sources when techniques such as PCA fail completely (Oja, 2004).

2.2. An example

To illustrate some important differences between PCA and ICA, consider the following example. Here, we combine two independent signals which each have a uniform distribution $U[-\sqrt{3},\sqrt{3}]$, denoted s_1 and s_2, using a mixing matrix **A** defined as follows:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \tag{2}$$

The bounds of the uniform distribution were selected so that the signals had unit variance. The joint probability density function of the mixed components, x_1 and x_2, for a sample of size 1000 is shown in Figure 1. As expected based on the central limit theorem, the mixed signals appear to be much closer to a Gaussian distribution compared to the original signals, which by construction are uniformly distributed. The PCA directions are also shown, with $\mathbf{e_1}$ representing the direction of maximum variance, and $\mathbf{e_2}$ constrained to be orthogonal to $\mathbf{e_1}$.

To simplify the ICA process, it is common to 'pre-whiten' the data, which means projecting the data on the principal directions shown in Figure 1, and then standardizing the data so that each direction has unit variance. This is frequently achieved using PCA, since this also allows for the dimension of the data to be reduced, thereby facilitating the optimization process for ICA. We thus have the following relationships between the independent components, **s**, the observed variables, **x**, and the whitened data, **z**:

$$\mathbf{z} = \mathbf{Vx} = \mathbf{VAs} \tag{3}$$

where **V** is a linear transform used to whiten the data. The whitened variables are shown in Figure 2, and show that marginal probability distributions are clearly not uniformly distributed, thus demonstrating that the original (uniformly distributed) source signals, **s**, still have not been found. The advantage of this process, however, is that since we are now dealing with a transformed variable whose elements have zero mean, are mutually uncorrelated and have unit variance, the ICA solution is now limited to some orthogonal rotation of the whitened data set about the origin.

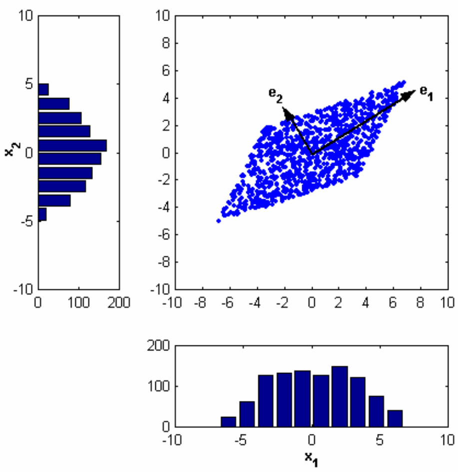

Figure 1: Plot of $\mathbf{x} = [x_1, x_2]$. The directions of eigenvectors $\mathbf{e_1}$ and $\mathbf{e_2}$ of the data are shown, and represent the principal directions of the bi-variate dataset. The principal components are the projections of **x** onto the principal directions $\mathbf{e_1}$ and $\mathbf{e_2}$.

Denoting a unit vector defining a line passing through the origin of the data in Figure 2 by **w,** then the projection of **z** on the line is given by $\mathbf{y} = \mathbf{w}^T\mathbf{z}$. It has been shown (Oja, 2004) that due to pre-whitening the data, no matter what the angle of the projection, it always holds that **y** has zero mean and unit variance. The object of ICA, therefore, is to find a suitable vector, **w**, that ensures the resulting components **y** are independent, which is obtained through a maximisation of the higher-order moments of $\mathbf{w}^T\mathbf{z}$ as described in section 2.3.

The ICA solution is shown in Figure 3, and illustrates that the optimum solution recovers the uniform distribution of the original signals. This process is repeated until all the ICs are found, and

w approximates one of the rows of the matrix $[\mathbf{VA}]^{-1}$. Thus, **y** becomes an estimator of the original independent components, **s**.

