Statistical prediction of Sri Lankan rainfall during October to December

PABODINI KARUNAPALA*, ** and CHANGHYUN YOO*, ***

*Department of Climate and Energy Systems Engineering,

Ewha Womans University, Seoul, Korea

**Department of Meteorology, Colombo, Sri Lanka

***Center for Climate/Environment Change Prediction Research,

Ewha Womans University, Seoul, Korea

e mail : pabodini.karunapala@gmail.com

सार – श्रीलंका में अक्टूबर से दिसंबर (ओ.एन.डी) के दौरान सबसे अधिक वर्षा होती है, लेकिन विश्वसनीय ऋतुनिष्ठ पूर्वानुमान प्रणाली अभी तक परिचालित नहीं हुई है। यहां हमने अक्टूबर, नवम्बर, दिसंबर में श्रीलंका की 1979-2012 के दौरान 1 और 2 महीने के अग्रकाल के दौरान हुई वर्षा का पूर्वानुमान करने के लिए बहु रैखिक समाश्रयण मॉडलों का निर्माण किया है। श्रीलंका की अक्टूबर, नवम्बर, दिसंबर की वर्षा और वैश्विक समुद्र सतह तापमान (एस.एस.टी.) विसंगतियों के बीच संबंधों की जांच करने के लिए सह-संबंध विश्लेषण का उपयोग किया गया। आंशिक रूप से सबसे कम वर्ग समाश्रयण विधि के माध्यम से दो अलग-अलग लीड समय परतीन स्वतंत्र अनुमानों की पहचान की गई, जिसमें दक्षिणी अटलांटिक एस.एस.टी. प्रवृति, दक्षिणी प्रशांत एस.एस.टी. प्रवृति और पश्चिमी प्रशांत और समुद्री महाद्वीप एस.एस.टी. प्रवृति शामिल हैं। तीन-वर्ष के पुनः सत्यापन से यह निष्कर्ष निकलता है कि कई रैखिक प्रतिगमन मॉडल क्रमशः 1 और 2 महीने के लीड समय के लिए 0.69 और 0.68 के सहसंबंध गुणांक कौशल पर अक्टूबर, नवम्बर, दिसंबर की वर्षा का पूर्वानुमान लगा सकते हैं। इन तीनों पूर्वसूचकों से जुड़ी भौतिक प्रक्रियाओं से श्रीलंका की अक्टूबर, नवम्बर, दिसंबर की वर्षा में वृद्धि में योगदान का पता चलता है।

ABSTRACT. Sri Lanka receives most rainfall during October to December (OND). Here we construct multiple linear regression models to forecast the OND Sri Lankan rainfall during 1979-2012 for lead times of 1 and 2 months. Correlation analysis was used to examine the relationship between Sri Lankan OND rainfall and global sea surface temperature (SST) anomalies. Three independent predictors were identified through partial least square regression method which includes the southern Atlantic SST tendency, southern Pacific SST tendency and western Pacific and Maritime Continent SST tendency at two different lead times. Three-year-out cross validation concludes that the multiple linear regression models can produce forecast the OND rainfall forecast at correlation coefficient skill of 0.69 and 0.68 for the 1 and 2 month lead times respectively. The physical processes associated with these three predictors show that they contribute to increase in OND rainfall of Sri Lanka.

Key words – Multiple-regression models, Seasonal forecast, OND Sri Lankan rainfall.

1. Introduction

Sri Lanka receives the major portion of its annual rainfall during October to December (OND) (Zubair and Ropelewski, 2006). This primary rainy and agricultural season is locally known as "Maha" period. In addition to agricultural productivity, hydropower planning, water resource management and disaster preparedness of Sri Lanka are strongly influenced by the OND seasonal rainfall. During the season, the rainfall rate is intense [approximately 150 ~ 200 mm month^{-1}; Figs. 1(a&b)], which often causes floods and landslides (Zubair et al., 2006). Therefore, prediction of the seasonal rainfall during OND is important, yet one of the most challenging tasks in Sri Lanka. Although interannual relationships of seasonal rainfall with leading climate modes, such as the El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD), have been examined (Zubair and Ropelewski, 2006; Zubair et al., 2003). The main objective of this study is hence to construct a statistical model, which will be a valuable tool to reduce disaster risk and economic losses in Sri Lanka.

The rainfall during the OND season can be attributed to multiple meteorological phenomena, such as the formation of tropical cyclones and depressions in the Bay
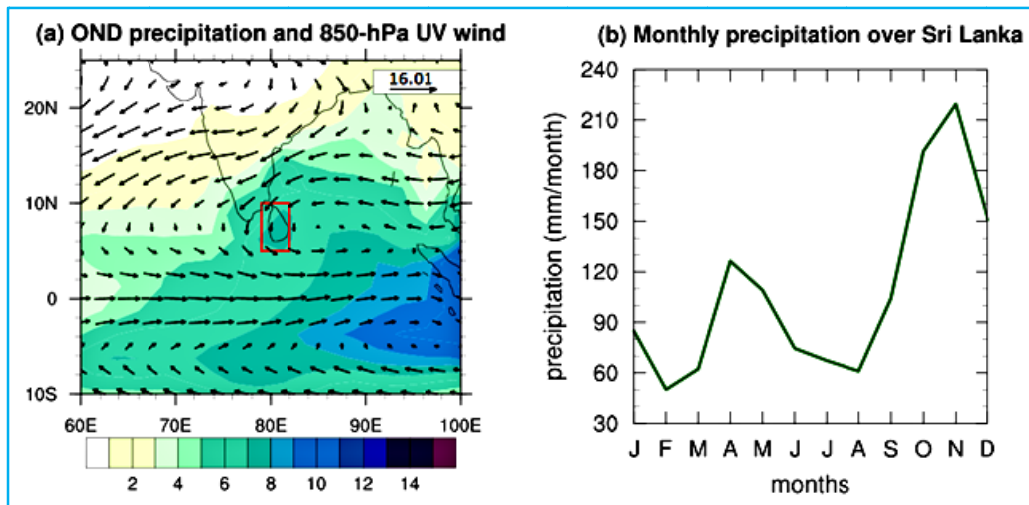
Figs. 1(a&b). (a) Climatological OND 850 hPa winds (ms^{-1}) and mean rainfall (mm day^{-1}) are shown. The red box indicates the domain of Sri Lanka used in this study (5° N to 10° N and 79° E to 82° E) and (b) Monthly climatology of rainfall amount averaged over Sri Lanka region for the period 1979-2012 are shown

of Bengal, southward passage of the Inter-Tropical Convergence Zone (ITCZ), the easterly waves, the orographic rainfall and coastal convergence (Zubair *et al.*, 2003). Climatologically, during June to September, the low-level winds over the south Asia blow from south west. Beginning late September, the south westerly wind weakens, allowing Arabian Sea and Bay of Bengal to warm. The warmed north Indian Ocean creates widespread deep convections and heavy thunderstorms. In October and November, warmest temperatures and lowest pressures are observed in the Bay of Bengal, which is favourable for the formation of low-pressure areas over the Bay of Bengal. In fact, the largest number of tropical cyclones has occurred in the northern Indian Ocean during OND (Singh and Sontakke, 1999). In December to February, the center of heating then moves to the Maritime Continent-northern Australian region. The heating over the oceans and cooling over Siberian-Mongolian region facilitate low level north-easterlies over the Bay of Bengal (Prasanna and Yasunari, 2008). During this season, moist, north-easterlies are established over the Bay of Bengal, providing favorable conditions for heavy precipitation over Sri Lanka.

