

SST phases and Australian rainfall

Wasył Drosdowsky

Bureau of Meteorology Research Centre, Australia

(Manuscript received November 2000; revised November 2001)

Two analogue or phase selection methods are regularly examined, but not officially included, in the Seasonal Climate Outlook (SCO) issued by the National Climate Centre; the non-linear Southern Oscillation Index (SOI) analogue developed by Drosdowsky (1994), and the SOI phase system of Stone and Auliciems (1992). An SST phase classification system, similar to that of Stone and Auliciems has been developed based on the two leading modes of Indian / Pacific Ocean variability. The hindcast skill of these three systems in predicting seasonal rainfall with a one month lead is examined using a cross-validation methodology, and compared with the current operational discriminant analysis forecast system. All systems have similar levels of hindcast skill and could potentially be combined to produce a system with a higher level of skill than that of any individual system.

Introduction

Each month the National Climate Centre (NCC) of the Australian Bureau of Meteorology issues a Seasonal Climate Outlook (SCO) for the following three-month season. This outlook, in the form of the probability of exceeding various thresholds for rainfall and maximum and minimum temperatures, is derived using discriminant analysis with large-scale patterns of sea-surface temperature as predictors (Drosdowsky and Chambers 1998; Jones 1998). An alternative seasonal forecast, based on phases of the Southern Oscillation (Stone and Auliciems 1992), is also widely available from the Queensland Department of Primary Industry (QDPI) via the internet and the commercial RAINMAN software package. This phase system is popular with users who can relate to conditions in the analogue years with similar phase to the current year. Since these forecasts are based on different predictors and different statistical methodologies, they occasionally give conflicting results. This paper describes an attempt to reconcile these differences by developing a phase system based

on the SST predictors. An additional SOI analogue system, developed by Drosdowsky (1994) and based on non-linear dynamical systems theory, is also examined during the formulation of the SCO.

To compare the hindcast skill of these systems a single rainfall dataset which covers the 100-year period from 1900 to 1999 with monthly temporal resolution is used. This dataset was created by the National Climate Centre (Jones and Weymouth 1997) by analysing all the available quality controlled station data onto a one-quarter degree grid using a successive correction scheme. A reduced resolution subset (at one degree spatial resolution) is used in this study.

Two versions of the Global Ice and Sea Surface Temperature data set (GISST) developed by the UK Meteorological Office are used in this study; Version 1.1; (Parker et al. 1995), available from 1949 to 1991, and Version 3.1 (Rayner et al. 2000) from 1871 to 1997. More recent data is obtained from the Bureau of Meteorology operational SST analyses (Smith 1995).

A detailed description of the SST phase system, and a comparison with the SOI phases, is presented in the next section. The forecast methodology and the evaluation of hindcast skill are then outlined in the following section, and finally, the conclusions and discussion of the results are presented.

Corresponding author address: Wasył Drosdowsky, Bureau of Meteorology Research Centre, GPO Box 1289K, Melbourne, Vic. 3001, Australia.
Email: W.Drosdowsky@bom.gov.au

Description and comparison of the three systems

SOI phases (Stone and Auliciems 1992)

In this scheme the state of the climate system is defined in terms of the past two months' SOI values. The phases are defined through a Principal Component analysis of these two values and their difference. The first two components, which represent the mean value and the month-to-month trend, are used to place each month into one of five possible clusters or phases describing the state of the SOI over the current and previous month. These phases are:

- phase 1: consistently negative SOI;
- phase 2: consistently positive SOI;
- phase 3: rapidly falling SOI;
- phase 4 rapidly rising SOI; and
- phase 5: near zero SOI.

Phases for each month were extracted from Stone and Auliciems (1992), and updated from the QDPI web page (<http://www.dnr.qld.gov.au/longpdk/lpphases.htm>). The relative frequency distribution and conditional transition probabilities for these phases (Table 1) show that (a) the clusters are not of equal size, with phase 5 being the most common, and (b) certain transitions are more likely, while two (phase 1 to or from phase 2) are explicitly not possible. Transitions from phase 3 (rapidly falling) to phase 2 (consistently positive), and from phase 4 (rapidly rising) to phase 1 (consistently negative) are theoretically possible, but have not been observed to date. Phases 1, 2 and 5 (near zero) show a high degree of persistence, while as expected phases 3 and 4 are transitional. Phase 3 is almost equally likely to be followed by phase 1 or 4, while phase 4 leads to phases 2, 3 or 5.

Non-linear analogues (Drosowsky 1994)

This system is similar to the Stone and Auliciems system in that it uses the mean value and trend in the SOI, but with the difference that the mean value and trend over a nine-month rather than two-month peri-

od are used. Each month effectively defines its own unique phase consisting of its ten closest analogues. These ten analogues are presented in the SCO, but not used explicitly in the rainfall or temperature forecasts.

SST phases

This system, developed and described for the first time in this paper, is also conceptually similar to the SOI phases of Stone and Auliciems, the major difference being that rather than using a lagged value of the same index, two different indices of the large-scale climate state are used. These indices are the time series of the first two rotated Principal Components (SST1 and SST2; Fig. 1) of the Indian/Pacific Ocean sea-surface temperature anomaly pattern described by Drosowsky and Chambers (1998). These patterns were derived from the GISST1.1 dataset for the period 1949-91. To extend the period of record the GISST3 data for 1871 to 1997 and the Bureau of Meteorology operational analyses since January 1998 are projected onto the GISST1.1 patterns. Both these time series (thin curves in Fig. 2) show multi-decadal variability over the 129-year record, with SST2 also warming significantly over the past 50 years. This low frequency variability, represented by the first twelve Fourier components (heavy curves in Fig. 2), is removed from both time series before the SST phases are defined.

The SST classes or phases are defined by simply assigning each monthly value of each (filtered) time series into one of three equally probable classes (terces); the lowest third or below normal values (terce 1), the middle third or near normal (terce 2) and the highest third or above normal values (terce 3). The terciles (the values that separate the terces) are calculated from the 100-year period 1900-99 as this coincides with the period of availability of Australian gridded rainfall data. The first 29 years of the dataset prior to 1900 are then classified based on these tercile values. If the raw time series are used, the SST2 time series (being anomalies from the 1961-90 mean) is biased towards negative values, and most of the terce

Table 1. Empirical probabilities (based on 1200 months data from 1900-1999) of each SOI phase, and the conditional probability of transition from SOI phase A in current month to SOI phase B in next month. Bold figures show most preferred transitions with probability greater than 0.3

Current phase (A)	Prob(A)	B=1	B=2	B=3	B=4	B=5
1 -	0.169	0.660	0.000	0.089	0.143	0.108
2 +	0.231	0.000	0.639	0.130	0.069	0.159
3 \	0.131	0.376	0.000	0.070	0.357	0.197
4 /	0.190	0.000	0.395	0.202	0.114	0.289
5 ~0	0.279	0.030	0.030	0.137	0.293	0.510

Fig. 1 Spatial pattern of loadings of the first two rotated principal components of the standardised monthly anomalies of the GISST1.1 dataset for the 1949-1991 period. Contour interval is 0.2, with zero contour heavy, negative contours dashed and areas above +0.2 and below -0.2 shaded.