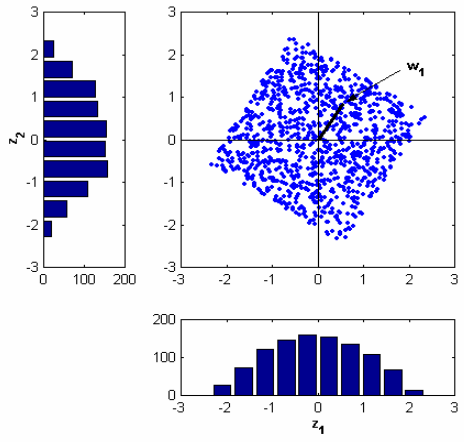

Figure 2: Plot of the whitened data time series $\mathbf{z} = [z_1, z_2]$. To find the ICA solution, we search for a vector **w** such that the projection $\mathbf{y} = \mathbf{w}^T\mathbf{z}$ has maximal divergence from a Gaussian distribution.

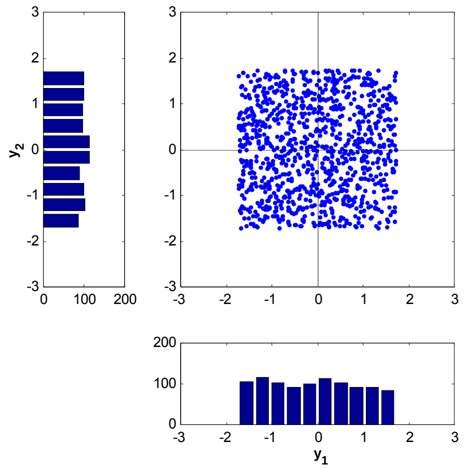

Figure 3: Plot of the estimated independent components, $\mathbf{y} = [y_1, y_2]$. The probability distributions approximate a uniform distribution.

2.3. Estimating the Independent Components

As mentioned previously, the objective of ICA is to find projections which yield components that are as independent as possible. It has furthermore been illustrated that, as inferred from the central limit theorem, this objective is equivalent to finding the directions of maximum divergence from a Gaussian distribution. ICA can thus be thought of as consisting of two basic elements:

1) Identifying a measure of divergence from a Gaussian distribution of the projection $\mathbf{w}^T\mathbf{z}$, often referred to as an objective function or contrast function; and
2) Finding an algorithm that optimises this divergence from a Gaussian distribution.

While a variety of measures of divergence from a Gaussian distribution are available, this study will focus on a method that uses a quantity known as negentropy (Comon, 1994, Hyvarinen *et al*, 2001), and is based on the information theoretic result that a Gaussian variable has the largest entropy of all random variables of equal variance. Negentropy J for a random variable **y** is defined as

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}), \tag{4}$$

where $H(\mathbf{y})$ is the differential entropy of **y** and $H(\mathbf{y}_{gauss})$ is the entropy of a Gaussian variable $\mathbf{y}_{gauss}$ of the same covariance matrix as **y** (Comon, 1994). The benefit of this quantity is that it is always nonnegative, and zero only when **y** has a Gaussian distribution (Hyvarinen *et al*, 2001). The problem, however, is that using negentropy is computationally difficult, as it requires the estimate of the probability density function of **y**. An approximation of negentropy can be used instead, which is given as:

$$J(\mathbf{y}) \approx [E\{G(\mathbf{y})\} - E\{G(\mathbf{y}_{gauss})\}]^2 \tag{5}$$

where G is an appropriately chosen nonlinear function often referred to as a contrast function, and the second term in the parentheses is a normalisation constant that makes the negentropy J equal to zero if **y** has a Gaussian distribution. It has been shown that G can be almost any non-quadratic, well behaving even function (Hyvarinen and Oja, 1997). In the present case, we use:

$$G(\mathbf{y}) = \log\cosh(\mathbf{y}) \tag{6}$$

as this is regarded as a good general-purpose contrast function due to its convergence properties and robustness against outliers. Since the second term in Equation 5 is constant, maximising the divergence from a Gaussian distribution for a projection $y = \mathbf{w}^T\mathbf{z}$ can be achieved by looking at the extrema of the contrast function $E\{G(y)\} = E\{G(\mathbf{w}^T\mathbf{z})\}$ over the unit sphere $\|\mathbf{w}\|$. The FastICA algorithm (Hyvarinen and Oja, 1997) is an efficient method for finding this extrema, with further details found in Hyvarinen et al (2001).