Charney and Shukla (1981) have demonstrated that tropical interannual variability is largely determined by slowly varying boundary conditions such as sea surface temperature (SST). There are seasonal rainfall forecast models that are based on predictability of SST. For example, Yim (2015) predict "Meiyu" (May-June) rainfall in Taiwan using three predictors of the 2-m air temperature and SST. Their prediction models were developed based on the physical understanding of lead-lag relationship between their predictors and the Meiyu

rainfall. Similarly, Lee and Seo (2013) develop an SST-based statistical model to forecast Changma (mid-June to end of July) precipitation over South Korea for a lead time of a month. Their model has recently revisited by Kim (2017), who improved the forecast skill by utilizing additional predictors based on oceanic, land & atmospheric interactions. For South Asia, Sahai (2003) finds that the SST can be a good predictor for Indian summer monsoon rainfall. As having demonstrated by these studies, it has also been reported that for Sri Lanka, the ocean is an important modulator of moisture transport and hence rainfall (Zubair *et al.*, 2003; Prasanna & Yasunari, 2008).

Motivated by these previous studies, we examine the lead-lag relationships between Sri Lankan OND rainfall and SSTs and construct statistical prediction models to predict the OND rainfall. To do so, our study aims first to identify the predictors by examining the relationship between the global sea surface temperature anomaly (SSTA) tendencies and the Sri Lankan OND rainfall. In addition, we investigate the physical and dynamical processes associated with the predictors and develop multiple regression models to predict OND rainfall at 1- and 2-month lead times.

The remainder of the study is organized as follows. Section 2 describes the data and methodology. Section 3 discusses the relationship between the Sri Lankan OND rainfall and global SSTA, the statistical forecast model & related physical and dynamical processes which explain the link between the predictors and Sri Lankan OND rainfall. Summary and the conclusions are provided in section 4.

2.   Data and methodology

2.1. *Data*

For precipitation, we use the monthly Global Precipitation Climatology Project (GPCP) Version 2.2 (Huffman *et al.*, 1997) at the spatial resolution of 2.5° latitude by 2.5° longitude. The monthly SSTs are obtained from the Version 1.1 of Hadley Centre Global Ice and Sea Surface Temperature dataset (HadISST1) provided by the United Kingdom Meteorological Office (UKMO) (Rayner *et al.*, 2003). The SST dataset is originally at the 1.0° latitude by 1.0° longitude resolution but is regridded to the 2.5° by 2.5° resolution of the GPCP. For monthly zonal wind and meridional wind (at what level?), the ERA-Interim dataset produced by European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee *et al.*, 2011) is used. The ERA-Interim is downloaded at a native horizontal resolution of ~ 60 km but is also interpolated to the 2.5° by 2.5° resolution.

2.2. *Methodology*

Two multiple regression models are constructed to predict the OND rainfall anomalies over Sri Lanka. Setting up the Sri Lanka rainfall index (SLRI) as the predict and is the first step to build the prediction model. The SLRI is defined as the normalized time series of OND rainfall anomaly averaged over Sri Lanka region (79° E to 82° E and 5° N to 10° N) for the period of 1979-2012. The index will be presented in Section 3.

To provide stable and effective prediction at seasonal time scale, SST is often chosen as one of slowly varying boundary conditions (Lau *et al.*, 2000; Park *et al.*, 2015; Yim *et al.*, 2013). Most of the statistical seasonal prediction models use lead-lag relationships between SSTA tendency and predict and (Yim *et al.*, 2015; Lee and Seo, 2013; Kim *et al.*, 2017; Yim *et al.*, 2013). We use the Pearson correlation coefficients, r, to identify the relationships between the Sri Lankan OND rainfall and SSTA tendency. The predictors are selected when the correlation coefficient exceeds the 95% significance level. Estimation of the statistical significance is based on a *t*-test that uses the *t*-statistic, $t = r\sqrt{(n-2)/(1-r^2)}$. In this equation, the number of seasons is used as the degree of freedom (n).

To find the lead-lag linkage between the predictors and the predict and, the correlation coefficients between the SLRI and tendency of SSTAs were calculated at 1- and 2-month lead times. We first define the 1-month lead time as the difference between August minus June. This forecast model uses the predictors which have information before and during August and then the 1-month lead time

is defined as the difference between July minus May, which includes the information during and before July.

To make sure the predictors are independent from each other, the partial least square regression (PLSR) method, *e.g.*, Black (2017) is employed as follows:

(*i*)   Grid by grid correlation coefficients between the global SSTAs and SLRI are calculated to obtain the first correlation map,

(*ii*)   Statistically significant regions at 95% confidence level are selected as predictors. Based on the significance test, the first predictor field is identified [show a table giving a list of selected predictors (time period of derived predictor) along with correlation coefficients with significant level],

(*iii*)   Area-weighted predictor field is normalized by subtracting its mean and dividing it by its standard deviation,

(*iv*)   First partial regression is obtained by using conventional least squares fitting and regressing the SSTAs against the first predictor,

(*v*)   First partial regression is linearly removed from both the predictor field and all SSTA field. The residual predictor field became as the new predictor and the residual SSTA field became as the new predictor field and

(*vi*)   The residual SSTA field is used to find the second predictor.

