

Prediction Skill of NCEP CFSv2 for Seasonal Precipitation and Surface Air Temperature Forecast over Southeast Asia

(Kemahiran Peramalan NCEP CFSv2 untuk Kerpasan dan Suhu Udara Permukaan Musiman di Asia Tenggara)

SITI AMALIA, FREDOLIN TANGANG, SHEAU TIEH NGAI & LIEW JUNENG*

ABSTRACT

Climate Forecast System version 2 (CFSv2) is the ocean-atmosphere-land coupled model and the latest version of seasonal climate forecast from National Centers for Environmental Prediction (NCEP). This study presents the prediction skill of seasonal precipitation and surface air temperature forecasting from CFSv2 over Southeast Asia. The objective of the study was to verify the prediction accuracy of CFSv2 by quantifying the deterministic quantities in term of correlation coefficients with respect to different lead times and target seasons based on a the 28-year ensemble means (1983/84 - 2010/11) for each variables. Additionally, the prediction skill of 20 sub-regions over Southeast Asia are verified with observation for regional assessment of the accuracy of seasonal precipitation and surface air temperature forecasted by CFSv2. In general, the result showed that the prediction skill of CFSv2 for seasonal precipitation and surface air temperature forecasting is reasonable, especially prediction skill after lead month-0 for all target seasons compared to other lead months. The lowest prediction skill is after lead month-6. Overall, the prediction skill of seasonal surface air temperature forecasting is better than precipitation. Moreover, the result obtained in this study highlights the advantages of using an ensemble technique for seasonal forecasting in Southeast Asia.

Keywords: Ensemble techniques; forecast skill assessment; NCEP CFS; seasonal forecasting; Southeast Asia

ABSTRAK

Sistem Ramalan Iklim versi kedua (CFSv2) merupakan sebuah model laut-atmosfera-darat berpasangan dan versi terkini model peramalan iklim bermusim dari Pusat Kebangsaan untuk Ramalan Sekitaran (NCEP). Kajian ini mengemukakan kemahiran ramalan kerpasan dan suhu udara permukaan bermusim dari CFSv2 di Asia Tenggara. Objektif kajian ini adalah untuk mengesahkan ketepatan ramalan CFSv2 dengan mengukur kuantiti deterministik dalam istilah pekali korelasi berdasarkan analisis terhadap min kelompok 28 tahun (1983/84-2010/11) bagi setiap pemboleh ubah pada masa bulan kedepan dan sasaran musim yang berbeza. Tambahan pula, kemahiran ramalan 20 subkawasan di Asia Tenggara dikesahkan dengan pemerhatian untuk melihat secara terperinci ramalan kerpasan dan suhu udara permukaan bermusim yang telah diramalkan oleh CFSv2. Secara umumnya, keputusan menunjukkan bahawa kemahiran ramalan CFSv2 bagi ramalan kerpasan dan suhu udara permukaan bermusim adalah munasabah terutama sekali kemahiran ramalan selepas 0-bulan kedepan untuk kesemua sasaran musim berbanding dengan bulan kedepan yang lain. Kemahiran ramalan paling rendah adalah selepas 6-bulan kedepan. Secara keseluruhannya, kemahiran ramalan bagi peramalan suhu udara permukaan adalah lebih baik daripada kerpasan. Tambahan pula, kajian ini telah menekankan kelebihan menggunakan teknik kelompok bagi peramalan bermusim di Asia Tenggara.

Kata kunci: Asia Tenggara; NCEP CFS; penilaian kemahiran ramalan; peramalan bermusim; teknik berkelompok

INTRODUCTION

Seasonal forecasting has improved in recent years by using the ensemble forecasting technique. Climatologist, in particular, developed a better seasonal forecast model by grouping a number of forecasts of different initial conditions simultaneously due to socioeconomic demands in terms of climate forecasting accuracy, such as precipitation. Specifically, in this area of research, climate forecast systems are referred to as ensemble climate forecasting, which evaluate and forecast the quality of climate elements like temperature and precipitation in order to enhance forecasting performances. Ensembles are only

finite sets of deterministic forecast realizations that initiated from different primary conditions or are subject to different boundary conditions, thought to represent samples from an underlying flow-dependent forecast probability distribution (Weigel 2012). Weigel (2012) indicated that ensembles are usually interpreted and applied as probabilistic forecasts practically and necessarily involving further statistical assumptions. Some modelling with ocean-atmosphere coupled model systems has improved and updated seasonal prediction systems to robust and trusted second generation by physics improvements and increased resolution (Saha et al. 2006; Yuan et al. 2011), for instance, National Oceanic

and Atmospheric Administration's (NOAA) National Centers for Environmental Prediction (NCEP) Climate Forecast System version 2 (CFSv2) and the European Centre for Medium-Range Weather Forecasts (ECMWF) System 4. There has been a substantial improvement after lead month-1 forecast skill of 2 m temperature (T2m) and precipitation (PRCP) relative to CFSv1 over the conterminous United States (CONUS) (Yuan et al. 2011).

A few years ago, forecast skill of precipitation and sea surface temperature (SST) were examined and systematic biases were found relative to CFSv2 for the Northern Hemisphere winter and Asian summer monsoon over the globe (Kim et al. 2012). CFSv2 has also been found to have a systematic cold bias in the central-eastern equatorial Pacific during summer/fall in SST for the periods 1982-1998 and 1990-2010, respectively (Xue et al. 2013). Jiang et al. (2013) found that CFSv2 increased skill in predicting precipitation and large-scale monsoon circulation features but decreased skill for the South Asian monsoon, although some biases in the CFSv1 (Yang et al. 2008) still exist in the CFSv2, especially the weaker-than-observed western Pacific subtropical high and the exaggerated strong link of the Asian Summer Monsoon (ASM) to El-Nino Southern Oscillation (ENSO).