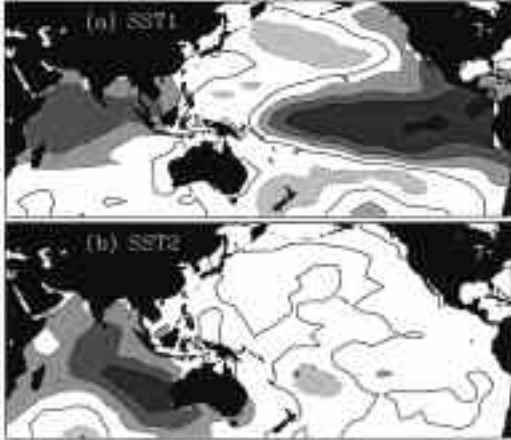
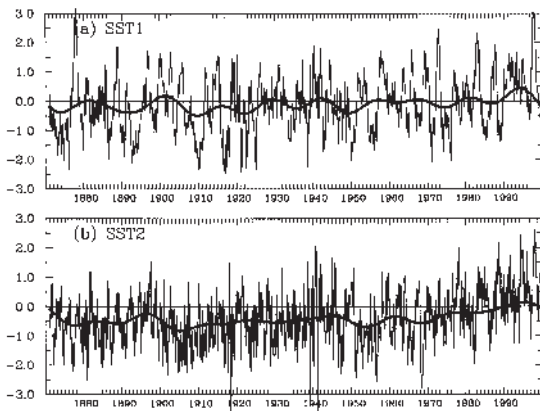


Fig. 2 Time series of the first two principal components, projected onto the GISST3 dataset for the period 1871 to 1997 and Bureau of Meteorology SST analyses for 1998-9. Each panel shows the original data (thin line) and the low frequency component, defined as the first 12 Fourier components calculated from the entire 129 year record (heavy line). These components consist of variability on time scales greater than 10.5 years.



1 values are found in the early part of the record. This type of classification, based on percentiles rather than standardised departures from mean values, is less sensitive to (perhaps erroneous) outliers, such as that observed in the SST2 time series in 1918.

Combining the two time series results in a nine (three by three) way classification, summarised in Table 2 which gives the tercile boundaries for each of the two SST components, the number of cases in each phase and the mean value of the two components in each phase for the 100-year (1200 month) record from 1900-99. The composite SST anomaly patterns based on 98 years of GISST3 data for 1900 to 1997 are shown in Fig. 3. While these maps show an overall cold bias, since the anomalies are calculated relative to the 1961-90 climatology, the major anomalies are found in the regions with high loadings on SST1 and SST2, i.e. the equatorial Pacific and the subtropical Indian Ocean.

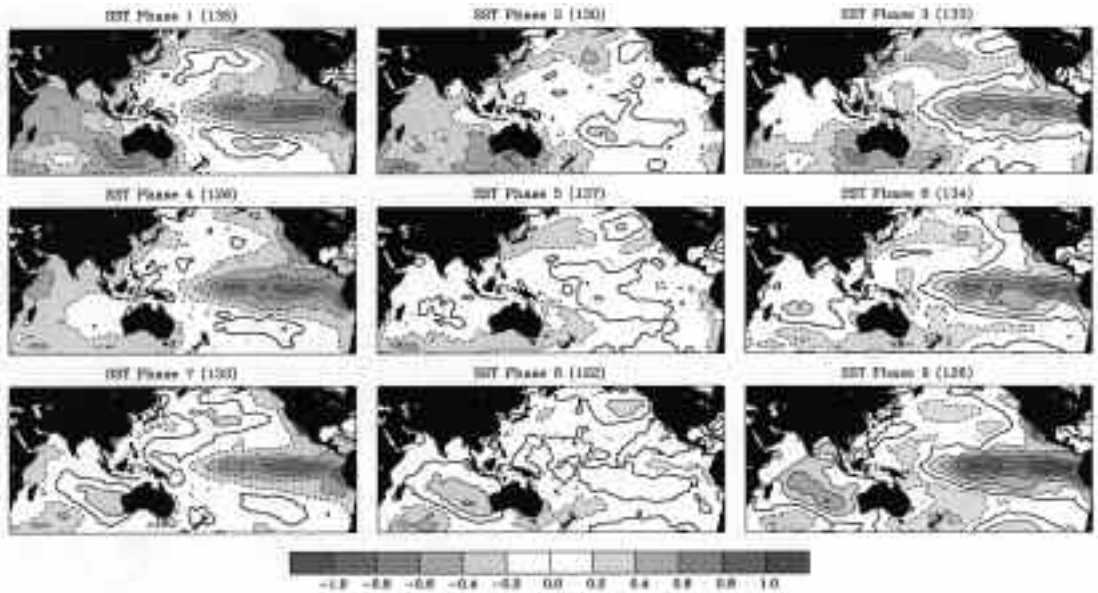
Since the two indices are largely uncorrelated, each of the nine phases is nearly equally likely; i.e. if SST1 is above normal then there is equal likelihood of SST2 being above, near or below normal, and similarly for the other SST1 classes. However, in both the SOI and SST phase systems the classification has been based on all months taken together, and there is an uneven number of occurrences of each phase in different months. The classification of each month from January 1871 to September 2000 is given in Table 3.

Table 4 gives the conditional probability of transitions between the nine phases. Although simultaneous transitions of both indices from one extreme class to the other (phase 1 to 9, or phase 3 to 7, and the reverse in each case) are not prohibited, they have not been observed in the 129-year record. Higher transition probabilities are found between phases with the same SST1 state than with the same SST2 state, and many of the possible transitions between extreme states of SST1 are also not found (e.g. phase 7 to 9 and the reverse). The typical evolution of El Niño events (extended periods of SST1 above normal) as documented by Rasmussen and Carpenter (1982), Harrison and Larkin (1998) and Reason et al. (2000) include a warming of the equatorial Indian Ocean. Although the SST2 index is mainly measuring subtropical eastern Indian Ocean SSTa, it does extend into the tropics and evolves, during many El Niño events, from below normal (phase 3) through near normal (phase 6) to above normal (phase 9), as shown by the bold text in Table 3. This evolution, which is clearly seen in the three right panels of Fig. 3, is consistent with a significant correlation of 0.29 between SST1 and SST2 lagged by 6 to 12 months, although this is reduced slightly to 0.23 when the low frequency component is removed from both time series.