Steps 1-5 are repeated to obtain the other predictors. This procedure is terminated when there are no further significant predictor fields. We limit the number of predictors up to 3 so as to avoid the over fitting problem (Lee and Seo, 2013; Kim *et al.*, 2017)

2.3. *Cross-validation*

The regression coefficients remain stable when using a cross validation method, which is widely used in climate prediction. That is, to examine the performance of the model, we employ two cross validation approaches. The first validation approach is following Blockeel & Struyf (2002), who suggest that 50%--70% data can be used to construct the regression model and the remaining data can be used to validate the model. For this approach, we divide the entire 34-year data into two subsets as the training period and the validation period. For the training period, first 21 years data (1979-1999) are used to obtain the regression coefficients for the model. The remaining 13 years data (2000-2012) are then used to make the independent forecast.
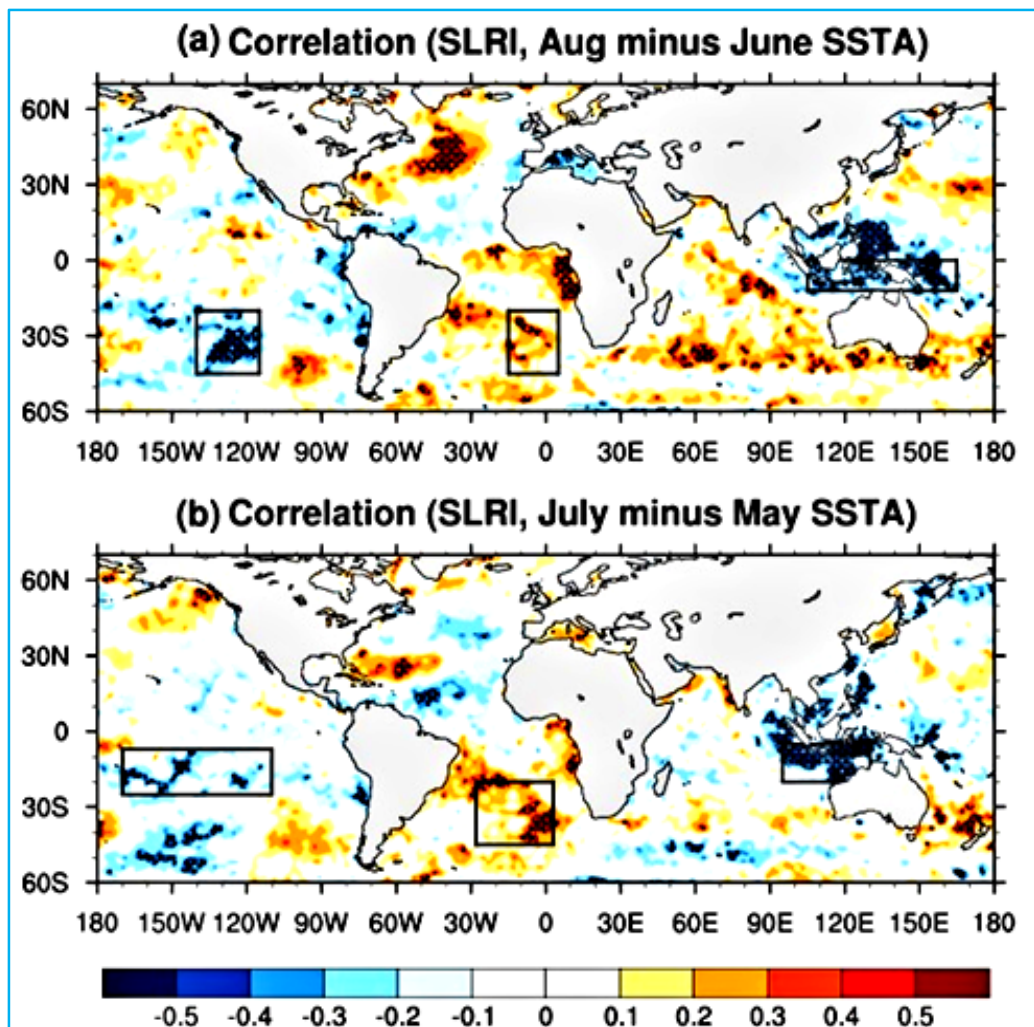
Figs. 2(a&b). Correlation coefficient maps between the SSTA and SLRI are shown at (a) 1- and (b) 2-month lead times, respectively. The boxes indicate the regions of the three predictors. Black crosses mark the areas that are statistically significant at the 95% confidence level

The second approach is based on the three years out cross validation method (Yim *et al.*, 2013). For this approach, we develop the model all years, but excluding the three years centered at the year that the prediction and hence the validation is performed. The procedure is repeated by taking 3-year out around each predicted year. That is, this leave-three-years-out cross validation involves using 3 observations as the validation set and the remaining observations as the training set.

3.   Results and discussion

3.1. *Selection of the best predictors*

To investigate the monthly rainfall evolution, monthly climatology of the Sri Lankan rainfall is computed by averaging the rainfall amount over the domain (79° E to 82° E and 5° N to 10° N; red box in Fig. 1(a) during the period of 1979-2012 [Fig. 1(b)]. Bimodality is apparent in the mean annual cycle of the rainfall with a primary peak from October to December and a subsidiary peak from April to June. The rainfall amount gradually increases from September and attains the highest rainfall from October to November. The OND mean low-level winds at 850 hPa [vectors in Fig. 1(a)] show cyclonic circulations and easterly / northeasterly trade winds over Bay of Bengal. The formation of the low-level cyclonic circulation to the east of Sri Lanka and the moist northeasterly winds blowing across Sri Lanka are favorable for heavy rainfall over the island. The OND mean precipitation is centered near the

TABLE 1

Definitions and the domains of the predictors selected for the prediction of Sri Lanka OND rainfall respectively at 1- and 2-month lead times

Predictor	Meaning	August minus June (1-month lead)	July minus May (2-month lead)
WP & MC	Western pacific and maritime continent SST tendency	0-12S 105E-165E	05S-20S 95E-130E
SP	Southern Pacific SST tendency	20S-45S 115W-140W	07S-25S 110W-170W
SA	Southern Atlantic SST tendency	20S-45S 05E-15W	20S-45S 03E-28W

TABLE 2

Correlation coefficients between the SLRI and predictorsrespectively at 1- and 2-month lead times

Predictor	August minus June (1-month lead)	July minus May (2-month lead)
WP &MC	-0.40*	-0.40*
SP	-0.47**	-0.40*
SA	0.32	0.37*

* Significance correlation at 95% confidence level,
** Significance correlation at 99% confidence level.

TABLE 3

Temporal correlation coefficients for the prediction respectively at 1- and 2-month lead times

Lead time	Correlations		
	Training period (1979-1999)	Validation period (2000-2012)	Cross validation (1979-2012)
1-month lead	0.69**	0.65**	0.69**
2-month lead	0.68**	0.71**	0.68**

** Significance correlation at 99% confidence level.

Maritime Continent and decreases toward Sri Lanka [shading in Fig. 1(a)].

