Journal of Hydroscience and Hydraulic Engineering
Vol. 20, No.2 November, 2002, 41-49

STATISTICAL CHARACTERISTICS OF SOUTHERN OSCILLATION INDEX AND ITS BAROMETRIC PRESSURE DATA

By

Akira Kawamura

Institute of Environmental Systems, Kyushu University, Hakozaki, Fukuoka, Japan

Soichiro Eguchi

Institute of Environmental Systems, Kyushu University, Hakozaki, Fukuoka, Japan

Kenji Jinno

Institute of Environmental Systems, Kyushu University, Hakozaki, Fukuoka, Japan

and

Alistair I. McKerchar

National Institute of Water and Atmospheric Research Ltd (NIWA), Christchurch, New Zealand

SYNOPSIS

The effects of the El Niño Southern Oscillation (ENSO) phenomenon on climate are widespread and extend far beyond the tropical Pacific. This phenomenon can be characterized by Southern Oscillation Index (SOI) which is derived from values of the monthly mean sea level pressure barometric difference between Tahiti and Darwin, Australia. In this study, for long-term data gathered since 1866, general statistical characteristics of SOI and the data from which they are derived (i.e. mean sea level pressure data at Tahiti and Darwin) are presented as basic information when using SOI for other analyses. These characteristics include the availability of the barometric pressure data, statistics of monthly pressure data, the correlation of Southern Oscillation (SO) intensity, frequency analysis of SOI by magnitude and by month (January-December), and duration properties of SOI by run analysis.

INTRODUCTION

El Niño is the condition in which sea surface temperature rises 1 to 2°C (sometimes 2 to 5°C) above normal in the eastern and central equatorial Pacific Ocean. It lasts typically 12-18 months, and occurs irregularly at the intervals of 2-7 years. On the other hand La Niña is a condition where the temperature becomes lower than normal. The Southern Oscillation is an atmospheric see-saw phenomenon in the tropical Pacific sea level pressure between the eastern and western hemispheres and is associated with the El Niño and La Niña oceanographic features (Sakurai (1), Japanese Study Group for Climate Impact & Application (2)). This oscillation can be measured by a simple index, the Southern Oscillation Index (SOI) (Kawamura et al. (3)) which is used by NOAA (National Oceanic and Atmospheric Administration) to determine whether the El Niño and La Niña events are occurring (Japanese Study Group for Climate Impact & Application (2)). The features are known collectively as the El Niño Southern Oscillation (ENSO) phenomenon.

The effects of ENSO on climate are widespread and extend far beyond the tropical Pacific, a phenomenon known as teleconnection (Sakurai (1), Japanese Study Group for Climate Impact & Application (2)). There have been many reports of abnormal weather conditions worldwide which are thought to have been caused by ENSO. These abnormal weather conditions and climatic changes have raise concerns about stable water resources management. As the result, qualitative and quantitative analyses of SO and its influence on local hydro-meteorological phenomena have become very important areas of research.

There are many research papers on the relationship between SOI and hydro-meteorological phenomena (e.g., Ropelewski and Halpert (4), Halpert and Ropelewski (5), Uvo et al. (6), Yoshino (7)). We have described the chaotic characteristics of SOI (Kawamura et al. (3)) and the correlation between SOI and precipitation and temperature data in Fukuoka, Japan (Eguchi et al. (8), (10), Kawamura et al. (9), (11), (12)). However, the properties of the barometric data from which SOI is derived (i.e. mean sea level pressure data at Tahiti and Darwin) and general statistical characteristics of SOI have not been published, as far as the authors know.

We have already published a paper concerning the autocorrelation and the spectral characteristics of SOI (Kawamura et al. (3)), therefore in this study, other statistical characteristics of SOI and the barometric pressure data are discussed. Firstly, the availability of the monthly mean sea level pressure data at Tahiti and Darwin, the statistics of the monthly pressure data, and the correlations between Tahiti and Darwin pressure anomalies are presented. In this study, uniquely long-term continuous monthly mean sea level pressure data at Tahiti and Darwin since 1866 are used for the above analyses. Next, the characteristics of SOI, including frequency analysis of SOI by magnitude and by month (January-December), and duration properties of SOI by run analysis are discussed. These characteristics of SOI have not investigated in detail as far as the authors know. We expect that these characteristics will offer some useful background information when using SOI for other analyses.