Table 2. Definition of the 9 SST phases in terms of three equally likely categories for each of SST1 and SST2. Table entries give the number of occurrences of each phase in the 100-year record from 1900 to 1999 (expected value of 1200/9 ~ 133), and mean value of SST1 and SST2 for each phase. Phases 1, 4 and 7 represent La Niña events (Cold Pacific), while phases 3, 6 and 9 represent El Niño (Warm Pacific) events.

<i>Indian Ocean (SST2)</i>	<i>Pacific Ocean (SST1)</i>		
	<i>Cold SST1 < -0.394</i>	<i>Neutral -0.394 < SST1 < 0.414</i>	<i>Warm SST1 > 0.414</i>
Cold SST2 < -0.338	phase 1 135 -0.884, -0.909	phase 2 132 0.016, -0.822	phase 3 133 0.911, -0.951
Neutral -0.338 < SST2 < 0.367	phase 4 126 -0.922, 0.018	phase 5 140 0.043, 0.023	phase 6 134 1.032, 0.028
Warm SST2 > 0.367	phase 7 139 -0.920, 0.943	phase 8 127 -0.011, 0.919	phase 9 134 1.124, 0.955

Fig. 3 Composite SST anomalies for each of the nine SST phases using 98 years of data from the GISST3 dataset. Anomalies are relative to a 1961-90 base period. Contours at 0.2°C intervals, zero contour heavy and negative contours dashed.



Comparison of SOI and SST phases

As noted above, the number of cases in each of the nine categories is nearly constant since the two time series are largely uncorrelated at zero lag. In contrast a similar procedure applied to two lagged values of the SOI or either of the SST components alone essentially produces a five category classification similar to that of Stone and Auliciems. Since there is a strong lag-one autocorrelation in all these time series, a scatter plot of current month against past month will show most of the data points falling into the three classes

with both months in the same terce, giving the consistently positive, consistently negative and near-zero phases. The remaining off-diagonal points (high in past month, low in current month and the reverse) will define the two rapidly rising or falling phases.

As both these systems are based in part on the El Niño/Southern Oscillation (ENSO) cycle, there should be some correspondence between the classification of each month. Table 5 shows the joint distribution of the two classifications, expressed as the number of occurrences in the common 118-year (1416

Table 3. SST phase for each month from January 1871 to September 2000. Major El Niño events, defined as extended nearly continuous periods of phases 3, 6 or 9 are in bold. Note the evolution from phase 3 to phase 6 to phase 9 found in many of these events (e.g. 1982/3 and 1997/8)

YEAR	J	F	M	A	M	J	J	A	S	O	N	D
1871	4	2	2	6	5	7	7	7	7	4	5	4
1872	2	2	2	1	1	2	1	1	1	4	7	4
1873	7	7	7	4	4	7	4	1	5	5	8	8
1874	7	7	7	4	1	1	1	1	1	1	1	4
1875	4	4	4	4	5	2	4	4	4	5	7	4
1876	7	7	4	7	7	2	3	5	2	2	2	3
1877	2	3	3	3	3	3	3	3	6	6	9	6
1878	9	9	9	9	9	6	9	5	5	8	8	8
1879	5	1	1	1	1	1	4	4	1	1	4	1
1880	1	4	8	5	5	2	4	8	5	3	6	9
1881	2	3	6	6	6	2	8	8	7	7	7	4
1882	4	7	5	8	8	1	4	4	5	2	1	4
1883	4	1	5	5	5	8	7	5	4	5	2	1
1884	1	3	3	3	2	4	2	5	5	6	3	3
1885	3	2	2	5	5	1	4	5	2	3	6	6
1886	5	2	5	8	7	7	7	7	4	1	4	1
1887	1	1	1	1	1	1	5	4	2	2	5	5
1888	2	6	9	9	9	9	9	9	6	9	9	9
1889	6	9	9	6	6	5	5	4	4	1	1	4
1890	1	4	4	1	4	4	1	4	1	1	1	1
1891	2	3	3	6	6	6	6	8	8	8	8	8
1892	8	8	5	5	8	7	7	7	4	4	7	4
1893	1	1	4	1	1	1	4	4	1	1	1	4
1894	4	1	4	7	7	7	7	7	7	7	7	5
1895	2	2	5	2	2	2	5	5	3	9	9	9
1896	8	5	6	6	6	5	6	9	6	9	9	6
1897	6	6	9	9	8	8	8	8	5	2	2	2
1898	2	4	1	1	4	7	4	4	4	8	7	7
1899	4	4	7	2	2	4	2	3	3	3	3	6
1900	9	9	9	6	6	6	3	3	2	3	8	9
1901	8	5	2	2	2	2	1	4	1	2	2	2
1902	1	1	1	5	5	9	6	9	9	9	9	9
1903	6	6	3	3	5	7	4	7	4	4	4	4
1904	1	4	4	7	4	4	2	5	2	3	3	2
1905	3	3	9	2	3	3	3	3	3	3	3	6
1906	3	6	3	2	8	8	7	4	1	1	1	1
1907	1	7	1	8	8	5	5	7	5	2	2	6
1908	3	2	2	2	4	2	1	1	1	2	1	4
1909	1	4	1	4	1	4	1	4	4	1	4	7
1910	7	4	7	7	1	1	1	4	1	1	4	1
1911	7	4	4	1	4	1	2	3	3	3	6	3
1912	6	6	3	9	9	8	8	8	2	2	5	7
1913	4	5	2	1	4	7	7	4	5	3	6	3
1914	3	3	3	3	3	6	6	9	6	9	9	3
1915	3	6	6	9	9	9	5	8	6	6	7	7
1916	1	1	7	7	1	1	4	7	4	1	1	1
1917	1	1	1	4	4	4	4	4	1	1	1	4
1918	7	7	7	4	2	3	3	5	8	6	3	3
1919	6	6	6	6	3	3	5	5	5	5	8	5
1920	9	6	9	8	8	5	5	6	6	9	8	4
1921	1	4	7	7	7	7	7	8	8	9	5	5
1922	5	5	2	3	2	4	4	4	1	7	4	1
1923	2	1	5	3	3	3	2	5	3	3	6	6
1924	2	5	5	7	4	1	2	1	1	4	7	7
1925	4	1	1	2	2	5	6	9	6	6	9	9
1926	9	6	9	6	6	6	6	8	5	8	5	5
1927	2	6	1	1	2	2	2	5	2	9	8	8
1928	8	8	8	5	7	7	8	5	2	2	8	2
1929	7	5	2	3	1	2	1	1	2	5	2	5
1930	2	3	2	8	8	7	2	5	3	3	6	6
1931	6	6	6	9	6	5	5	5	4	2	8	7
1932	8	5	6	6	2	5	6	2	3	2	2	2
1933	5	8	8	5	5	1	7	7	7	4	4	1
1934	4	4	1	2	5	8	2	8	5	5	5	7
1935	2	1	1	7	7	4	5	3	3	3	3	5
1936	3	5	5	3	6	2	5	7	5	6	6	5
1937	5	2	3	6	2	2	2	8	5	2	8	2

Table 3. Continued.