The questions raised in this paper ask whether the forecast skill (prediction skill) of CFSv2 over Southeast Asia is reasonably good for targeted lead month times in four target seasons. This study only focuses on assessing the accuracy of NCEP CFSv2 of Southeast Asia in term of correlation between CFSv2 with selected observational dataset. Southeast Asia is typically known as a tropical region that has variability in terms of precipitation and surface temperature due to unique famous phenomenon like monsoon and ENSO effects. Thus, this study only focuses on the precipitation and surface air temperature received in this region as forecasted by NCEP CFSv2. Many countries in Southeast Asia depend on seasonal forecasting systems advanced by the seasonal forecast system model development of countries such as the USA in order to help decision makers prepare for unwanted disasters like flood and drought, which are very common in certain seasons. The result of current study can be crucial to enhance the Southeast Asia adaptation capacity in providing timely climate information for disaster mitigation as regional climate variabilities are expected to enlarged under warmer global climate. Pin et al. (2013) showed that the identifying the necessity of climatic research and development avenues as a priority to a developing country as they have limited resources to address the diversity of climate change issues.

MATERIALS AND METHODS

The retrospective forecast (hindcast) data used in this study is the 6-hourly time series from the 9-month runs (CFS Reforecast 'First Look' Time Series from NCEP CFSv2). The reforecast of CFSv2 consists of the NCEP Global Forecast System at T126 ($\sim 0.937^\circ$) resolution, the

Geophysical Fluid Dynamics Laboratory Modular Ocean Model version 4.0 at 0.25° - 0.5° grid spacing coupled with a two-layer sea ice model, and the four-layer NOAA land surface model. CFSv2 is a fully coupled general circulation model (GCM) that provides a period of 28-years from 1982/83 to 2010/11 ensemble hindcast dataset with 24 members (Saha et al. 2006). This dataset is a set of 9-month reforecasts initiated from every 5th day, with four ensemble members (4 cycles) for the period 1982-2010, and the NCEP Climate Forecast System Reanalysis dataset is used as the dataset's initial condition for the atmosphere and ocean (Saha et al. 2010). NCEP compiled the monthly estimates as follows: The retrospective data with initial dates after 7th of that month were used as the ensemble members of the next month for each calendar month (Yuan et al. 2012). For instance, the starting dates for the February ensemble members are January 11th, 16th, 21th, 26th, 31st and February 5th, therefore, the total ensemble for six dates in February is 24. However, only 16 ensemble members of each precipitation (PRCP) and surface air temperature (SAT) variable is chosen in this study during the particular dates from the period 1983/84 until 2010/11, with consideration of lead time of month-0, month-1, month-3 and month-6 (LM0-6) for four seasons: December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON).

The reference data set used for the PRCP of the CFSv2 forecast evaluation is the Global Precipitation Climatology Centre (GPCC) monthly precipitation dataset, available from 1901 until the present and calculated from the global station data (Schneider et al. 2017) and Climate Prediction Centre (CPC) Global Temperature data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (CPC Global Daily Temperature 2018) for SAT (Kim et al. 2012; Yuan et al. 2011). All retrospective forecast data is interpolated first to standard $1.0^\circ \times 1.0^\circ$ resolution (360×180) to match the grids of the reference datasets. It must be noted that the calculation of the datasets between PRCP hindcast and GPCC are only over land grids on monthly time scales. The two variable hindcast datasets of CFSv2 are total precipitation rate ($\text{kg/m}^2/\text{s}$ converted to monthly precipitation mm/month) and surface temperature (from Kelvin to degree Celsius). Then, the correlation coefficients between the CFSv2 (Climate forecasting model) and reference data set selected were computed grid-by-grid to assess the prediction skill of CFSv2 (Figure 1).

RESULTS AND DISCUSSION

The deterministic quantities of the forecast model in this context of study include seasonal prediction skill, which is quantified by calculating each grid cell of the product-moment correlation coefficients between the reference data set and forecast ensemble mean series (hindcast) (Wilks 2006; Yuan et al. 2011). The correlation coefficient between GPCC (CPC) and PRCP (SAT) anomaly of CFSv2 hindcast products is calculated for the ensemble mean

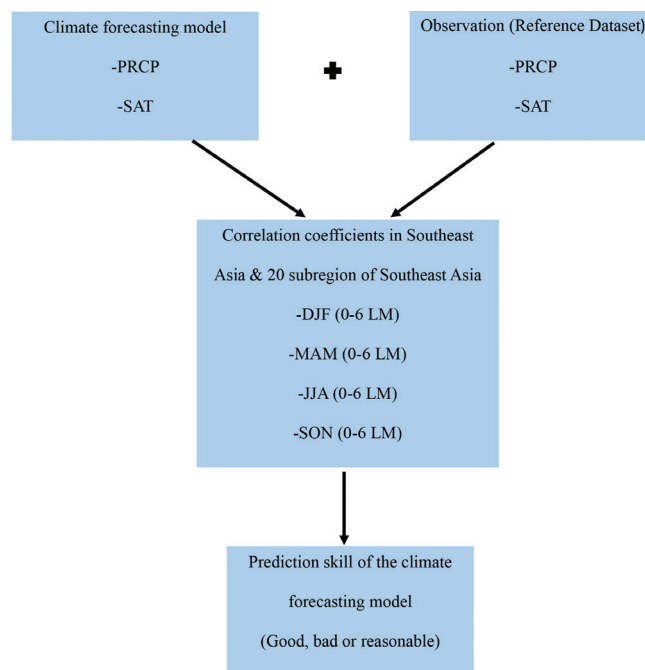


FIGURE 1. The flow chart of research methods

determined from the 28 of each target season (DJF, MAM, JJA and SON). Figure 2 illustrates geographic distributions of prediction skill for CFSv2 after lead times of month-0 (LM0), month-1 (LM1), month-3 (LM3) and month-6 (LM6) for precipitation forecast over land grids for DJF, MAM, JJA and SON. Overall, CFSv2 has high predictive skill in PRCP ($r > 0.20$; refer Figure 2(d), 2(h), 2(l) & 2(p)) in SON, especially over East Malaysia (Sabah and Sarawak), Brunei, Philippines, Cambodia, Laos, Vietnam, Timor-Leste and East Indonesia in August (LM0) (Figure 2(d)). The prediction skill after LM0 is the highest and decreases as the lead times of month increases. However, the prediction skill for other seasons is inconsistent as the lead times of month increases. For instance, prediction skill after LM3 (LM6) is the highest in JJA (MAM) and lowest prediction skill after LM1 in DJF.