PROPERTIES OF MONTHLY MEAN SEA LEVEL PRESSURE DATA AS THE BASE OF SOI

Availability of the pressure data at Tahiti and Darwin

SOI values are calculated by using the monthly mean sea level pressure (MSLP) data at Papeete, Tahiti (149.6°W, 17.5°S.) and Darwin, Australia (130.9°E, 12.4°S). The MSLP pressure data since 1882 are available through web cites such as NOAA Network Information Center. Some missing values, however, exist for the Tahiti pressure data as shown in Table 1. Ropelewski and Jones (13) infilled all missing values using newly found old pressure data in Tahiti. Furthermore, by supplementing and interpolating the data of Tahiti before 1882, they completed the pressure time series since 1866. Allan et al. (14) also infilled the pressure data in Darwin before 1882 by means of interpolation using the older records and by correlation with data from other observation stations. As a result, the pressure data in Darwin are also complete since 1866.

In this study, we use the monthly MSLP data over a period of 134 years (1608 months) from January 1866 to December 1999 at Tahiti and Darwin. However, it is acknowledged that the reliability of the pre-1935 pressure data may be slightly less than the later data (Ropelewski and Jones (13)).

Characteristics of MSLP data

Fig. 1 shows the mean values and their standard deviations of monthly MSLP for each month (January-December) at Tahiti and Darwin for over a period of 134 years. The mean values of MSLP are higher in the summer season in northern hemisphere and lower in the winter season for both sites. The difference between the highest and the lowest mean values is 6.8 hPa at Darwin, which is about twice as much as that (3.7 hPa) at Tahiti. The standard deviations are larger in January-March and smaller in April-June for both sites.

Table 1 MSLP data missing periods at Tahiti

Missing Period		Missing
From	To	Months
Sep 1892	Dec 1895	40
Mar 1906	May 1906	3
Dec 1906	Aug 1908	21
Apr 1914	Sep 1914	6
Nov 1914	Oct 1915	12
Mar 1921	Jun 1921	4
Jun 1927	Aug 1927	3
Aug 1931	Aug 1932	13
	Total	102

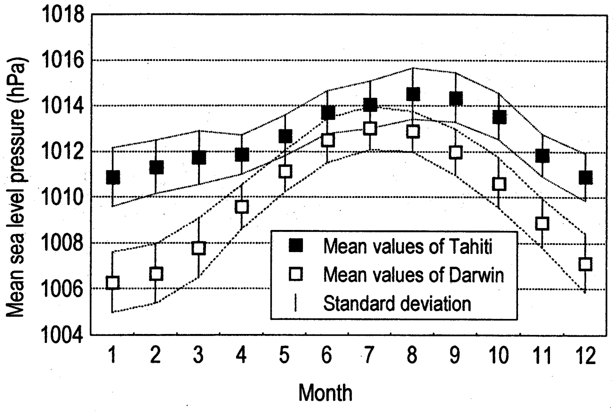

Fig. 1 The mean values and their standard deviations of monthly MSLP for each month (January-December) at Tahiti and Darwin for 134 years data

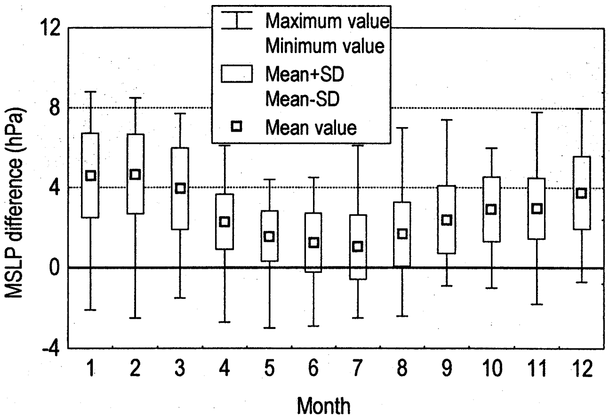

Fig. 2 Average values, standard deviations, maximum and minimum values of MSLP difference between Tahiti and Darwin for each month (January-December)