YEAR	J	F	M	A	M	J	J	A	S	O	N	D
1938	4	5	2	1	5	7	7	4	4	4	7	4
1939	4	5	1	3	3	5	5	8	2	2	8	3
1940	3	3	6	9	9	9	9	8	9	9	8	5
1941	5	1	3	3	3	4	7	8	5	9	9	6
1942	6	9	9	9	6	5	5	4	4	4	4	4
1943	1	1	1	1	2	2	8	8	2	2	8	7
1944	7	4	3	3	3	6	3	6	5	9	8	8
1945	8	2	2	1	1	8	8	8	5	5	8	8
1946	5	2	1	5	3	5	1	1	4	4	4	2
1947	3	3	3	2	3	3	2	1	1	1	2	5
1948	2	8	9	6	6	6	5	8	4	4	8	5
1949	8	9	8	9	6	7	8	7	7	7	7	7
1950	7	7	7	7	7	4	7	7	7	7	4	1
1951	1	1	1	2	2	2	3	3	6	6	6	3
1952	6	6	9	9	6	3	5	5	2	2	2	2
1953	2	6	6	6	3	3	6	6	3	5	6	6
1954	5	5	5	1	1	1	1	1	4	4	1	1
1955	7	4	4	4	1	1	1	1	1	1	1	1
1956	7	4	7	4	1	1	1	1	1	1	2	5
1957	5	5	6	3	6	3	6	3	3	3	6	9
1958	9	9	6	9	9	6	6	6	6	3	6	8
1959	9	9	9	6	9	8	7	7	8	8	8	5
1960	2	2	2	2	2	2	2	2	3	5	2	5
1961	8	8	5	3	2	2	2	5	5	8	8	5
1962	5	8	8	7	7	4	8	8	8	5	5	7
1963	8	2	6	5	6	3	3	3	6	9	9	9
1964	8	9	8	5	4	4	1	1	1	1	1	1
1965	4	4	2	2	2	2	6	3	3	3	3	3
1966	3	6	3	3	2	5	5	5	5	5	7	7
1967	7	7	8	7	7	7	7	7	7	7	7	7
1968	1	1	1	1	1	2	2	3	3	3	3	3
1969	6	6	9	6	6	3	6	8	6	9	9	9
1970	9	6	9	9	5	4	7	7	7	4	7	7
1971	4	4	4	1	4	1	4	4	1	1	1	1
1972	8	8	5	3	6	6	3	6	6	6	9	9
1973	9	6	9	6	5	2	5	4	4	4	5	1
1974	4	1	1	1	4	2	7	2	1	2	2	4
1975	4	7	7	5	8	7	7	4	4	4	4	1
1976	1	1	4	5	8	9	9	9	9	6	3	6
1977	6	9	6	9	5	8	9	8	8	8	9	9
1978	8	8	8	7	4	1	1	1	1	4	7	5
1979	8	5	6	3	2	5	8	8	5	8	8	8
1980	9	3	6	6	6	3	5	4	5	2	8	2
1981	4	1	5	4	4	4	4	4	4	8	1	3
1982	3	2	2	3	3	3	3	6	6	6	9	9
1983	9	6	9	9	9	9	9	9	9	8	8	8
1984	7	4	8	8	4	7	7	7	7	7	7	4
1985	4	4	7	7	7	7	7	7	7	4	8	5
1986	1	2	2	2	2	4	4	2	3	3	3	6
1987	3	3	6	6	6	9	9	9	9	6	9	3
1988	3	9	9	8	7	7	7	7	7	7	7	7
1989	7	7	7	7	7	1	7	7	8	5	8	8
1990	2	9	8	6	5	2	8	5	6	9	8	9
1991	9	3	6	9	9	9	9	9	8	9	3	9
1992	9	6	9	9	9	6	9	8	2	3	2	2
1993	5	3	9	6	6	5	5	5	5	5	8	5
1994	3	2	5	8	8	7	7	7	8	9	9	9
1995	9	9	9	8	8	6	9	8	8	5	5	5
1996	4	7	8	8	7	8	7	5	4	4	5	7
1997	4	2	3	3	6	6	6	6	9	9	9	9
1998	9	9	9	9	9	9	9	9	8	5	5	2
1999	2	5	8	8	7	7	7	8	8	7	7	7
2000	7	7	4	5	8	7	7	8	8	8	8	8

months) record from 1882 to 1999. As expected SOI phase 1 is mainly associated with SST phases 3, 6 and 9; the 'El Niño phases', while SOI phase 2 occurs with SST phases 1, 4 and 7, the 'La Niña' phases. The

other SOI phases are spread more uniformly across the SST phases, with a tendency for SOI phase 5 to occur with SST phases 2, 5 and 8 corresponding to neutral SST1 (Pacific) conditions.

Table 4. Conditional probability of transition from SST phase A in current month to SST phase B in next month (based on 1200 months data from 1900-1999) of each of the nine SST phases. Highest probabilities are found on the diagonal elements (bold) and represent the persistence of each phase. The strongest transitions between different phases (off diagonal elements) occur between phases with the same state of SST1, e.g. between phases 1 and 4, phases 3 and 6 etc. (italic).