To select the predictors for the 1-month lead time, the correlation coefficients between the SLRI and August minus June SSTA tendencies are calculated [Fig. 2(a), Table 1]. The correlation pattern shows some regions being significantly correlated with the SLRI. When there are many potential predictors while their physical relationships with the predict and are not well defined, a
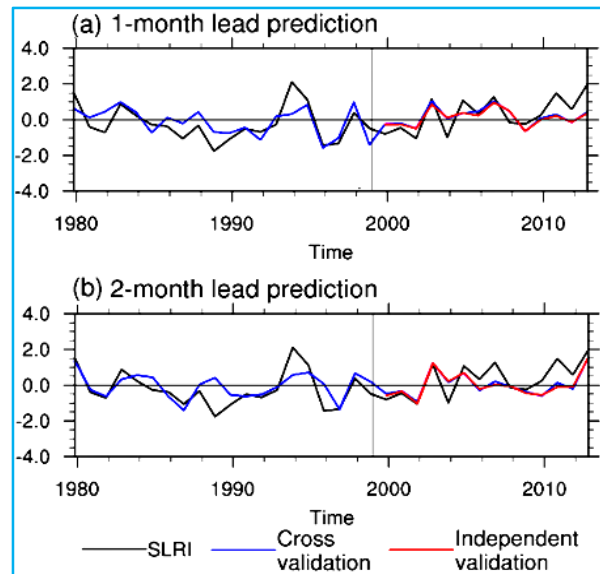
Figs. 3(a&b). Seasonal rainfall predictions are made using the multiple regression models at (a) 1-month and (b) 2-month lead times. The observation, *i.e*., SLRI is shown in black. The three-year-outcross-validated prediction is shown in blue. The cross validation is performed by taking 3-year out around the predicted year. Independent prediction for the validation period (2000-2012) is shown in red. The model for the independent prediction is built using the data of the training period (1979-1999)

few of them should be selected based on statistical methods to avoid collinearity (Sahai *et al.*, 2003). This procedure begins with the construction of simple linear regression models for each potential predictor variable. The predictor field, which has the maximum correlation coefficient with the minimum root mean square error (RMSE), is selected as the first predictor field (Sahai *et al.*, 2003; Del Sole and Shukla, 2002). Based on this condition, an area over the southern Atlantic SSTA tendency (SA) is selected as the first predictor field [box near the date line in Fig. 2(a)]. The southern Atlantic SSTA tendency field is indexed through, first, area averaging and, then, normalization. The index will be referred to SA hereafter (Table 1). The first partial regression coefficient is obtained by regressing the SA against the SLRI and the value is 0.292 [Eqn. (1)].

The second predictor is obtained via the correlation coefficients between the SLRI and the residual SSTA field, where the residual is defined by linearly removing the SA from the SSTA field. Through this procedure, the southern Pacific area [box near 140° W in Fig. 2(a)] is chosen as the second predictor field. Similarly, for the SA, the southern Pacific SSTA tendency (hereafter SP, Table 1) is area averaged and normalized, before it is regressed against the SLRI to obtain the second partial
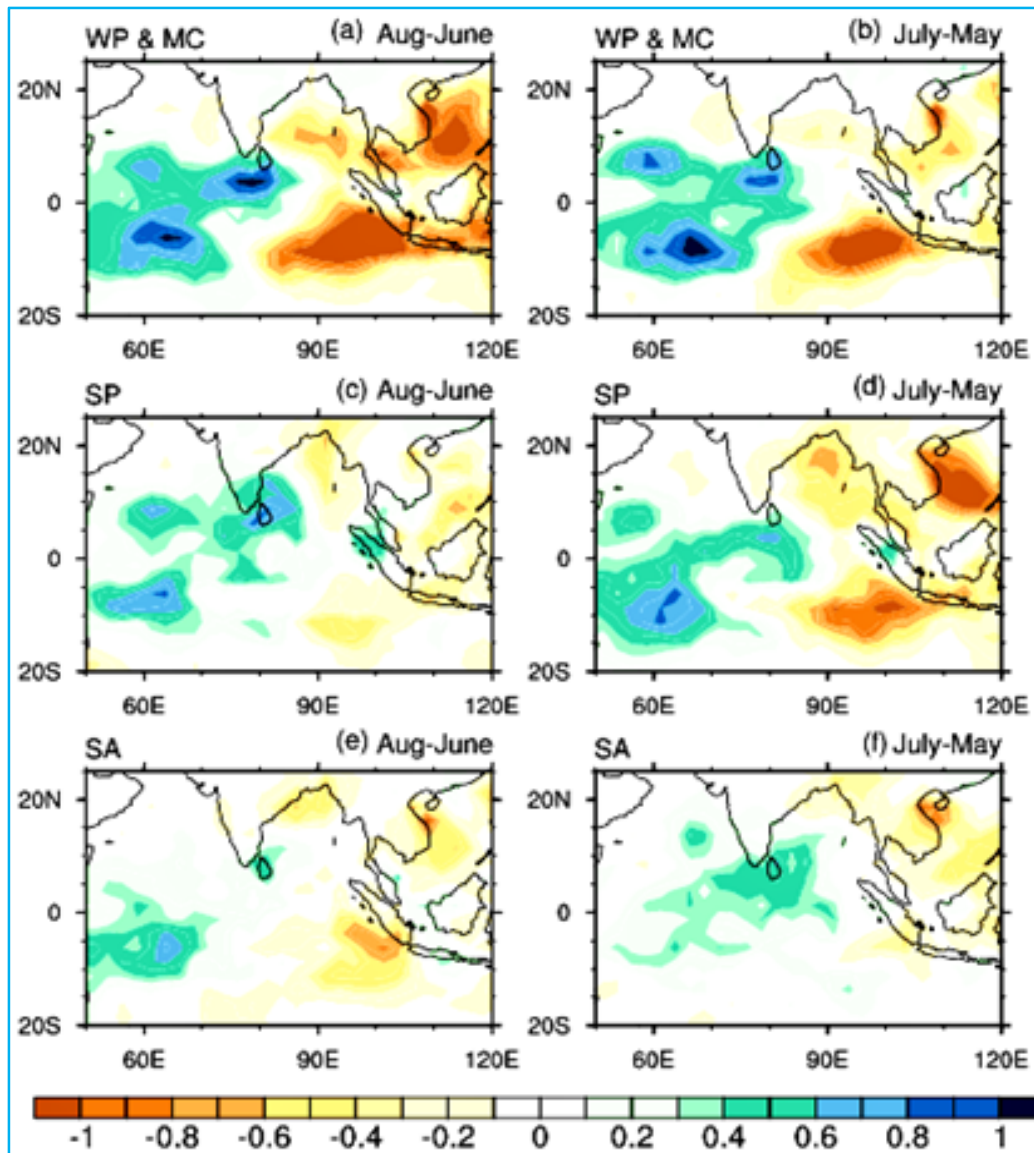
Figs. 4(a-f). Regressed OND precipitation anomalies (mm day^{-1}) against the reversed WP & MC time series (top row), reversed SP time series (middle row) and SA time series (bottom row) are shown at (left column) the 1- and (right column) 2-month lead times

regression coefficient [*i.e.*, -0.357 in Eqn. (1)]. By repeating the same process, the western Pacific and Maritime Continent region (box between 105° E-165° E) is selected as the third predictor (hereafter WP & MC, Table 1). Note that the three predictors are constructed to ensure independence from each other and hence their inter-correlations between the predictors are negligible.