Fig. 2 shows the average values, their standard deviations, maximum and minimum values of MSLP difference between Tahiti and Darwin for each month (January-December). From the figure, the average values of MSLP differences are positive for all months, but they are close to zero in the northern hemisphere in summer. The mean minus standard deviation values are negative in June and July. The minimum values are negative for all months.

To check the reliability of the MSLP data, we divided the series into two groups, 1866-1934 (69 years), and 1935-1999 (65 years). The average values and their standard deviations for both groups are almost the same. The mean value of the difference of monthly average values is 0.18 hPa, and the mean value of the standard deviations is 0.13 hPa.

Cross-correlation of Tahiti and Darwin pressure deviations

Deviation time series from the average MSLP of each month (January-December) are shown in Fig. 3 for the last 30 years. From this figure, we can see a clear tendency that when the pressure deviations rise above the monthly averages at Tahiti, the pressure deviations at Darwin fall below the monthly averages, and vice versa. This phenomenon was discovered by Sir Gilbert Walker early in the 20th century, and it was named the Southern Oscillation.

Fig. 4 shows the scatter plots for Fig. 3, that is, the relationship between pressure deviations from the

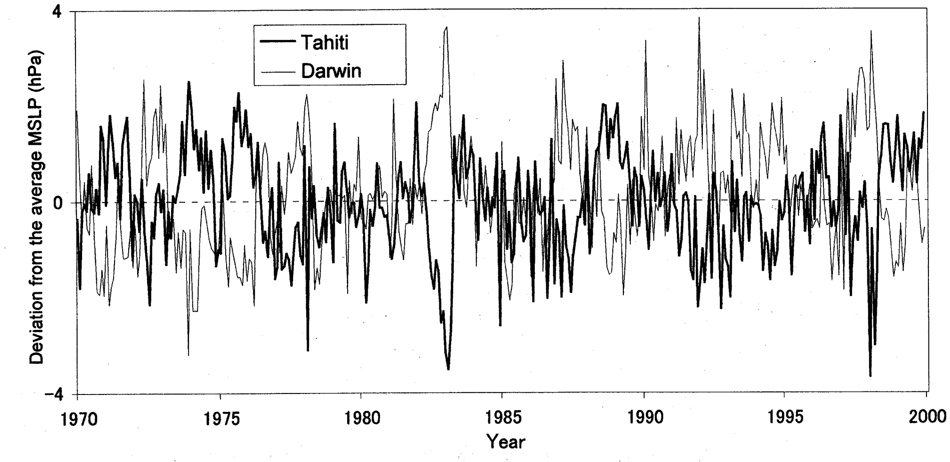

Fig. 3 Deviation time series from the average MSLP for the recent 30 years

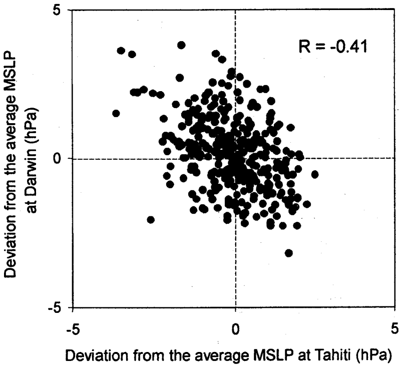

Fig. 4 Scatter plots for Fig. 3

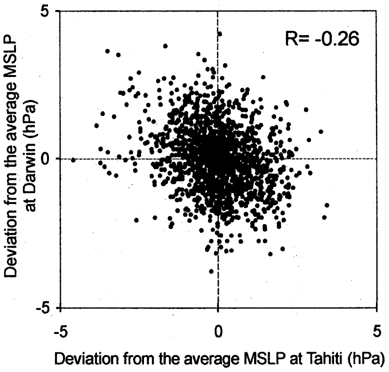

Fig. 5 Scatter plots of deviation from the average MSLP for the whole 134 years

average MSLP at Tahiti and Darwin. This figure shows that we can obtain the cross-correlation coefficient of -0.41 for the most recent 30 years. The correlation is not so high as we expect. For the whole series (Fig. 5), the correlation coefficient of pressure deviations is -0.26 which is even smaller. However, it should be noted that the significant level of those correlation coefficients is extremely high because of the large number of the data. This means that when the statistical hypothesis test is executed, the null hypothesis of "correlation coefficient = 0" is easily rejected at the significant level of 0.001%. Furthermore, the correlation coefficient becomes larger when the Darwin data are lagged by a one-month lag: it becomes -0.29 for the whole data.

STATISTICAL CHARACTERISTICS OF SOI

Calculation of SOI

Two commonly used methods to compute the SOI from the MSLP data at Tahiti and Darwin are Troup's method and the Climate Prediction Centre's method. The difference between two methods is very small as pointed out by McBride and Nicholls (16), Ropelewski and Jones (13) and Kawamura et al. (3). In this study, we use Troup's method (Troup (15)). The $SOI(y,m)$ in year y, month m(m=January to December) is calculated by the following equation:

$$SOI(y,m) = [\{P_T(y,m)-P_D(y,m)\}-M_{30}(m)]/S_{30}(m) \qquad (1)$$

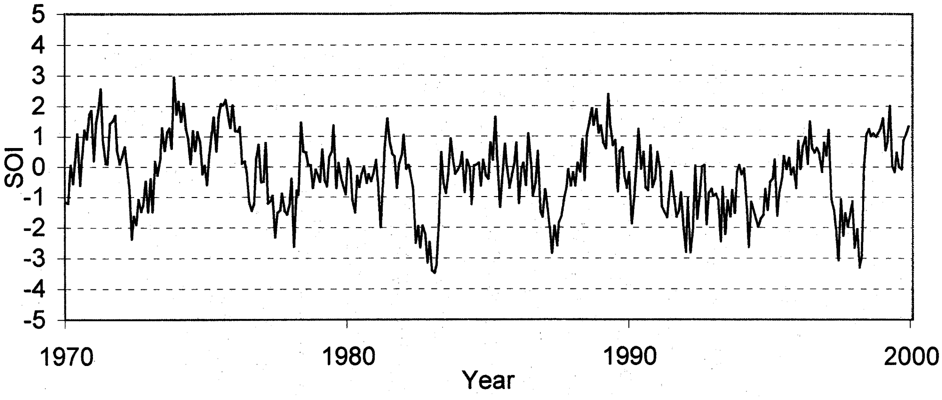

Fig. 6 SOI time series for the recent 30 years

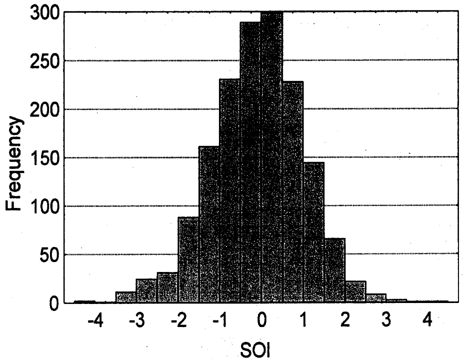

Fig. 7 Histogram of original SOI time series

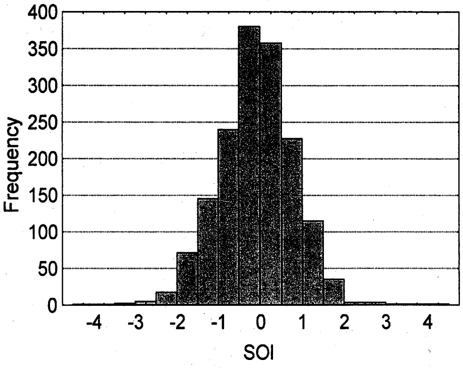

Fig. 8 Histogram of 5 months moving average time series of SOI

Here, $P_T(y,m)$, $P_D(y,m)$ = MSLP (hPa) at Tahiti and Darwin, respectively ; $M_{30}(m)$, $S_{30}(m)$ = the mean value (hPa) and its standard deviation (hPa) of MSLP difference between Tahiti and Darwin (cf. Fig. 2) for the base period of 30 years (usually 1951-1980). As is indicated in Eq.1, SOI is expressed as the MSLP difference series between Tahiti and Darwin, which is normalized to mean zero and a standard deviation of one. Note that a standard deviation of 10 is also commonly used. Generally, El Niño conditions occur when the SOI is less than -1, and La Niña conditions occur when the SOI is more than 1. Strong El Niño/La Niña conditions prevail when the SOI absolute value exceeds 2. Fig. 6 shows the SOI time series for the past 30 years. From this figure, significant El Niño events are found in 1982/1983 and 1997/1998.

Frequency characteristics of SOI