	B=1	B=2	B=3	B=4	B=5	B=6	B=7	B=8	B=9
A=1	0.526	0.126	0.022	<i>0.200</i>	0.037	0.000	0.067	0.022	0.000
A=2	0.106	0.356	0.152	0.045	<i>0.174</i>	0.038	0.015	0.098	0.015
A=3	0.008	0.120	0.474	0.008	0.083	<i>0.263</i>	0.000	0.008	0.038
A=4	<i>0.270</i>	0.071	0.008	0.381	0.071	0.000	<i>0.159</i>	0.040	0.000
A=5	0.050	<i>0.164</i>	0.086	0.064	0.293	0.071	0.079	<i>0.164</i>	0.029
A=6	0.007	0.037	<i>0.209</i>	0.000	0.075	0.381	0.015	0.022	<i>0.261</i>
A=7	0.043	0.022	0.000	<i>0.230</i>	0.043	0.000	0.561	0.094	0.000
A=8	0.008	0.087	0.008	0.024	<i>0.244</i>	0.039	0.134	0.339	0.118
A=9	0.000	0.007	0.037	0.000	0.030	<i>0.209</i>	0.000	0.172	0.545

Table 5. Joint distribution of SOI and SST phases for 1416 months from January 1882 to December 1999. Conditional probabilities can be calculated by dividing each cell by the appropriate row (for Pr(SST | SOI)) or column (for Pr(SOI | SST)) total. Bold figures show the correspondence between SOI phase 1 (consistently negative) and SST phases 3, 6 and 9 (El Niño), SOI phase 2 (consistently positive) and SST phases 1, 4 and 7 (La Niña), and SOI phase 5 (near zero) and SST phases 2, 5 and 8 (SST1 neutral).

SOI	SST phase									Total
	1	2	3	4	5	6	7	8	9	
1	5	12	40	5	18	51	5	20	77	233
2	87	35	11	70	38	8	62	16	5	332
3	14	20	27	12	15	34	23	17	21	183
4	34	31	24	37	35	19	28	36	20	264
5	29	58	45	38	62	43	43	56	30	404
Total	169	156	147	162	168	155	161	145	153	1416

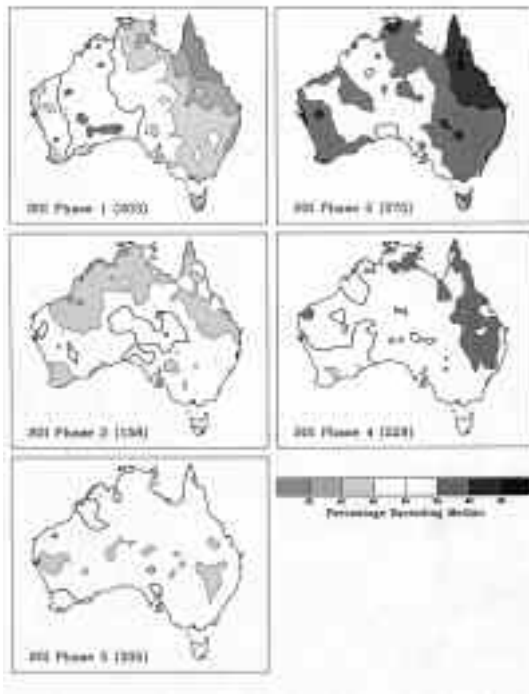
Forecast methodology

The conditional probabilities given in Tables 1 and 4 can be taken as forecast probabilities for the evolution of the climate system as represented by the SOI or the SST phases. Alternatively, forecasts of other parameters such as seasonal rainfall can be produced by examining the distribution of that parameter in the following season or seasons. The phases or analogues can be used to produce either categorical or probabilistic forecasts. Categorical forecasts can be generated by simply averaging across all analogues or members of a phase. This is essentially the compositing technique used in many diagnostic studies of climate variability (e.g. Rasmusson and Carpenter 1982). Probability forecasts can be produced in a similar manner to the conditional transition probabilities of the phases themselves, and regarding the observed

conditional empirical frequencies as forecast probabilities. This simply involves counting the number of members of the phase or class that fall into the forecast category, e.g. the number above or below the median. For the SOI or SST phases these will be unchanged from year to year, providing the same base set of data is always used. If the database is updated each year the forecasts will gradually alter over time as more cases or analogues become available. For the SOI analogues each forecast will be different, as the analogue years will almost certainly differ from year to year, and even from month to month.

The conditional probabilities for above median rainfall in the three-month season beginning the month after the observed phase for each of the five SOI phases are shown in Fig. 4, for each of the three categories of the two SST components separately in Fig. 5, and for the nine SST phases in Fig. 6.

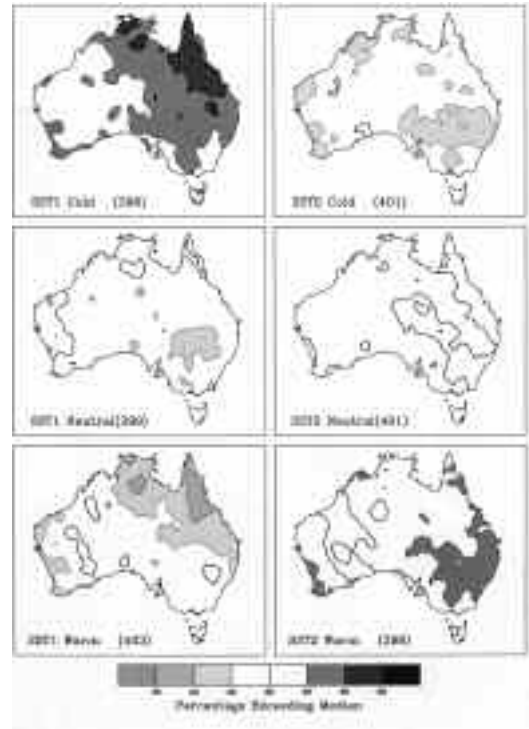
Fig. 4 Conditional probabilities (expressed as percentage) for above median rainfall in the season beginning the month after the observed phase for each of the five SOI phases. Numbers in brackets are the number of cases in each phase. Contour interval is 5%, with the 50% contour heavy and contours below 50% dashed.



The SOI phases (Fig. 4) show a reduced probability of exceeding the median following phases 1 and 3, especially over Queensland, and an enhanced probability following phases 2 and 4, with no significant shifts in phase 5. The Pacific Ocean SST1 (Fig. 5, left panels) shows a similar pattern to the SOI phases, with reduced (enhanced) probabilities associated with the warm (cold) phase. The Indian Ocean SST2 (Fig. 5, right panels) is partially consistent with the dipole pattern previously noted by Pittock (1975), Nicholls (1989) and Drosowsky (1993), with the cold (warm) phase associated with drier (wetter) conditions over the southeast of the continent, however, the expected wetter (drier) conditions over southern South Australia and western Victoria are not as evident.