Additionally, this study executes the calculation of the product-moment correlation coefficients in 20 sub-regions covering Southeast Asia (Figure 4) based on Ngai et al. (2017) for further review of CFSv2 prediction skill. Figure 3 presents the correlation coefficient for the PRCP anomaly from hindcast after LM0-6 over 20 sub-region land grids in DJF, MAM, JJA & SON. The prediction skill for PRCP in DJF (Figure 3(a)) is greatest (correlation larger than 0.40) over Region 7 (East Indonesia) from hindcast after all lead times of month except LM1. On average, Region 11 (Java) is the lowest skill from hindcast after all month lead times but the lowest prediction skill is after LM3 & LM6 in Region 17-18 (Vietnam-Thailand) and LM1 in Region 8 (West Papua New Guinea). Region 1-2 (Sarawak & Philippines), Region 5 (Borneo) and Region 16-20 (Vietnam, Thailand, Laos, Cambodia and Burma) has inconsistent low prediction skill almost after all lead months. Kim et al. (2012) found wet bias in CFSv2 model

system from hindcast after LM0 in winter season along the South Pacific Convergence Zone (SPZC) and Southern Indian Ocean as well as Western Pacific.

Roughly, Region 9-10 (Timor Leste-East Indonesia), Region 13 (Peninsular Malaysia) and Region 14-15 (North Sumatra-South Thailand) have average prediction skill. Wet bias also commonly found in East Asia and equatorial Atlantic in CFSv2 and SYS4 model systems from hindcast after LM0 in winter season (Kim et al. 2012). The most skillful prediction skill for PRCP in this season is from hindcast after lead times of month-0 followed by month-3 and month-6. Tangang et al. (2012) reported that climate in Malaysia is strongly influenced by natural climate variabilities associated with two large oceans i.e. the Pacific Ocean to the east and the Indian Ocean to the west and there was a shift resulting in the strengthening of the ENSO-Malaysia precipitation relationship and the weakening of the ENSO-Indonesia precipitation relationship (Tangang & Juneng 2004) which could affect the skillfulness of the CFSv2 prediction skills in DJF over these sub-regions in Southeast Asia.

The prediction skill for PRCP in MAM (Figure 3(b)) is greatest (correlation larger than 0.40) over Region 3 (Sarawak & North Kalimantan) from hindcast after LM0 and LM1. Region 15-16 (South Thailand-South Cambodia) has no skill from hindcast after LM3 whilst Region 18 (Laos) after LM0-1 and Region 19-20 (Thailand-Burma) after LM1. Region 6 (Sulawesi) almost for all month lead times. For other regions like Region 7 (East Indonesia), Region 12 (South & West Sumatra) and Region 14 (North Sumatra) have inconsistent low and high prediction skill from hindcast after all lead times of month. Overall, the most skillful prediction skill for PRCP in this season is

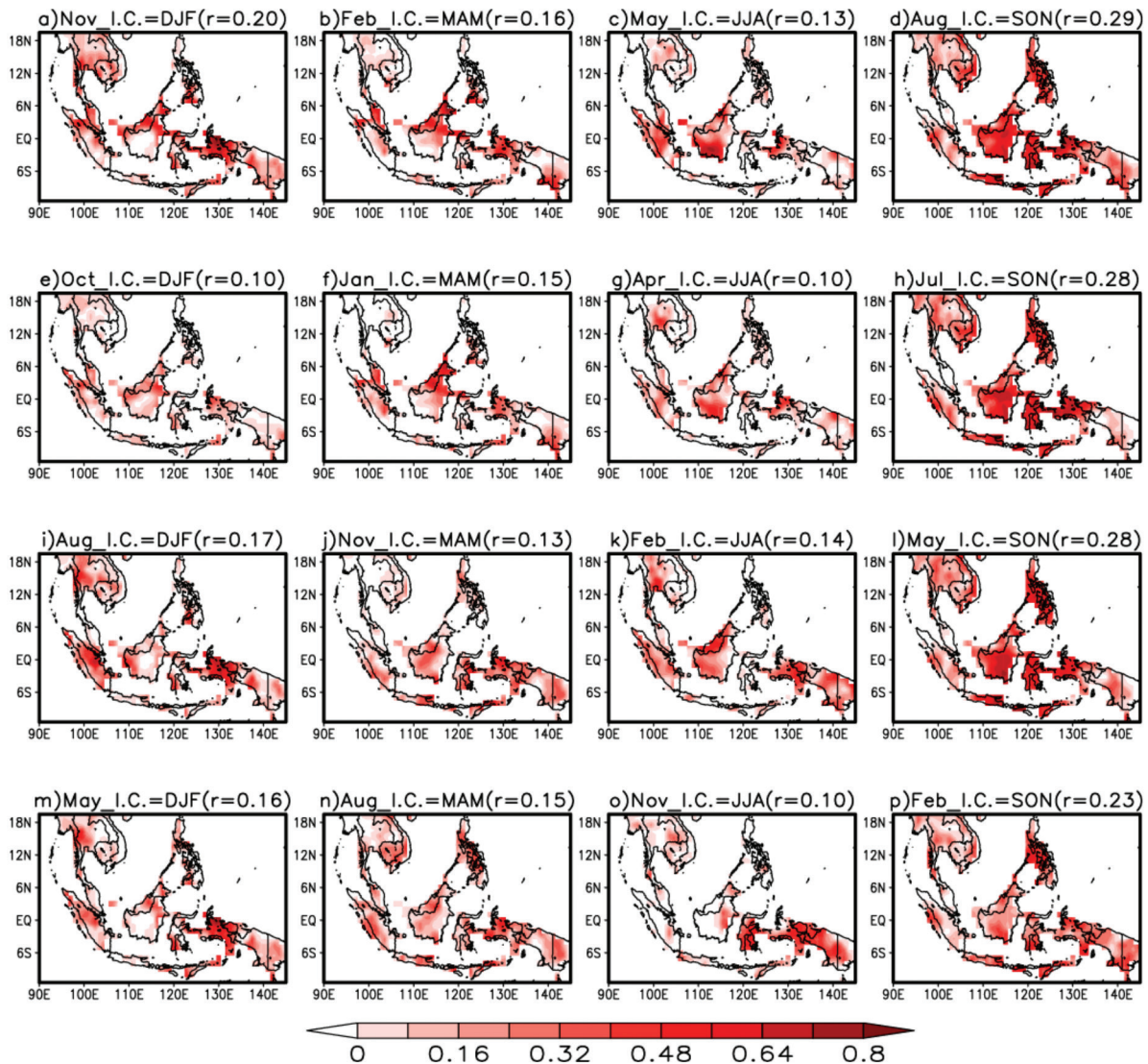


FIGURE 2. The spatial patterns of correlation coefficients between CFSv2 and observed precipitation (PRCP) anomaly during 1983-2010 at (a)-(d) LM0, (e)-(h) LM1, (i)-(l) LM3 and (m)-(p) LM6 over land grids for four target seasons and four initial months, shown on the top of each panel. The numbers are the Southeast Asia mean correlations

from hindcast after lead month-0 followed by month-1 time.