The same procedure is repeated for the 2-month lead time, where the SSTA tendency of the 2-month lead time is defined as July minus May. We note that despite the different lead times, significant correlations of the SSTA

tendencies with SLRI are found over similar regions [Figs. 2(b)]. The exact locations of the three predictors for each lead time are listed in Table 1. We also note that all the variables are normalized so that the regression coefficients in the statistical model represent the relative weighting among the predictors (Lee and Seo, 2013; Kim *et al.*, 2017).

The correlation coefficients between the SLRI and the predictors are summarized in Table 2. The correlation coefficients between the SLRI and normalized SA index are 0.32 and 0.37 at the 1- and 2-month lead times,
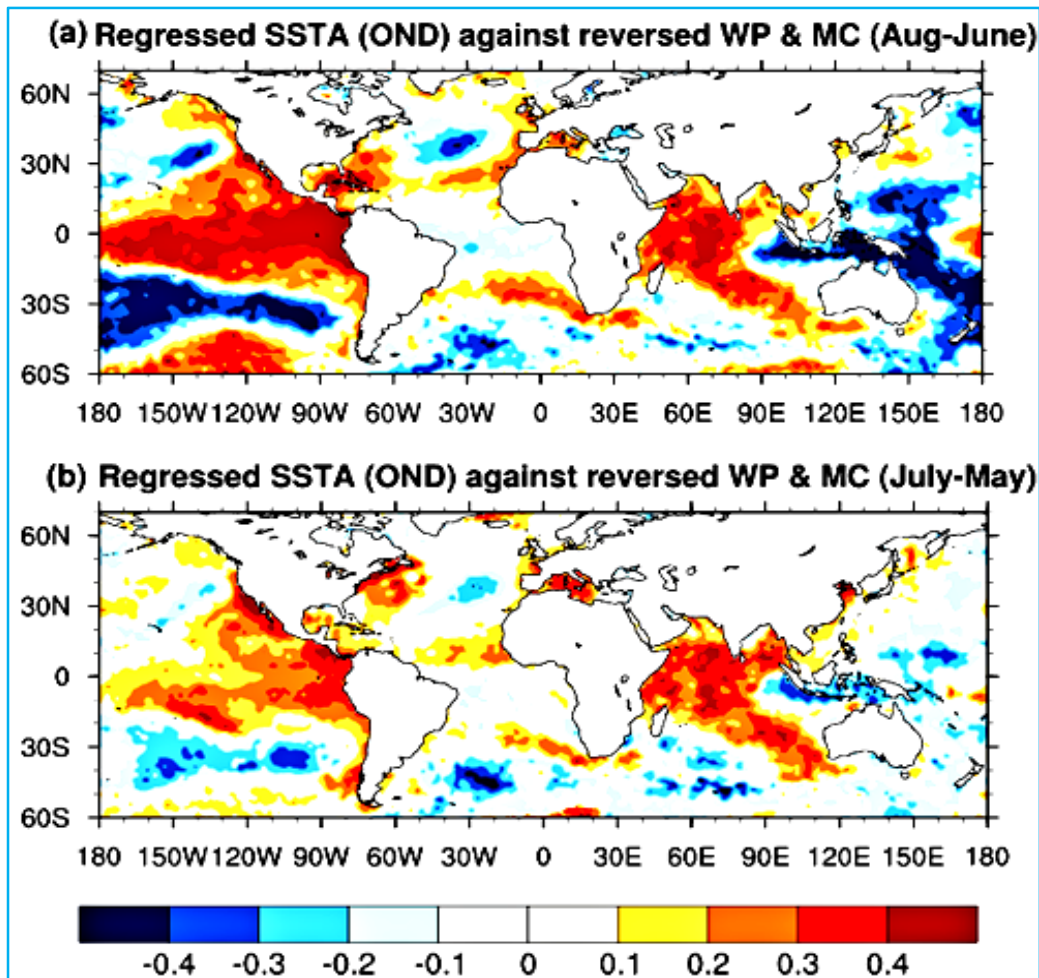
Figs. 5(a&b). Regressed OND SSTAs (°C) against the reversed WP & MC time series at (a) the 1- and (b) 2-month lead times

respectively. After the SA signal is linearly removed from the SLRI, the correlation coefficients between the residual time series of the SLRI and the normalized SP index are -0.47 and -0.40 at the 1- and 2-month lead times, respectively. Lastly, we linearly remove the SP from the residual time series and compute its correlation coefficients with normalized WP & MC index, which are -0.40 and -0.40 at the 1- and 2-month lead times, respectively. The values that exceed the 95% confidence level are marked by asterisks (Table 2).

3.2. *Prediction skills of PLSR forecasts*

Having established the three predictors for each 1- and 2-month lead time, multiple linear regression models are constructed. First, the model is built for the training period, *i.e*., 1979-1999. The equation (1) represents the multiple regression equation for the prediction of Sri Lankan OND rainfall at 1-month lead time.

$$SLRI = 0.292 \,(SA) - 0.357 \,(SP) - 0.411(WP \,\& \,MC) \tag{1}$$

The regression models for the 2-month lead time is formulated as (2):