When these two effects are combined, using the nine way two index classification, the resulting probabilities are as shown in Fig. 6. While the Pacific Ocean is clearly the dominant influence on Australian

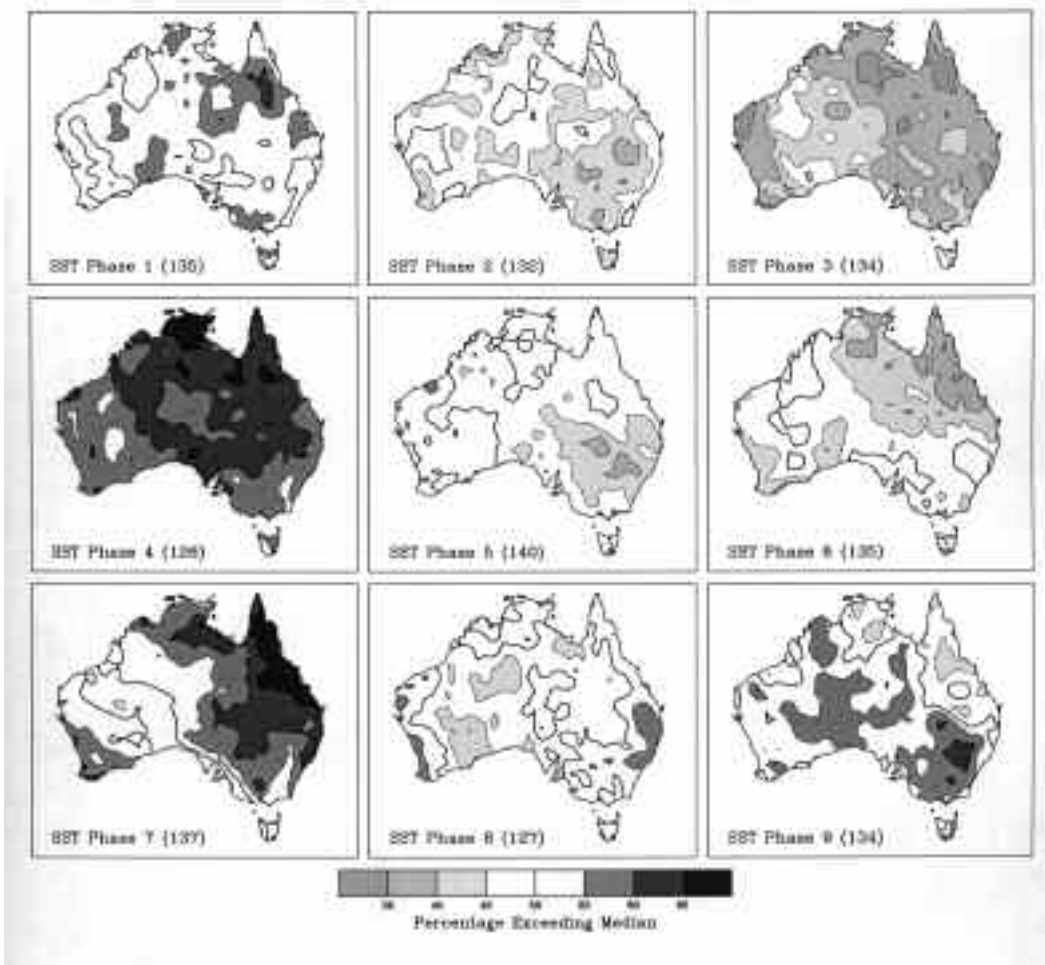
Fig. 5 As for Fig. 4, but for each of the tercile categories for SST1 (left-hand panels) and SST2 (right-hand panels).



rainfall, an important modulating effect due to the Indian Ocean is evident over southeast Australia. In particular, in phase 9, with both oceans in the warm phase, the wetter conditions over eastern Australia associated with the Indian Ocean dominate over the drier tendency of the Pacific Ocean, resulting in a very weak dry signal over northeast Australia and wetter conditions in the southeast.

The relationship between predictors such as the SOI or the SST principal components and Australian rainfall varies through the annual cycle, as shown by McBride and Nicholls (1983) and Drosowsky and Chambers (1998). The conditional probabilities therefore need to be calculated on seasonally stratified data. This results in 60 possible 'forecast' maps for the SOI phase system (12 seasons by 5 phases, the maps for zero-lag forecasts are available on the QDPI web site) and 108 for the SST system (12 seasons by 9 phases, not shown here due to space considera-

Fig. 6 As for Fig. 4, but for each of the nine SST phases.



tions). However, this stratification into individual seasons results in some season/phase combinations having very few members, and may lead to unreliable probability estimates.

The hindcast skill of this forecast process is estimated through a cross-validation process. An independent hindcast is produced for each of the 100 years in the dataset by repeated partitioning of the data into a 99-year training set, and the single year to be hindcast. For each partition the empirical frequency distribution based on the 99 training years is recalculated, and the status of the omitted year (i.e. above/below median) determined for subsequent verification. The hindcast for this year is calculated as described above, but based on the conditional probabilities derived from the 99-year training set. The hindcast skill is estimated using the LEPS methodology applied to probability forecasts, as described by Ward and Folland (1991) and Potts et al. (1996).

The resulting seasonally stratified cross-validated hindcast skill for each scheme is shown in Figs 7 (SOI phases), 8 (SOI Analogues) and 9 (SST phases), while the spatial averages of each system, and those of the operational discriminant analysis system are given in Table 6. Overall, no system is consistently superior to the others. All show a similar distribution of skill in both space and time, with highest levels of skill over eastern Australia during the spring to early summer season. The SST phases are perhaps the most consistent with positive spatially averaged skill in all seasons, particularly through the late summer and autumn period when the SOI systems perform less well.

Discussion and conclusions

A new method of generating probabilities of above/below median Australian seasonal rainfall,

Fig. 7 Cross-validated hindcast LEPS skill scores, based on 100 years from 1900-99, for the SOI phase system. Contour interval is 2.5, with zero contour heavy and negative contours not plotted. See Potts et al. (1996) for a description of the LEPS scoring system.