The prediction skill for PRCP in JJA (Figure 3(c)) is greatest (correlation larger than 0.30) over Region 4-5 (Sabah-Borneo) from hindcast after LM1 and LM6, Region 8 (West Papua New Guinea) after LM6, and Region 13 (Peninsular Malaysia) after LM0. Region 1 and 2 (Philippines & Sarawak), Region 10-11 (Java & Timor-Leste) and Region 16-20 (Vietnam, Thailand, Laos, Cambodia and Burma) has inconsistent low prediction skill almost after all lead months but Region 20 (Burma) has no skill after LM0-1. Kim et al. (2012) found the summer mean precipitation shows an excessive precipitation along the Inter-Tropical Convergence Zone (ITCZ), equatorial Atlantic, equatorial Indian Ocean and maritime continent. Kim et al. (2012) also determined that both systems of SYS4 and CFSv2 show a dry bias over the East Asia monsoon region and northern part of South America. In SYS4, a

strong dry bias is found over equatorial central Pacific. Asian monsoon region, especially the Indian Ocean, has low skill in precipitation where the correlation coefficients do not exceed the significant confidence level (Kim et al. 2012).

The skill of precipitation prediction in this season often leads to monsoon prediction by the CFS, which is mainly a result of ENSO (Yang et al. 2008). Jiang et al. (2013) proved that the CFS produces weaker-than-observed large-scale monsoon circulation, partially due to the cold bias over the Asian continent, and tends to over emphasize the relationship between ENSO and the Asian monsoon, as well as the impact of ENSO on the Asian and Indo-Pacific climate.

Yang et al. (2008) stated that a higher-resolution version of the CFSv2 (T126) captures the climatology and variability of the Asian monsoon more realistically than the current resolution version (T62-CFSv1). The largest

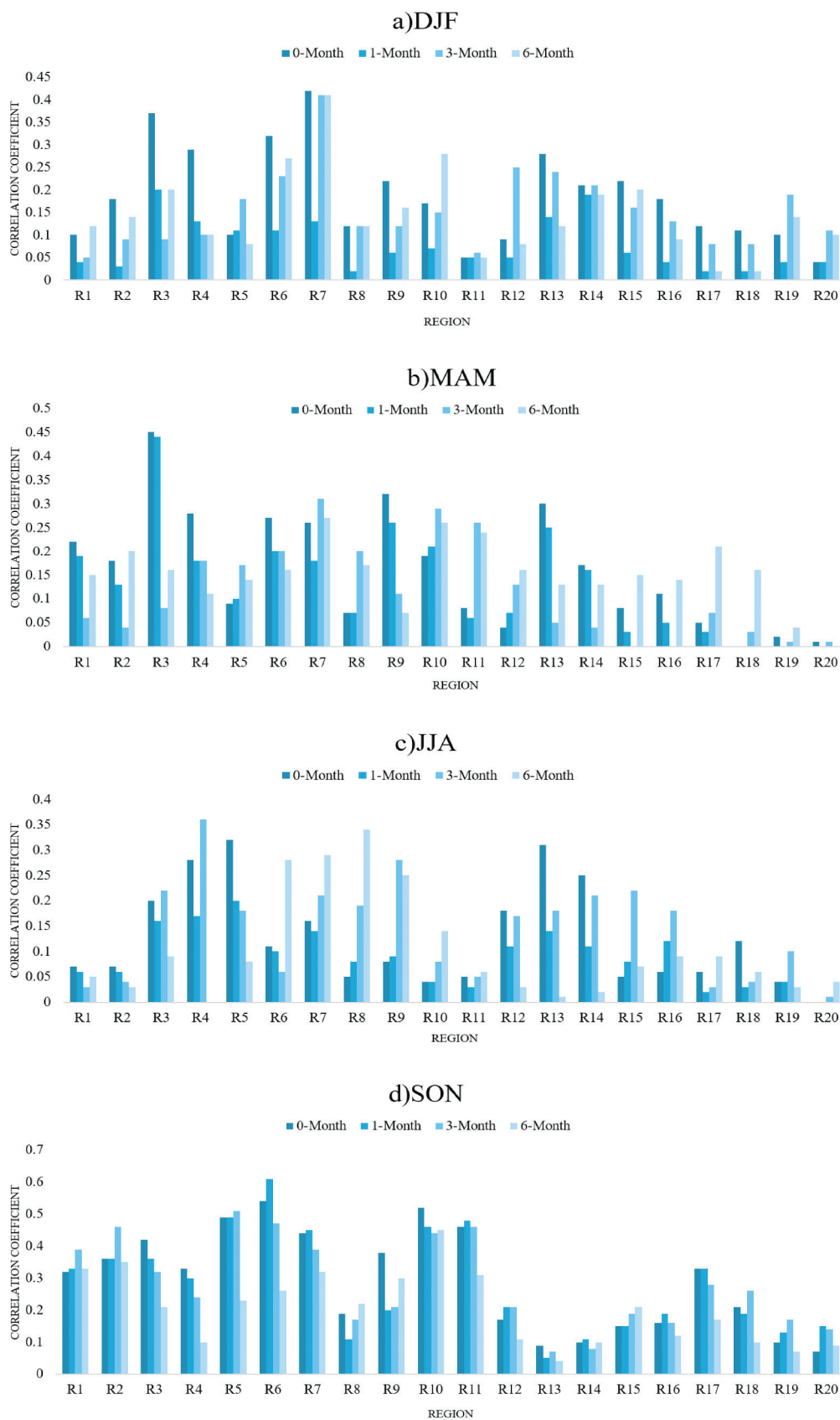


FIGURE 3. CFSv2 and observed precipitation (PRCP) anomaly during 1983-2010 at LM0 (dark blue), LM1 (blue), LM3 (light blue) and LM6 (pale blue) for four target seasons: (a) DJF, (b) MAM, (c) JJA and (d) SON