$$SLRI = 0.446 \,(SA) - 0.302 \,(SP) - 0.409(WP \,\& \,MC) \tag{2}$$

Using the Eqns. (1&2), we perform the seasonal rainfall forecasts for the validation period (2000-2012) at the two lead times (red lines in Fig. 3). The observed SLRI is shown in black lines in Fig. 3. Temporal correlation coefficients are computed between the SLRI and the prediction models and their statistical significance are examined (Table 3). At the 1-month lead time, the correlation coefficient reaches 0.69 for the training period and it is 0.65 for the validation period. Similarly, at the 2-month lead time the correlation
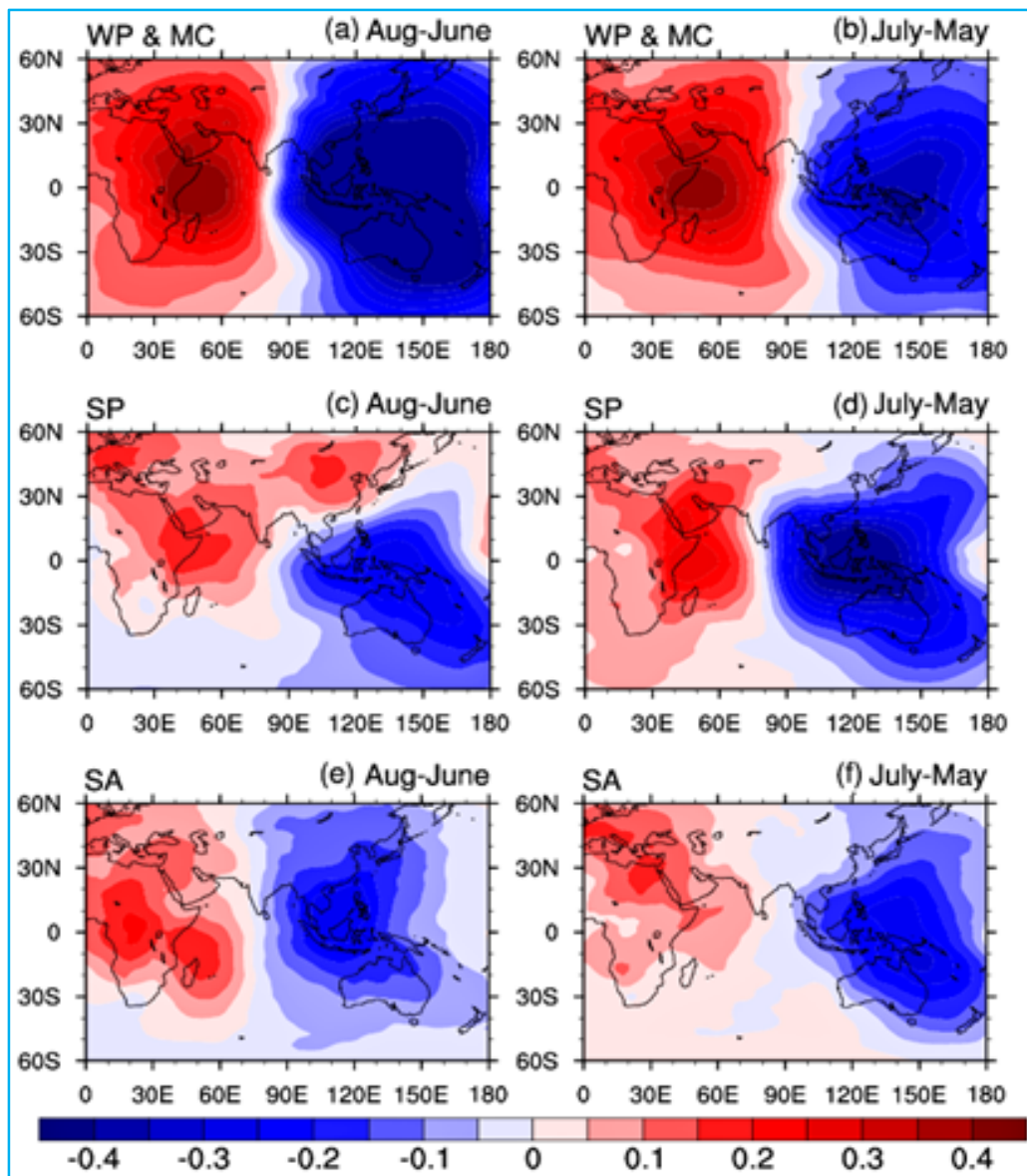
Figs. 6(a-f). Same as Fig. 4, but the regressed OND 850-hPa velocity potential anomalies (10^6 m^2 s^{-1}) are shown

coefficient is 0.68 for the training period and is 0.71 for the validation period.

To further verify the predictive capability of the statistical models, the cross-validation method with three years out scheme is used. That is, as explained earlier, the model is constructed for each year using the entire time series except for the three years centered at the year where the prediction is made. As a result, the predicted time series (blue lines in Fig. 3) are correlated with the SLRI by values of 0.69 and 0.68 at the 1- and 2-month lead times, respectively. These values are very similar to those that are

obtained from Eqns. (1&2). In Fig. 3, one can also notice that the blue and red lines are overlapped by each other.