3. APPLICATION TO AUSTRLIAN RAINFALL

The analysis described in the earlier section was conducted on a synthetic data set in which the independent components were already known. In reality, the objective of ICA is to find these independent components so that the dataset is represented in a way that is easy to analyse and interpret. We now apply ICA to a dataset containing time series of Australian rainfall, and examine the relationship between the extracted ICs and the source signals we expect would have the maximum influence on the rainfall characteristics in Australia. Based on recommendations in the literature (e.g. Chiew *et al*, 1998, and references therein), the dominant source of variability in Australian rainfall is the El Niño Southern Oscillation (ENSO). In addition, a second source of variability is the modulation of the ENSO at a decadal or longer time scale (Zhang *et al*, 1997).

Our analysis seeks to identify the dominant independent components of seasonal rainfall data from a number of point locations spread over Australia, with the presumption that these ICs will be related to the two main sources of rainfall variability – ENSO and inter-decadal climate fluctuations. We use the Southern Oscillation Index (SOI) as an indicator of the ENSO, and an index describing the strength of the Interdecadal Pacific Oscillation (IPO) as an indicator of the inter-decadal signal outlined above. Our hypothesis is that if ICA leads to components that are more strongly related to the underlying source signals that influence Australian rainfall, the relationship between the ICs and above mentioned source signals (SOI and IPO) should be stronger than what would result if PCA were used. This hypothesis is based on the theoretical advantages of ICA over PCA in extracting physically meaningful signals from mixed data, which was demonstrated in the example described above.

3.1. Data

The rainfall data used in this analysis is based on a set of high quality rain gauges located throughout Australia that were identified by Lavery *et al.* (1997). For the purpose of this study, only those locations that contain records between 1921 and 2000 were used. In regions where the records were sparse, some infilling of data was undertaken using near-by rain gauges, so that the final time series consisted of less than 1 % missing data. In total, rainfall time series from 201 gauging stations were used, the locations of which are shown in Figure 4. We used seasonal data for the analysis, where the seasons were defined as autumn (MAM), winter (JJA), spring (SON) and summer (DJF).

The SOI data set was obtained from International Research Institute for Climate Prediction (IRI) / Lamont Doherty Earth Observatory (LDEO) climate data library (http://iri.columbia.edu) and is defined as the difference between the standardised Tahiti sea level pressure (SLP) and the standardised Darwin SLP. As with the rainfall data, seasonal average SOI values were used. Only concurrent relationships were examined.

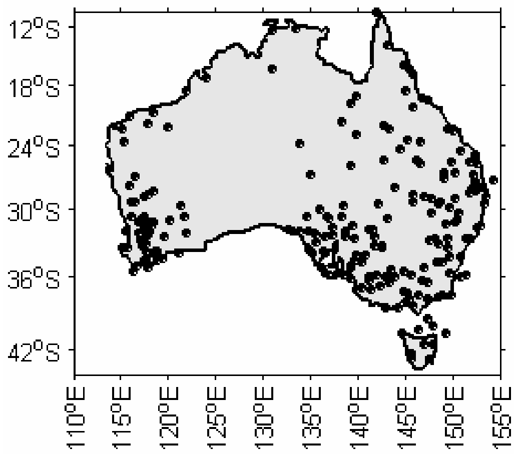

Figure 4: Location of rain gauge sites.

3.2. Results

To establish the benefits of ICA over PCA in representing Australian rainfall time series in a manner that is easy to interpret, we examine the correlation of the extracted components with the SOI, since a successful analysis of the rainfall data should result in at least one component that represents the contribution of the ENSO phenomenon on Australian rainfall. Specifically, we examine the following cases:

1. We compare the PCs and ICs extracted from the Australian rainfall data set from 1921 to 2000 with concurrent SOI data.
2. We then examine the links between the PCs and ICs of Australian rainfall with the SOI during an IPO negative phase, which spans from 1946 to 1977. Rainfall during this period has been shown to be more closely correlated with ENSO than during the IPO positive phase (Power *et al*, 1999, Verdon *et al*, 2004), and this should be reflected in the ICA results.