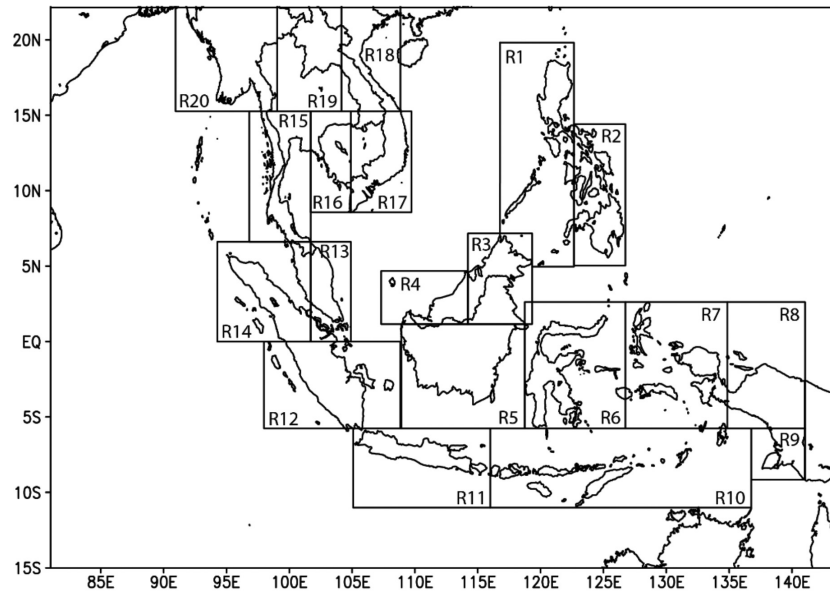


FIGURE 4. The 20 sub-regions that cover the study area (Ngai et al. 2017)

improvement occurs in the simulations of precipitation near the Tibetan Plateau and over the tropical Indian Ocean associated with the zonal dipole model structure (Yang et al. 2008). On average, the most skillful prediction skill for PRCP is from hindcast after lead month-0 followed by lead month-1.

Figure 3(d) presents the prediction skill for PRCP in SON is greatest (correlation larger than 0.45) over Region 5-6 (Borneo-Sulawesi) and Region 10-11 (Java & Timor-Leste) from hindcast after LM0, LM1 and LM3. Region 1-4 (Philippines-North Borneo-Sabah-Sarawak) has average prediction skill from hindcast after all lead times. Region 13-14 (Peninsular Malaysia-North Sumatra) almost has no skill from hindcast after all lead times. Region 15-17 (South Thailand, South Cambodia & South Vietnam) has inconsistent low prediction skill from hindcast after all lead times. The most skillful prediction skill for PRCP in this season is from hindcast after lead time of month-0 followed by month-1 and month-3 for all regions except Region 18-20 (North Thailand-Laos-Cambodia-Burma).

Overall, the correlation coefficient of the PRCP anomaly shows good indication of the actual model skill, denoting the CFSv2 modelling system forecast well. Forecasts from CFSv2 after LM0 for all target seasons are generally skillful compared to other month leads and are not very skillful after LM6. Kim et al. (2012) and Peng et al. (2011, 2000) proved that prediction skill of the CFSv2 modelling system for PRCP at LM0 is much greater over the tropics than over the extra-tropics, and greater over ocean than over land when compared to GPCP. Yang et al. (2008) found that CFSv1 captures the onset of the monsoon (high precipitation) better than the retreat of the monsoon (low precipitation), and it simulates the seasonal march of monsoon precipitation over Southeast Asia more realistically than over South Asia. However, prediction skill for precipitation in CFSv2 for Southeast

Asia is generally lower than 2 metres temperature (2mT), showing greatest skill over the equatorial Pacific due to ENSO (Kim et al. 2012). This could perhaps be due to forecast error including the error caused by uncertainty in the initial state as well as the error caused by model imperfection (Peng et al. 2000) existing in CFSv2 forecasts. Rai and Krishnamurthy (2011) found the growth of errors over the land points of India for precipitation in CFSv1. Yuan et al. (2011) presented obvious enhancement of CFSv2 compared to the previous version of CFSv1 even with fewer grid cells, significant correlation (>0.4) of CFSv2 prediction skill in PRCP after LM1 is demonstrated.

Similar with Figure 2, Figure 5 demonstrates geographic distributions of the LM0-6 prediction skill of CFSv2 in the surface air temperature (SAT) anomaly forecast over land grids for DJF, MAM, JJA and SON. Overall, CFSv2 has high prediction skill ($r = 0.55$; Figure 5(b) & 5(f)) in MAM, especially over Peninsular Malaysia, East Malaysia (Sabah and Sarawak), Brunei, Philippines, Cambodia, Vietnam, and Sumatra in February (LM0 & LM1) (Yuan et al. 2011). The prediction skill at lead month-0 is the highest compared to other lead months, however, the prediction skill in DJF and SON from hindcast after all lead months nearly the same.

Similar to precipitation in Figure 3, this study also executes the calculation of the product-moment correlation coefficients in 20 sub-regions for SAT (Figure 6). The prediction skill for SAT in DJF (Figure 6(a)) is greatest (correlation larger than 0.60) over Regions 3-5 (Sabah-Sarawak-Borneo) and Region 11 (Java) from hindcast after all month lead times. This is comparable to the high prediction skill of CFSv2 over Eastern Asia and Southern Asia in November (LM0) for DJF, as shown by Yuan et al. (2011). Region 7 (East Indonesian) has an inconsistently low prediction skill from hindcast after all month lead times. Interestingly, prediction skill after