In several previous papers SOI values were classified into various categories according to their magnitudes, and the relationship between the categorized SOI and hydro-meteorological elements were examined (e.g., Moss et al. (17), McKerchar et al. (18), Eguchi et al. (8), (10), Kawamura et al. (9), (11), (12)). In this section, using their findings, frequency analysis of classified SOI will be studied.

Fig. 7 shows the histogram SOI data over a period of 1608 months. It looks normally distributed as expected. However, frequencies of the negative values are larger than those of positive values except when the absolute values are less than 0.5. For example, the frequency less than -2 is 69, while the frequency more than 2 is 34, i.e., El Niño tendencies occur more frequently than La Niña tendencies. This tendency becomes more dominant when the SOI time series is smoothed. For example, in the histogram of 5 months moving average time series of SOI (Fig. 8), although the frequencies themselves more than 2 absolute value decrease, the frequency more than 2 is 8, whereas the frequency less than -2 is 25. The reason for using 5 months for

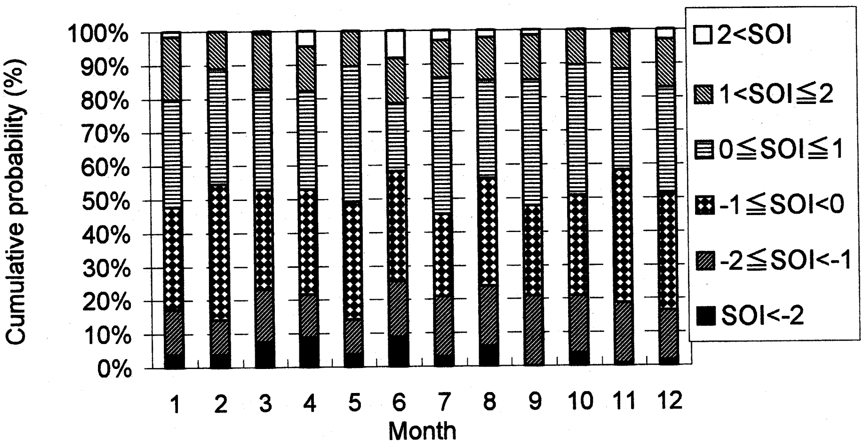

Fig. 9 Occurrence of SOI classified into six categories according to their magnitude

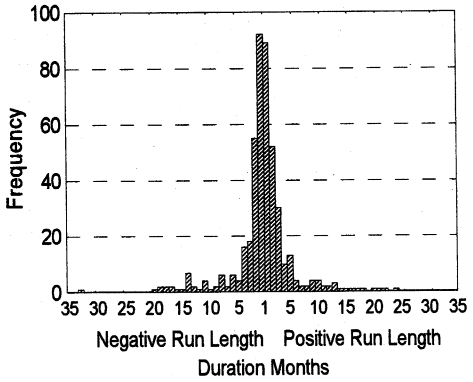

Fig.10 Histogram of positive and negative run length for SOI time series

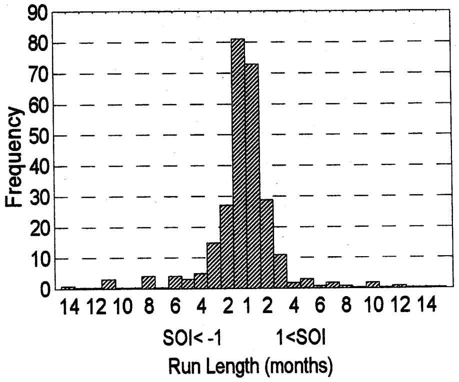

Fig. 11 Histogram of duration analysis result for the case in which the absolute value of SOI is more than 1