3.3. *Processes associated with the predictors*

In this subsection, we attempt to understand the physical linkage between the predictors of the Sri Lankan OND rainfall. First to verify the large-scale precipitation pattern associated with the precipitation anomalies over Sri Lanka, we compute the regressed precipitation anomalies against the SA, reversed SP and reversed WP & MC indices for the two lead times (Fig. 4). The signs of
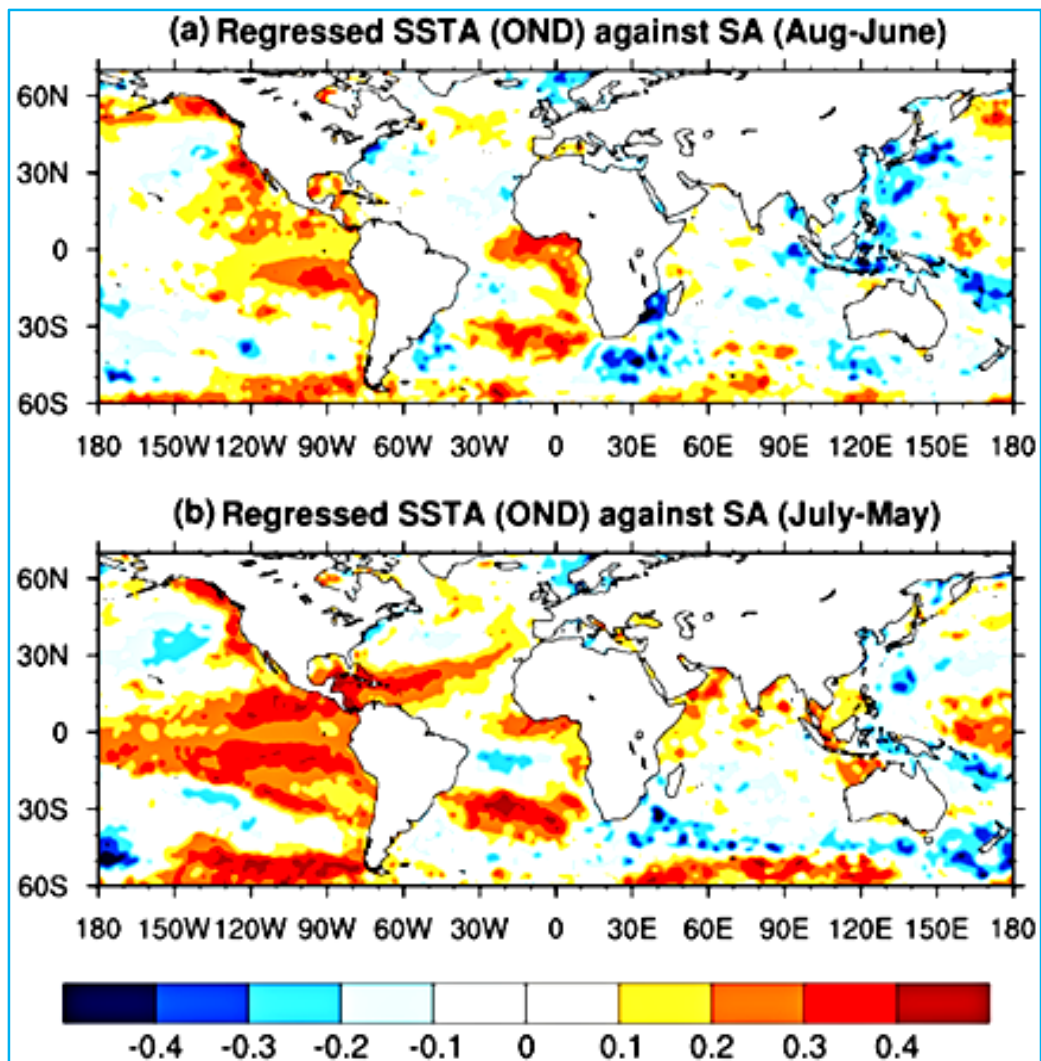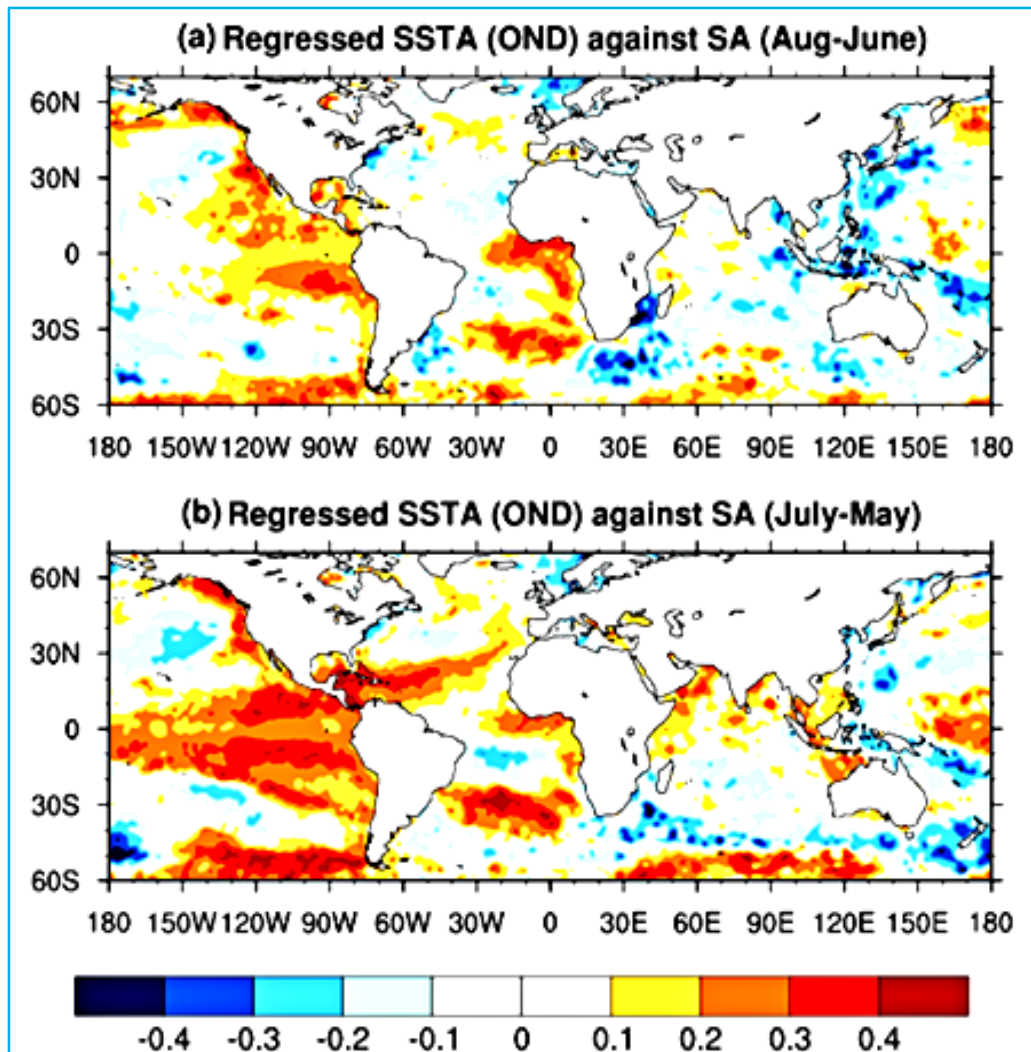
Figs. 7(a&b). Regressed OND SSTAs (°C) against the reversed SP time series are shown at (a) 1- and (b) 2-month lead times

the SP and WP & MC are reversed in order to express their contribution to the regression models as positive values. For all the indices, positive precipitation anomalies are apparent over the western equatorial Indian Ocean, as well as over Sri Lanka, at 1- and 2- month lead times. Negative anomalies are seen at the same time over the Maritime Continent and the Bay of Bengal. This suggests that a dipole-like precipitation pattern, which is influenced by the SSTs over the selected regions, plays an important role on Sri Lankan OND rainfall. Therefore, the physical linkage can be examined by investigating the processes that induce the large-scale dipole pattern of precipitation over the Indian Ocean.

Regressed OND SSTAs against the reversed WP & MC index similarly illustrates the dipole pattern with the

warm anomalies over Indian Ocean and cold anomalies over the Maritime Continent and western Pacific [Fig. 5(a)]. The SSTAs retain their structure throughout the lags, contributing to a continuous warming and cooling at the corresponding locations [Fig. 5(b)]. The widespread warm SSTA scan act as a forcing that favors deep convection over Indian Ocean. Also, the SST gradient between warm north Indian Ocean and cool western Pacific induces easterly winds over the Bay of Bengal with favor moisture transport toward Sri Lanka. The positive velocity potential anomalies, which are also obtained through the regression analysis against the reversed WP & MC index, represent large-scale lower level convergence over the western Indian Ocean when the SSTs over the Maritime Continent and western Pacific decreases [top row in Figs. 6(a&b)]. The velocity potential

Figs. 8(a&b). Regressed OND SSTAs (°C) against the SA time series are shown at (a) the 1- and (b) 2-month lead times

	IOD	Niño3.4	Niño3.0
SLRI	0.39*	0.33	0.32

* Significance correlation at 95% confidence level.

shows a baroclinic structure with a negative anomaly at the upper troposphere (not shown), indicating the upper level divergence. Over the Maritime Continent and western Pacific, the sign for the velocity potential is reversed, exhibiting the dipole structure that has previously been shown in the large-scale precipitation pattern (Fig. 4), as well as the SSTAs (Fig. 5). The strong lower-level convergence anomaly over the western Indian Ocean and the strong lower-level divergence anomaly over the western Pacific are consistent with the strong positive and negative SSTAs over the respective regions.

Despite its smaller amplitude compare to that for the WP & MC, the regressed OND SSTAs against the SA and reversed SP indices show positive anomalies over Indian Ocean [Figs. 7&8(a&b)]. Similarly, over the Maritime Continent and western Pacific, negative anomalies can be seen, which forms again the dipole-like structure. A positive anomaly can be seen over the western Indian Ocean in the regressed field of the OND 850 hPa velocity potential against the reversed SP index [Figs. 6(c&d)]. As was for the WP & MC, this low-level large-scale

convergence is vertically associated with an upper-level divergence (not shown) and horizontally with a low-level positive velocity potential anomaly over the western Pacific and Maritime Continent. The similar features are observed for the SA index [Figs. 6 (e&f)].