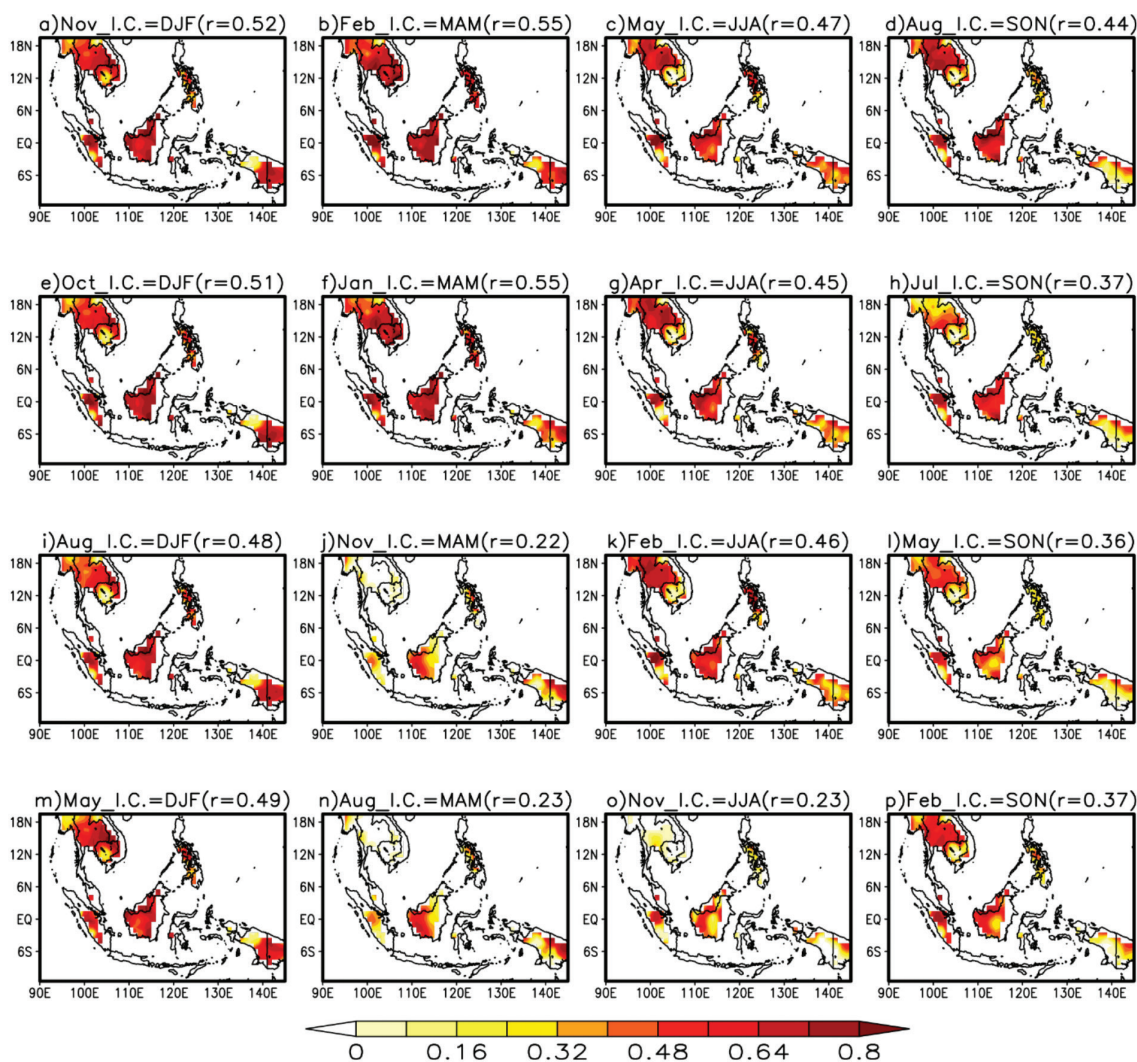


FIGURE 5. The spatial patterns of correlation coefficient between CFSv2 and observed surface air temperature (SAT) anomaly during 1983-2010 at (a)-(d) LM0, (e)-(h) LM1, (i)-(l) LM3 and (m)-(p) LM6 over land grids for four target seasons and four initial months, shown on the top of each panel. The numbers are the Southeast Asia mean correlations

all month lead times in this season, is skillful at most regions, especially prediction skill after lead month-6. Figure 6(b) shows the prediction skill for SAT in MAM is greatest (correlation larger than 0.60) over Regions 3-5 (Sabah-Sarawak-Borneo) and Regions 13-18 (Peninsular Malaysia-North Sumatra-South Thailand-Cambodia-Laos-Vietnam) from hindcast after LM0 and LM1 only. Prediction skill for many of the sub-region in MAM is slightly worst (correlation below than 0.20) compared to other target seasons. Region 1 (East Philippines) and Regions 15-19 (Thailand-Cambodia-Laos-Vietnam) have the worst prediction skill from hindcast after LM3 and LM6 lead times. The most skillful prediction for SAT in this season is from hindcast after lead month-0 followed by lead month-1. Whilst in JJA (Figure 6(c)), the prediction skill for SAT is greatest (correlation larger than 0.60) over Region 4 (Sabah), Region 13-15 (Peninsular Malaysia-North Sumatra-South Thailand) and Region 19 (Vietnam) from hindcast after LM0, LM1 and LM3. The prediction skill for SAT in this season was interestingly skillful after

LM3 whilst is the worst after LM6, especially in Region 9 (West-South Papua New Guinea) and Region 17-20 (Vietnam-Cambodia-Laos-Thailand-South Burma).

Lastly, the prediction skill for SAT in SON (Figure 6(d)) is greatest (correlation larger than 0.60) over Region 4 (Sabah) and Region 14 (North Sumatra) from hindcast after LM0, LM1 and LM3. Region 6 (Sulawesi) has inconsistently low prediction skill from hindcast after all lead time of months. The prediction skill for SAT in this season after LM6 was interestingly skillful compared to LM1, especially in Regions 15-20 (Vietnam-Cambodia-Laos-Thailand-South Burma). Overall, the correlation coefficient of the SAT anomaly shows good indication of the actual model as opposed to PRCP. CFSv2 forecasts SAT better than PRCP to some extent by improved representation of physical processes and data assimilation, statistical or dynamical downscaling techniques (Yuan et al. 2011). Similar to precipitation, forecasts from CFSv2 after LM0 for all target seasons were reasonably skillful compared to other month leads and as predicted, were