moving average anomalies of sea surface temperature In the El Niño monitoring area exceed 0.5 degrees and they last more than 6 months. Similarly, La Niña is defined as a condition where the anomalies become less than -0.5 degrees.

Fig. 9 shows cumulative frequency of SOI classified into six categories for each month (January-December) during the past 134 years. From this figure, in spite of a small variation of the frequency of each category from month to month, it can be seen that El Niño and La Niña could occur in any month. However, during the month of June the frequencies of more than 2 and less than -2 are extremely high compared with those of other months. Also, the frequencies of more than 1 and less than –1 are much higher, whereas the ratio between -1 and 1 is very low. On the other hand, the frequencies of high absolute values are low in February and May, that is, the frequency of more than 2 for those months are zero, and the frequencies of values less than -1 are less than for other months. This phenomenon can be attributed to seasonality in the evolution of El Niño and La Niña events (Trenberth and Shea (19)).

Duration properties of SOI time series

In order to investigate the duration properties of SOI, a run analysis of the time series is carried out. Fig. 10 shows the histogram of positive run length (La Niña side) and negative run length (El Niño side). A Positive run length is defined by the number of consecutive months in which the positive values of SOI

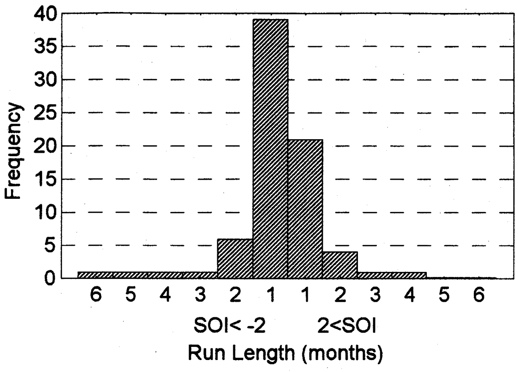

Fig. 12 Histogram of duration analysis result for the case in which the absolute value of SOI is more than 2

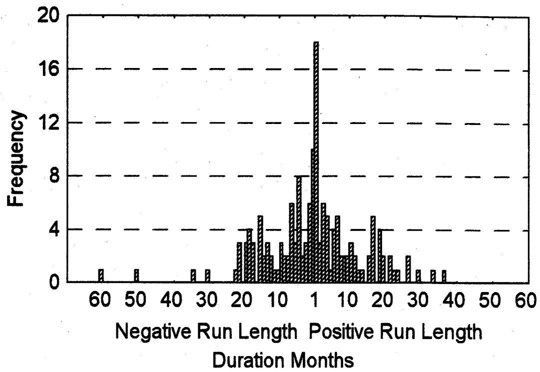

Fig.13 Histogram of positive and negative run length for 5 months moving average SOI time series

continue, and similarly, negative run length is for the negative values. From Fig. 10, it can be seen that both positive and negative run lengths have the biggest frequency in one month, and that the frequency decreases drastically with each increase of run length. The average of positive run length is 3.4 months, and 3.7 months for negative run length. The longest negative run length of 33 months is from August 1939 to April 1942. The negative run length for the large-scale El Niño event covering the period from the spring of 1982 to the summer of 1983 is 13 months, and 14 months for another large-scale El Niño event covering the period from the spring 1997 to the summer 1998.