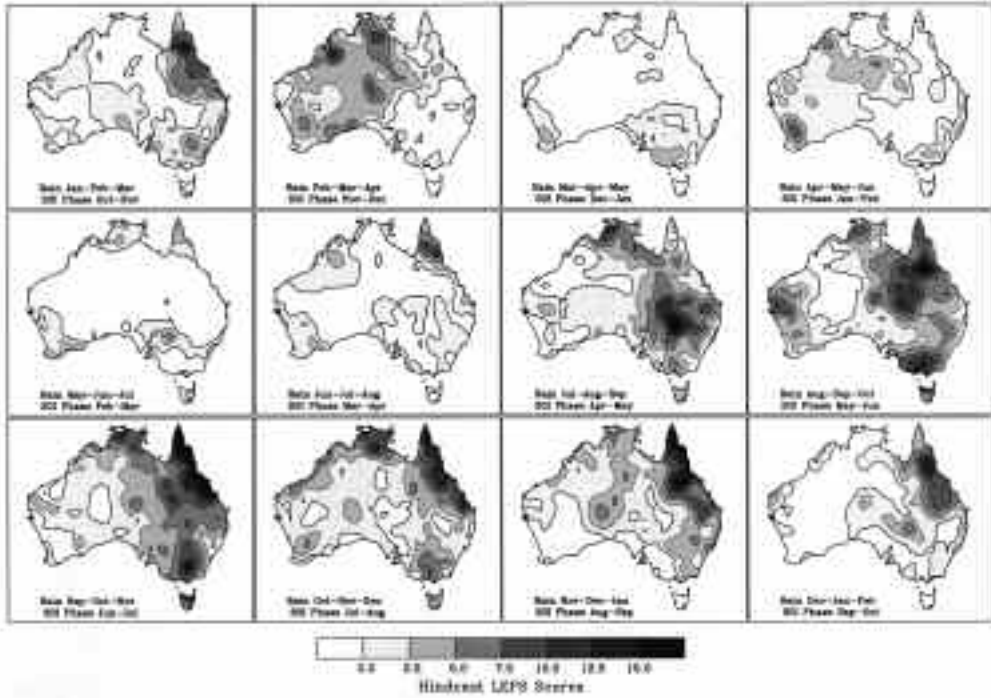


Fig. 8 As for Fig. 7, but for the non-linear SOI analogue system.

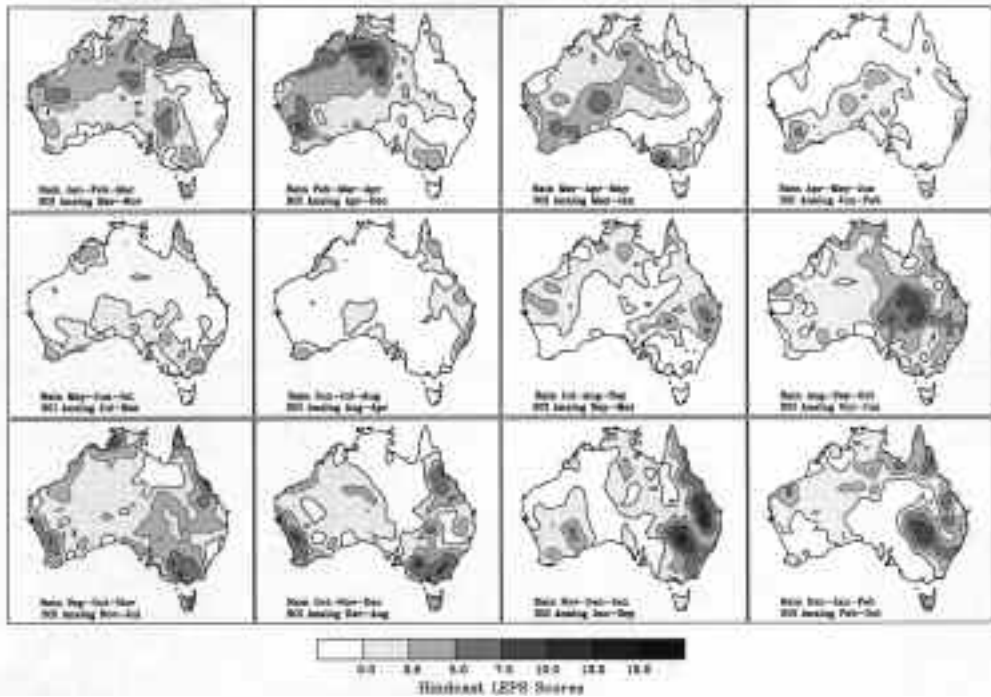
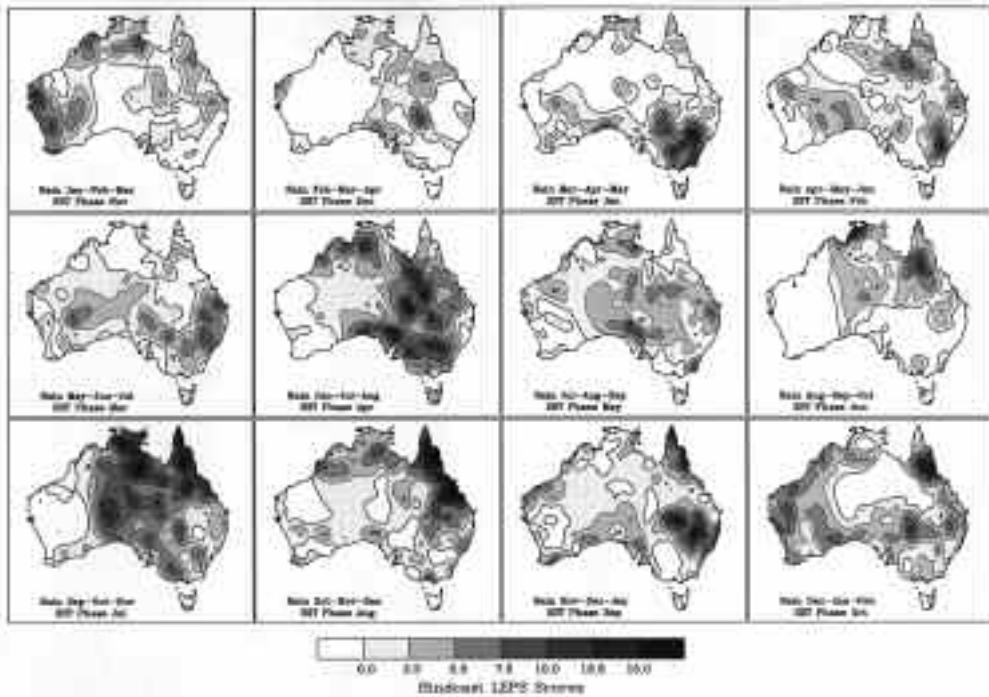


Fig. 9 As for Fig. 7, but for the SST phase system.



based on the state of Indian / Pacific Ocean SST anomalies has been documented. The classification based on this method has been compared to the SOI phases developed by Stone and Aulicciems (1992). The hindcast skill of both these systems have been calculated, as well as that of the SOI non-linear analogue technique developed by Drosowsky (1994), and all compared to that obtained by the discriminant analysis system currently used operationally by the NCC. All systems show similar levels and spatial patterns of hindcast skill.

A number of areas of possible concern regarding these analogue schemes, and their use as seasonal forecasts should be noted.

Low frequency variability and long term trends

The data on which the SST phases are based have been high-pass filtered to remove the low-frequency variability and long-term warming trend. Similarly there are weak long-term trends in the SOI, with more negative values observed in the recent past, especially during summer. A possible solution to this problem is to completely recalculate the phases as new data become available, although this will almost certainly result in changed classification for many months.