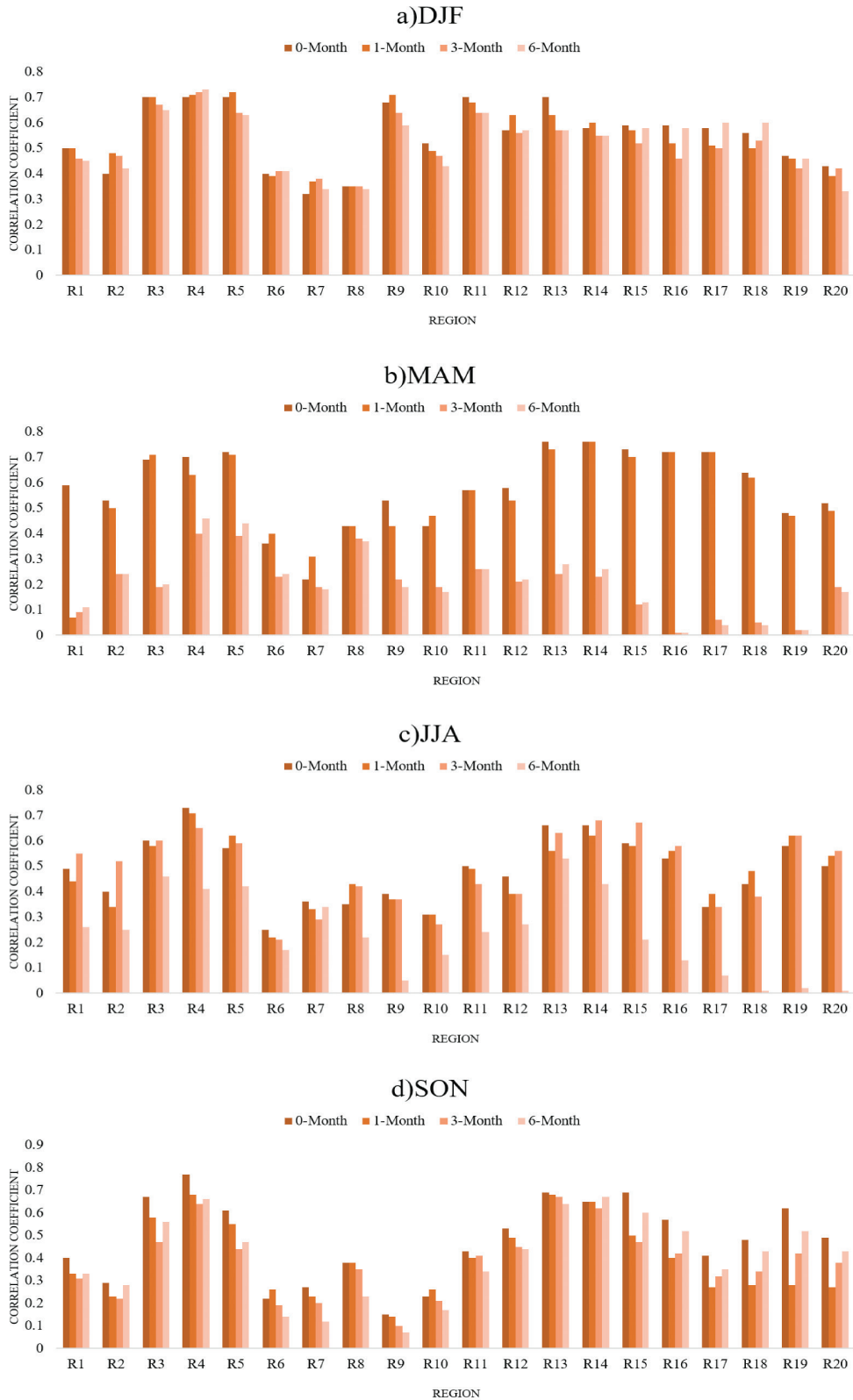


FIGURE 6. CFSv2 and observed surface air temperature (SAT) anomaly during 1983-2010 at LM0 (maroon), LM1 (red), LM3 (light red) and LM6 (pale red) for four target seasons: (a) DJF, (b) MAM, (c) JJA and (d) SON

slightly not skillful after lead month-6 in MAM and JJA. Perhaps, due to the predictability errors in CFSv2 and this will increase as the forecast lead increases (Drbohlav 2010).

CONCLUSION

CFSv2 is a second generation coupled ocean-atmosphere-land model of NCEP that was implemented by an

initialization of a partially coupled ocean, atmosphere, land and sea ice climate reanalysis from 1979, known as CFSR (Saha et al. 2010) producing retrospective forecast data (hindcast) spanning a period of 28 years from 1983/84 to 2010/11 (Saha et al. 2014). This newly-coupled dynamical model improves seasonal climate forecasts with advanced physics, increased resolution and refined initialization. Therefore, the purpose of this study was to evaluate the prediction skill of CFSv2 for different lead time of months in different target seasons by using retrospective predictions. Seasonal predictions after month-0, month-1, month-3 and month-6 lead times for four target seasons (DJF, MAM, JJA and SON) have been investigated with 16 ensembles for 1983-2011. This study examined the seasonal predictive skill of precipitation (PRCP) and surface air temperature (SAT) by calculating the correlation coefficients between observation (reference data set) and the reforecast anomalies (hindcast) for the ensemble mean over 28 years.

CFSv2 shows significant correlation in SON for precipitation, especially over East Malaysia (Sabah and Sarawak), Brunei, Philippines, Cambodia, Vietnam, Timor-Leste and East Indonesia in August (LM0). However, CFSv2 shows better significant correlation in SAT than PRCP especially in MAM over East Malaysia (Sabah and Sarawak), Brunei, Philippines, Cambodia, Vietnam, Timor-Leste and East Indonesia in February (LM0). The prediction skill after LM0 is the highest and decreases as the month lead-time increases for seasonal forecasting variables, precipitation and surface air temperature. However, the prediction skill after LM3 (LM6) is the highest in JJA (MAM) for precipitation.

For better review of CFSv2 prediction skill, this study has calculated the product-moment correlation coefficients in 20 sub-regions of Southeast Asia for both seasonal forecasting variables. The prediction skill for precipitation in DJF is low among other seasons after any month lead times for most of the sub-region, except for Philippines, and high in SON, except for northwest of Southeast Asia (North Sumatra, Peninsular Malaysia, Thailand, Burma, Cambodia, Laos and Vietnam). As studies proved there is systematic biases in CFSv2 during winter, strong wet biases along SPCZ as well as in the Southern Indian Ocean, dry biases over Northern Australia and wet biases in East Asia. Nevertheless, the prediction skill in JJA was only skillful at equatorial, East and South Southeast Asia. Another systematic biases also found during winter, dry bias over the East Asia monsoon region and wet bias (excessive precipitation) along ITCZ, equatorial Atlantic, equatorial Indian Ocean and maritime continent. As for MAM, prediction skill for precipitation was only skillful after LM0 and LM1.