For each case we present correlation results of the PC or IC that yields the maximum correlation coefficient with the SOI. Based on the conclusions from a synthetic study relating to the number of ICs that can be extracted from a given length of data (results not shown), we extract 2 ICs both for the data set from 1921 to 2000 (80 data points), and for the data set from 1946 to 1977 (32 data points). Before using the ICA technique, we apply a PCA reduction to the data so that the number of PCs was the same as the number of ICs. In a separate analysis, we also retained an additional PC, and found the conclusions to be very similar.

The correlation coefficients between the extracted components (PCs and ICs) and the SOI are presented in Table 1 for each season of the year. In each case we only present the absolute value of the correlation coefficient, since ICs can only be determined up to a scalar multiple so that the sign of the correlation coefficient is meaningless. Correlation coefficients greater than 0.22 and 0.35 are statistically significant at the 95 percent level for the 80 year and 32 year datasets, respectively.

Comparison of the ICA results with the PCA results shows that in all cases, the ICs display a greater level of correlation with the SOI compared with the PCs. In fact, correlation coefficients for all the ICs are statistically significant at the 95 percent level, whereas correlation coefficients for all the PCs (except for spring from 1921 to 2000) are not statistically significant.

Furthermore, a comparison between the 1921 to 2000 data and the 1946 to 1977 data sets show a substantially improved performance for the shorter, IPO negative data set, and this is consistent with the literature, which suggests that the influence of ENSO is enhanced during the IPO negative phase (Power *et al*, 1999; Verdon *et al*, 2004). The consensus of these results therefore suggest that the manner in which ICA represents the data is more in line with finding the sources of variability of Australian rainfall than with PCA.

	Autumn	Winter	Spring	Summer
PCA 1921-2000	0.19	0.22	0.28	0.17
ICA 1921-2000	0.25	0.41	0.51	0.39
PCA 1946-1977	0.25	0.33	0.30	0.24
ICA 1946-1977	0.44	0.53	0.66	0.64

Table 1: The absolute value of the correlation coefficient between the PCs and the SOI, and the ICs and the SOI, for an 80 year period from 1921 to 2000, and a shorter 32 year period from 1946 to 1977 corresponding to the IPO negative phase. Statistically significant correlation coefficients at the 95th percentile confidence level were 0.22 and 0.35 for the 1921 to 2000 and the 1946 to 1977 datasets, respectively. The results show that, for the majority of PCA results, the relationship between the PCs and the SOI is not statistically significant, while for each of the ICA results, the relationship is clearly statistically significant. Finally, a marked improvement in performance is observed during the IPO negative time period.

To provide a reference point for the ICA and PCA results, we also compare the correlation coefficients from each of the individual 201 rainfall time series with the SOI. Due to the large number of results, we sort the correlation coefficients from smallest to largest, and only consider the 50th and the 95th percentile correlation coefficients. These results are presented in Table 2.

A comparison of the results in Tables 1 and 2 indicates that, whereas the correlation coefficients between the PCs and the SOI are similar to the median (50th percentile) correlation coefficients, the ICA results are closer to the 95th percentile results. This should not be surprising, since the objective of PCA is to find a transformation that faithfully represents the variance of the original data, where ICA aims to find a transformation that facilitates the interpretation of the dataset.

	Autumn	Winter	Spring	Summer
50 percentile 1921-2000	0.16	0.24	0.25	0.18
95 percentile 1921-2000	0.35	0.44	0.47	0.33
50 percentile 1946-1977	0.28	0.25	0.41	0.23
95 percentile 1946-1977	0.54	0.58	0.63	0.50

Table 2: Correlation between individual rain gauges and the SOI. Correlation coefficients for each of the 201 rainfall time series were computed, and the 50 percentile and 95 percentile results are shown. Statistically significant correlation coefficients at the 95% confidence level were 0.22 for the data set from 1921 to 2000, and 0.35 for the data set from 1946 to 1977.

4. CONCLUSIONS

The results presented herein indicate that PCA and ICA represent fundamentally different solutions to the mixing problem. Mathematically we can summarise these differences as follows:

1) PCA is able to describe maximum variance of the data set (in a least squares sense), whereas under certain conditions ICA is able to separate distinct signals from mixed data;
2) PCA constrains the solution to an orthogonal transformation of the original data set, resulting in difficulties interpreting components beyond the first PC. This constraint does not apply to ICA;

3) PCA uses second-order statistics resulting in components that are mutually uncorrelated, whereas ICA is also able to reduce higher-order dependencies to find components that are statistically mutually independent.