Next, the duration analysis is carried out for the case in which the absolute value of SOI is more than 1, which means how many months SOI time series is continuously above one or below minus one. The result is shown in Fig. 11. This figure indicates that the biggest run length is 14 months of negative side, which corresponds to the 1997/1998 El Niño event. The biggest positive one is 12 months which corresponds to the 1917 La Niña event. Average run lengths are 2.0 and 2.2 months for positive and negative, respectively, in which El Niño side tends to be a bit longer. The similar duration analysis (Fig. 12) is also carried out for the case in which the absolute value of SOI is more than 2, which corresponds to strong El Niño and La Niña. From Fig. 12, the biggest run length is 6 months of negative side, which corresponds to 1982/1983 El Niño event, and the biggest positive one is 4 months which corresponds to the 1917 La Niña event. The same tendency of longer duration of the El Niño side is obtained.

In addition, the similar run analyses are carried out for 5 months moving average SOI time series, and one of the results is shown in Fig. 13. From this figure (compared with Fig. 10), the biggest frequencies are obtained in one month for both positive and negative run length, which is the same result as in Fig. 10. However, the frequencies themselves decrease especially for the negative run length. The average lengths naturally become longer, and they are 9.6, 11.2 months for positive and negative run length, respectively, which indicates El Niño side becomes longer than that of Fig. 10.

In Fig. 13, the longest negative run length of 61 months started from August 1990, whereas the second longest negative began in March 1911 and lasted 51 months. The longest negative run length of 33 months in Fig. 10 corresponds to 35 months in Fig. 13 with almost no increase of run length. As the result, the order of the run length periods does not correspond well between SOI time series (Fig. 10) and its 5 months moving average time series (Fig. 13).

CONCLUSIONS

In this paper, firstly the characteristics of MSLP data at Tahiti and Darwin as the base of SOI are investigated. Then, general statistical characteristics of SOI including frequency and duration properties are studied. The findings obtained in this study are as follows :

-1. The consecutive SOI data are available since January 1866. The pre-1935 data may be slightly less reliable.

-2. The MSLP is higher in the northern hemisphere summer season, and lower in winter season for both Tahiti and Darwin. The range of mean values of MSLP at Darwin is about twice as big as that at Tahiti. The

standard deviations of MSLP are largest in January-March and least in April-June for both sites.
-3. The average values of MSLP difference between Tahiti and Darwin are positive for all months, but they are close to zero in the northern hemisphere summer season.
-4. Absolute value of the correlation coefficient of Southern Oscillation intensity over the whole period of 134 years is less than 0.3, which is not as high as we expect. However, the significant level of the correlation is extremely high according to the result of statistical hypothesis test. A higher correlation coefficient is obtained when a one-month lag is considered for the Darwin data.
-5. The frequency of SOI less than -2 (strong El Niño) is about twice that of SOI more than 2 (strong La Niña).
-6. The occurrences of SOI more than 2 and less than -2 are remarkably higher in June compared with other months, and occurrences between -1 and 1 are relatively low. Extreme values of the SOI are least likely in February and May.
-7. Both positive and negative run lengths of SOI time series have the biggest frequency in one month, and the frequency decreases drastically with each increase of run length. The negative run length is longer than the positive run length on the average. This tendency persists for the absolute values of SOI more than 1 and more than 2.
-8 For 5 months moving average SOI time series, the frequencies of run length decrease compared with the original SOI time series especially for the negative run length. The negative value run length is longer than the positive value run length.
-9 The order of the run length periods does not correspond well between original SOI time series and its 5 months moving average time series.

We hope that the findings of this study will provide some useful background information when using SOI for other analyses.