Sharp cluster boundaries

Although the SOI and SST datasets have continuous,

essentially normal distributions with no natural sub-groups, both the SOI and SST phase systems impose rigid boundaries between the phases. Small changes in the SOI or SST values can result in changes in phase, with possibly very different forecast probabilities. One solution to this problem may be the use of 'fuzzy' clusters, with points near the boundaries having a weighted membership in two or more phases. The same weighting can be used to estimate the contribution of each month to the empirical frequencies of exceeding the median rainfall as shown in Figs 5 and 6. Applying this procedure using a buffer zone of width 0.30 centred on the tercile points shown in Table 2, results in a slight improvement in the overall LEPS skill shown in Table 6. Marked increases in skill in the two seasons with lowest values (Feb-Mar-Apr and Aug-Sep-Oct), is offset by slight reductions in some of the seasons with highest skill.

Cross-validation of phase selection

In the evaluation of hindcast skill described in the 'forecast methodology' section, the analogue or phase selection mechanism itself was not cross-validated. In a rigorous application of the cross validation methodology for the two SOI based systems, the monthly SOI values would first need to be recalculated using all the data except those from the year to be forecast. (Note that the particular form of the SOI used by NCC

Table 6. Cross-validated hindcast LEPS scores for each of twelve overlapping three-month seasons, averaged over all independent grid-points, based on 100 years for SOI phases, SOI Analogues and SST phases, and 44 years for the operational Discriminant Analysis forecast system. (The 631 independent grid-points are those used by Drosowsky and Chambers (1998) in their principal component analysis of Australian rainfall anomalies)

Season	SOI phases	SOI Analogues	SST phases	SST Disc Anal
Jan - Feb - Mar	1.422	1.541	1.474	0.722
Feb - Mar - Apr	1.225	1.148	0.046	-0.657
Mar - Apr - May	-0.486	0.822	1.536	-0.040
Apr - May - Jun	0.293	-0.986	1.363	1.567
May - Jun - Jul	-0.509	-0.375	1.478	1.006
Jun - Jul - Aug	0.464	-0.857	3.720	2.590
Jul - Aug - Sep	2.933	0.524	1.776	1.539
Aug - Sep - Oct	4.764	2.168	0.296	2.742
Sep - Oct - Nov	5.147	2.637	4.252	2.195
Oct - Nov - Dec	3.086	1.757	2.941	3.436
Nov - Dec - Jan	2.496	1.874	2.948	2.591
Dec - Jan - Feb	1.125	1.284	3.460	1.754

is based on data for the 1933 to 1992 period.) In addition we are using the published results for the SOI phases, and it would not be feasible to recalculate these, due to some of the subjective selections that are made in the clustering procedure (Stone and Auliciems 1992). For the SST analogues the principal components themselves should also be recalculated. However this also involves some subjective decisions on the number of components to rotate. Cross validation experiments with a similar, seasonally stratified data set (Drosowsky and Chambers, 1998) show that the principal component spatial patterns are very stable and the projection of the missing data is very strongly correlated with the full data set analysis.

Conflicting forecasts

The existence of at least four, potentially different, seasonal forecasts could increase the level of confusion amongst users of the forecasts. Given that no system is clearly superior, an alternative may be to construct a simple linear combination of all the forecasts which may produce more skilful forecasts than any individual method.

References

- Drosowsky, W. 1993. Potential predictability of winter rainfall over southern and eastern Australia using Indian Ocean sea-surface temperature anomalies. *Aust. Met. Mag.*, 42, 1-6.
- Drosowsky, W. 1994. Analogue (non-linear) forecasts of the Southern Oscillation Index time series. *Weath. forecasting*, 9, 78-84.
- Drosowsky, W. and Chambers, L.E. 1998. Near global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *BMRC Research Report No. 65*, Bur. Met., Australia, 39 pp.
- Harrison, D.E. and Larkin, N.K. 1998. El Nino-Southern Oscillation sea surface temperature and wind anomalies, 1946-1993. *Rev. Geophys.* 36, 353-99.
- Jones, D.A. 1998. The prediction of Australian land surface temperatures using near global sea surface temperature patterns. *BMRC Research Report No. 70*, Bur. Met., Australia, 44 pp.
- Jones, D.A. and Weymouth, G. 1997. An Australian monthly rainfall dataset. *Technical Report No. 70*, Bur. Met., Australia, 19 pp.
- McBride, J.L. and Nicholls, N. 1983. Seasonal relationships between Australian rainfall and the Southern Oscillation. *Mon. Weath. Rev.*, 111, 1998-2004.
- Nicholls, N. 1989. Sea surface temperatures and Australian winter rainfall. *Jnl Climate*, 2, 965-73.
- Parker, D.E., Folland, C.K., Bevan, A.C., Ward, M.N., Jackson, M. and Maskell, K. 1995. Marine surface data for analysis of climatic fluctuations on interannual-to-century time scales. *Natural Climate Variability on Decadal-to-Century Time Scales*, National Research Council, 241-52.
- Pittock, A.B. 1975. Climatic change and the patterns of variation in Australian rainfall. *Search*, 6, 498-504.
- Potts, J.M., Folland, C.K., Jolliffe, I.T. and Sexton, D. 1996. Revised "LEPS" scores for assessing Climate Model simulations and long-range forecasts. *Jnl Climate*, 9, 34-53.
- Rasmusson, E.M. and Carpenter, T.H. 1982. Variations in tropical sea surface temperature and surface winds associated with the Southern Oscillation / El Nino. *Mon. Weath. Rev.*, 110, 354-84.
- Rayner, N.A., Parker, D.E., Frich, P., Horton, E.B., Folland, C.K. and Alexander, L.V. 2000. SST and Sea-Ice fields for ERA40. *Second WCRP International Conference on Reanalysis*. WMO / TD 985, 18-21.
- Reason, C.J.C., Allan, R.J., Lindesay, J.A. and Ansell, T.J. 2000. ENSO and climatic signals across the Indian Ocean basin in the global context: Part I, Interannual composite patterns. *Int. J. Climatol.*, 20, 1285-327.
- Smith, N.R. 1995. The BMRC ocean thermal analysis system. *Aust. Met. Mag.*, 44, 93-110.
- Stone, R. and Auliciems, A. 1992. SOI phase relationships with rainfall in eastern Australia. *Int. J. Climatol.*, 12, 625-36.
- Ward, M.N. and Folland, C.K. 1991. Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea surface temperature. *Int. J. Climatol.*, 11, 711-43.