Therefore, while PCA may be beneficial for dimension reduction, ICA should in theory provide results with a much greater degree of interpretability. In Section 2 of this paper we illustrated this concept using a simple example in which two uniform signals were linearly mixed, demonstrating that ICA was clearly successful in recovering these signals whereas PCA was not.

Application of PCA and ICA to an Australian rainfall dataset showed that whereas the PCs were unable to resolve the influence of ENSO on Australian rainfall, at least one of the ICs from each analysis were found to be correlated with the SOI at or above the 95 percent significance level for each season. This result was further enhanced during the IPO negative phase from 1946 to 1977, where correlation coefficients between one of the ICs and the SOI were found to be as high as 0.66 and 0.64 for the spring and summer time series, respectively, illustrating that ICA is capable of identifying at least some of the dominant physical processes that give rise to rainfall in Australia.

These results therefore suggest that ICA has some significant advantages over PCA in analysing climatic time series such as rainfall, particularly with regards to the interpretability of the extracted signals. Areas for future research include exploring extensions to the simple ICA model presented here, such as developing models that account for noise, non-stationary mixing matrixes or nonlinear mixing. The application of ICA to alternative climate datasets, such as SSTs, is also an avenue worth pursuing. Finally, the performance of ICA should be tested as a basis for developing statistical forecasts for rainfall or other hydro-climatic variables.

5. ACKNOWLEDGMENTS

Funding for this research came from the Australian Research Council and the Sydney Catchment Authority. Their support for this work is gratefully acknowledged.

6. REFERENCES

Chiew F.H.S., Piechota, T.D., Dracup, J.A. & McMahon, T.A., El Nino / Southern Oscillation and Australian Rainfall, Streamflow and Drought: Links and Potential for Forecasting, *Journal of Hydrology, 204*, 138-149, 1998.

Comon, P., Independent Component Analysis, A New Concept?*, Signal Processing, 36,* 287-314, 1994.

Herault, J., and C. Jutten, Space or time adaptive signal processing by neural network models, in *Neural networks for computing: AIP conference proceedings 151,* edited by J.S. Denker, American Institute for physics, New York, 1986

Hyvarinen, J. and E. Oja, A fast fixed-point algorithm for independent component analysis, *Neural Computation, 9* (7), 1483-1492, 1997.

Hyvarinen, A., J. Karhunen. and E. Oja, *Independent Component Analysis,* 481 pp., John Wiley and Sons, New York, 2001.

Lavery, B., G. Joung, and N. Nicholls, An extended high-quality historical rainfall dataset for Australia, *Australian Meteorological Magazine*, *46*, 27-38, 1997.

Oja, E., Applications of Independent Component Analysis, *Lecture Notes in Computer Science, 3316,* 2004.

Power, S., F. Tseitkin, V. Mehta, B. Lavery, S Torok and N. Holbrook, Decadal climate variability in Australia during the twentieth century, *International Journal of Climatology, 192,* 169-184, 1999.

Richman, M.B., Rotation of principal components, *Journal of Climatology, 6*, 293-335, 1986.

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

Thompson, D.W.J. and Wallace, J.M., The arctic oscillation signature in the wintertime geopotential height and temperature fields, *Geophysical Research Letters*, 25, 1297-1300, 1998.

Wilks, D.S., *Statistical Methods in Atmospheric Sciences*, Academic Press, San Diego, 1995.

Verdon, D.C., Wyatt, A.M., Kiem, A.S. & Franks, S.W., Multidecadal variability of rainfall and streamflow: Eastern Australia, *Water Resources Research, 40*, 2004.

Zhang, Y., J.M. Wallace, and D.S. Battisti, ENSO-like interdecadal variability: 1900-93, *Journal of Climate, 10* (5), 1004-1020, 1997.

Zwiers, F.W. and Von Storch, H., On the role of statistics in climate research, *International Journal of Climatology, 24,* 2004.