4. Summary and discussion

Sri Lanka receives the highest amount of rainfall during October to December (OND). Our study aims to build an empirical model to predict the Sri Lankan OND rainfall. Through the partial least square regression approach, three predictors of sea surface temperature anomaly tendency fields are over the western Pacific and Maritime Continent, southern Atlantic and southern Pacific. Using the predictors, multiple linear regression models have forecasted the Sri Lankan OND rainfall at 1- and 2- month lead times. The three year out cross validated prediction skill for the period of 1979-2012 reaches 0.69 and 0.68 at the 1- and 2- month lead times respectively. Similar skills can be identified by dividing the entire 34 years into the 21-year calibration periods (1979-1999) and the 13-year verification period (2000-2012).

The physical processes associated with the reversed WP & MC, reversed SP and SA predictors show positive (negative) SSTAs over the western Indian (western Pacific) Ocean with a dipole structure at all three lead times. Low-level convergence and divergence in the warm western Indian Ocean and the cool western Pacific, respectively, tend to induce cyclonic circulations to the east of Sri Lanka and to enhance easterlies in Bay of Bengal, which are likely to increase precipitation in Sri Lanka. These results indicate that all three predictors have a crucial impact on the Sri Lankan OND rainfall.

In this study, the spatial characteristics of Sri Lankan rainfall is not considered when building a model. Although the seasonal prediction model relies on large-scale variability of climate, the local orography of Sri Lanka and climatological circulation makes the south western sector receive the highest rainfall. The annual rainfall of this region exceeds 2500 mm, which separates it from the other regions by 500-1000 mm per year. We plan to use higher resolution data to regionally evaluate our model's performance.

Our study develops a statistical prediction model of seasonally averaged precipitation for a targeted domain, Sri Lanka. As explained above, the predictors of the model are chosen by investigating linear correlations with the predictand. This approach has taken in many previous studies. However, one may instead pursue to find a source of predictability from well-established tropical climate

modes, such as the El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD). To explore possibility of this alternative method, we have checked the correlations between the OND Sri Lankan rainfall and the ENSO and IOD indices (Table 4). The correlation with the IOD is 0.39, which exceeds the 95% confidence level. This implies that the IOD is closely related to the Sri Lankan rainfall variability, which seems to be implicated in the effect of the WP & MC index of the present study. The correlations with the Niño3.4 and Niño3.0 indices are similarly above 0.3. Therefore, both of the climate modes may provide good source of skill for seasonal Sri Lankan rainfall. To do so, however, we feel that one needs to carefully consider the seasonality of the climate modes, as well as their interconnections.

References

Black, J., Johnson, N. C., Baxter, S., Feldstein, S. B., Harnos, D. S. and L'Heureux, M. L., 2017, "The predictors and forecast skill of northern hemisphere teleconnection patterns for lead times of 3-4 Weeks", *Mon. Wea. Rev.*, **145**, 2855-2877.

Blockeel, H. and Struyf, J., 2002, "Efficient algorithms for decision tree cross-validation", *J. Mach. Learn. Res.*, **3**, 621-650.

Charney, J. G. and Shukla, J., 1981, "Predictability of monsoons", *Monsoon Dynamics*, 99-109.

Dee, D. P. and Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer P., Bechtold, P., Beljaars, A. C. M., Van de Berg, L., Bidlot, J., Bormann, N., Delsol C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., HólmE. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, De Rosnay, C. P., Tavolato, C., Thépaut, J. N. and Vitart, F.,2011, "The ERA-Interim reanalysis : Configuration and performance of the data assimilation system", *Q. J. R. Meteorol. Soc.*, **137**, 553-597.

DelSole, T. and Shukla, J., 2002, "Linear prediction of Indian monsoon rainfall",*J. Clim.*, **15**, 3645-3658.

Huffman, G. J., Bolvin, D. T. and Adler, R. F., 1997, "Global Precipitation Climatology Project (GPCP) combined precipitation data set", *Bull. Amer. Meteor. Soc.*, **78**, 5-20.

Kim, J. Y., Seo, K. H., Son, J. H. and Ha, K. J., 2017, "Development of statistical prediction models for Changma precipitation: An ensemble approach", *Asia-Pacific J. Atmos. Sci.*, **53**, 207-216.

Lau, K. M., Kim, K. M. and Yang, S., 2000, "Dynamical and boundary forcing characteristics of regional components of the Asian summer monsoon", *J. Clim.*, **13**, 2461-2482.

Lee, S. E. and Seo, K. H., 2013, "The development of a statistical forecast model for changma", *Weather Forecasting*, **28**, 1304-1321.

Park, H. L., Seo, K. H. and Son, J. H., 2015, "Development of a dynamics-based statistical prediction model for the Changma onset", *J. Clim.*, **28**, 6647-6666.

Prasanna, V. and Yasunari, T., 2008, "Interannual variability of atmospheric water balance over south peninsular India and Sri Lanka during north east monsoon season", *Int. J. Clim.*, **28**, 15, 1997-2009.

Rayner, N. A., 2003, "Global analyses of sea surface temperature, sea ice and night marine air temperature since the late nineteenth century", *J. Geophys. Res.*, **108**, p4407.

Sahai, A. K., Grimm, A. M., Satyan, V. and Pant, G. B., 2003, "Long-lead prediction of Indian summer monsoon rainfall from global SST evolution", *Clim. Dyn.*, **20**, 855-863.

Singh, N. and Sontakke, N. A., 1999, "Post-Monsoon Season Over India", *Int. J. Clim.*, **339**, 309-339.

Yim, S. Y., Wang, B. and Xing, W., 2013, "Prediction of early summer rainfall over South China by a physical-empirical model", *Clim.Dyn.*, **43**, 1883-1891.

Yim, S. Y., Wang, B., Xing, W. and Lu, M. M., 2015, "Prediction of Meiyu rainfall in Taiwan by multi-lead physical - empirical models", *Clim. Dyn.*, **44**, 3033-3042.

Zubair, L. and Ropelewski, C. F., 2006, "The strengthening relationship between ENSO and northeast monsoon rainfall over Sri Lanka and southern India", *J. Clim.*, **19**, 1567-1575.

Zubair, L., Ralapanawe, V., Tennakoon, U., Yahiya, Z. and Perera, R., 2006, "Natural disaster risks in Sri Lanka : Mapping hazards and risk hotspots", *Nat. Disaster Hotspots Case Stud.*, 109-136.

Zubair, L., Rao, S. and Yamagata, T., 2003, "Modulation of Sri Lankan Maha rainfall by the Indian Ocean Dipole", *Geophys. Res. Lett.*, **30**, 2, p1063.