REFERENCES

1. Sakurai, K. : Aiming at forecasting El Niño events. Tropic Weather Affecting Japan (in Japanese), Meteorological Society of Japan, Kansai Branch, the 20th Summer School Textbook, pp.38-57, 1998.
2. Japanese Study Group for Climate Impact & Application : El Niño & Global Environment (in Japanese), Seizando, Japan, 1999.
3. Kawamura, A., A.I. McKerchar, R.H. Spigel and K. Jinno : Chaotic characteristics of the Southern Oscillation Index time series, Journal of Hydrology, Vol.204, pp.168-181, 1998.
4. Ropelewski, C.F. and M.S. Halpert : Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation, Monthly Weather Review, Vol.115, pp.1606-1626, 1987.
5. Halpert, M.S. and C.F. Ropelewski : Surface temperature patterns associated with the Southern Oscillation, Journal of Climate, Vol.5, pp.577-593, 1992.
6. Uvo, C.B., C.A. Repelli, S.E. Zebiak and Y. Kushnir : The relationships between tropical Pacific and Atlantic SST and northeast Brazil monthly precipitation, Journal of Climate, Vol.11, pp.551-562, 1998.
7. Yoshino, F : On the relation between the monthly rainfall in Shikoku island and the El Niño event (in Japanese), Proceedings of Shikoku Branch of JSCE, pp.90-91 , 1999.
8. Eguchi, S., A. Kawamura and K. Jinno : Correration between Southern Oscillation Index and precipitation in Fukuoka (in Japanese), Proceedings of Seibu Branch of JSCE, No.2, pp.178-179, 2000.
9. Kawamura, A., S. Eguchi and K. Jinno : Cross-Correlation between Southern Oscillation Index and precipitation in Fukuoka, Japan, Supplement to EOS, Transactions, American Geophysical Union, Vol.81, No.22, p.WP57, 2000.
10. Eguchi, S., A. Kawamura and K. Jinno : Correration between Southern Oscillation Index and temperature in Fukuoka (in Japanese), Proceedings of Annual Conference of Japan Society of Hydrology and Water Resources, pp.120-121, 2000.
11. Kawamura, A., S. Eguchi and K. Jinno : Cross-Correlation between Southern Oscillation Index and precipitation /temperature in Fukuoka, Japan, Proc. of Fresh Perspectives Symposium 2000, Christchurch, New Zealand, 2000.
12. Kawamura, A., S. Eguchi and K. Jinno : Correlation between Southern Oscillation and monthly precipitation in Fukuoka (in Japanese with English abstract), Journal of Hydraulic, Coastal and Environmental Engineering, Japan Society of Civil Engineers, 2001 (in press).
13. Ropelewski, C.F. and P.D. Jones : An extension of the Tahiti-Darwin Southern Oscillation index, Monthly Weather Review, Vol.115, pp.2161-2165, 1987.
14. Allan, R.J., N. Nicholls, P.D. Jones and I.J. Butterworth : A further extension of the Tahiti-Darwin SOI, early ENSO events and Darwin pressure, Journal of Climate, Vol.4, pp.743-749, 1991.

15. Troup, A.J. : The "southern oscillation", Quarterly Journal of the Royal Meteorological Society, Vol.91, No.390, pp. 490-506, 1965.
16. McBride, J.L. and N. Nicholls : Seasonal relationships between Australian rainfall and the Southern Oscillation, Monthly Weather Review, Vol.111, pp.1998-2004, 1983.
17. Moss, M.E., C.P. Pearson and A.I. McKerchar : The Southern Oscillation index as a predictor of the probability of low streamflows in New Zealand, Water Resources Research, Vol.30, No.10, pp.2717-2723, 1994.
18. McKerchar, A.I., C.P. Pearson and B.B. Fitzharris : Dependency of summer lake inflows and precipitation on spring SOI, Journal of Hydrology, Vol.205, pp.66-80, 1998.
19. Trenberth, K.E. and D.J. Shea : On the evolution of Southern Oscillation, Monthly Weather Review, Vol.115, pp.3078-3096, 1987.

APPENDIX – NOTATION

The following symbols are used in this paper :

$M_{30}(m)$, $S_{30}(m)$ = the mean value (hPa) and its standard deviation (hPa) of MSLP difference between Tahiti and Darwin for the base period of 30 years ;

$P_T(y,m)$, $P_D(y,m)$ = Mean Sea Level Pressure (hPa) at Tahiti and Darwin ; and

$SOI(y,m)$ = value of SOI in the year y and the month m (m=January to December) .

(Received September 13, 2001 ; revised June 18, 2002)