In season DJF, CFSv2 showed better predictive skill for SAT at all month lead times than other target seasons for over 20 sub-regions, but was slightly low in MAM and JJA at LM3 and LM6 based on high mean correlation, r that has been calculated (Figure 6(b) & 6(c)). Interestingly, the prediction skill in SON was more skillful after LM6

than LM1 over Thailand, Cambodia, Laos, Vietnam and South Burma. In season MAM, CFSv2 prediction skill for SAT was only skillful after LM0 and LM1 in most of the sub-regions. Yuan et al. (2011) found that CFSv2 reduced the bias in surface temperature forecasting for August by 53% from CFSv1 (Wang et al. 2010), and the reduction was much higher over high latitudes in Eurasia. Overall, this study has quantified the prediction skill of the most recently upgraded seasonal forecast system from NCEP that can be considered as reasonable since the mean correlation in Southeast Asia calculated show moderate values. Yuan et al. (2011) also stated that the CFSv2 does show promising features even though it has limited skill beyond LM1. The prediction skill of the CFSv2 model is much better than CFSv1, especially strong in surface temperature predictions (Peng et al. 2013). However, some studies suggested that there is a critical issue in evaluating the choice of the reference (observation) dataset used for model prediction skill assessment. Perhaps, future study should examine the sensitivity of the prediction skill by using station data observation rather than station gridded data, and divide the data spanning years into two periods (1983-1999 and 2000-2011) (Peng et al. 2011).

In spite of this, for further study must quantify the systematic bias of CFSv2 and the increase of predictability errors over Southeast Asia by using reference data set in this study as well as quantifying the probabilistic quantities of prediction skills.

ACKNOWLEDGEMENTS

This research was supported by MOSTI through a grant, 04-01-02-SF1377. This support is gratefully acknowledged. The authors also gratefully acknowledge the availability of the forecast information.

REFERENCES

- CPC Global Daily Temperature. 2018. <https://www.esrl.noaa.gov/psd/>. Accessed on 26 September 2018.
- Drbohlav, H.K.L. 2010. Spatial structure, forecast errors, and predictability of the South Asian Monsoon in CFS monthly retrospective forecasts. *Journal of Climate* 23: 4750-4769.
- Jiang, J., Yang, S., Li, Y., Kumar, A., Liu, X., Zuo, Z. & Jha, B. 2013. Seasonal-to-interannual prediction of the Asian summer monsoon in the NCEP climate forecast system Version 2. *Journal of Climate* 26: 3708-3727.
- Kim, H.M., Webster, P.J. & Curry, J.A. 2012. Asian summer monsoon prediction in ECMWF System 4 and NCEP CFSv2 retrospective seasonal forecasts. *Climate Dynamics* 39: 2975-2991.
- Kim, H.M., Webster, P.J., Curry, J.A. & Toma, V.E. 2012. Seasonal prediction skill of ECMWF system 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere winter. *Climate Dynamics* 39: 2957-2973.
- Ngai, S.T., Tangang, F. & Juneng, L. 2017. Bias correction of global and regional simulated daily precipitation and surface mean temperature over Southeast Asia using quantile mapping method. *Global and Planetary Change* 149: 79-90.

- Peng, P., Barnston, A. & Kumar, A. 2013. A comparison of skill between two versions of the NCEP Climate Forecast System (CFS) and CPC's operational short-lead seasonal outlooks. *Weather Forecasting* 28: 445-461.
- Peng, P., Kumar, A. & Wang, W. 2011. An analysis of seasonal predictability in coupled model forecasts. *Climate Dynamics* 36: 419-430.
- Peng, P., Kumar, A., Barnston, A.G. & Goddard, L. 2000. Simulation skills of the SST-forced global climate variability of the NCEP-MRF9 and the Scripps-MPI ECHAM3 models. *Journal of Climate* 13: 3657-3679.
- Pin, K.F., Pereira, J.J. & Aziz, S. 2013. Platforms of climate change: An evolutionary perspective and lessons for Malaysia. *Sains Malaysiana* 42(8): 1027-1040.
- Rai, S. & Krishnamurthy, V. 2011. Error growth in Climate Forecast System daily retrospective forecasts of South Asian monsoon. *Journal of Geophysical Research* 116(D3).
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M. & Becker, E. 2014. The NCEP climate forecast system version 2. *Journal of Climate* 27: 2185-2208.
- Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.T., Chuang, H.Y., Juang, H.M.J., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G. & Goldberg, M. 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* 91(8): 1015-1057.
- Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., van den Dool, H.M., Pan, H.L., Moorthi, S., Behringer, D., Stokes, D., Pena, M., Lord, S., White, G., Ebisuzaki, W., Peng, P. & Xie, P. 2006. The NCEP climate forecast system reanalysis. *Journal of Geophysical Research* 19(15): 3483-3517.
- Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M. & Becker, A. 2017. Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere* 8(3): 52.
- Tangang, F. & Juneng, L. 2004. Mechanisms of Malaysian rainfall anomalies. *Journal of Climate* 17: 3616-3622.
- Tangang, F., Juneng, L., Salimun, E., Sei, K.M., Le, L.J. & Muhammad, H. 2012. Climate change and variability over Malaysia: Gaps in science and research information. *Sains Malaysiana* 41(11): 1355-1366.
- Weisheimer, A., Doblas-Reyes, F.J., Palmer, T.N., Alessandri, A., Arribas, A., Déqué, M., Keenlyside, N., MacVean, M., Navarra, A. & Rogel, P. 2009. ENSEMBLES: A new multi-model ensemble for seasonal-to-annual prediction-Skill and progress beyond DEMETER forecasting tropical Pacific SSTs. *Geophysical Research Letters* 36: L21711.
- Wang, W., Chen, M. & Kumar, A. 2010. An assessment of the CFS real-time seasonal forecasts. *Weather Forecasting* 25(3): 950-969.
- Weigel, A.P. 2012. Verification of ensemble forecasts. In *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd ed., edited by Jolliffe, I.T. & Stephenson, D.B. United Kingdom: John Wiley and Sons. pp. 141-166.
- Wilks, D.S. 2006. *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Massachusetts: Academic Press.
- Xue, Y., Chen, M., Kumar, A., Hu, Z. & Wang, W. 2013. Prediction skill and bias of tropical pacific sea surface temperature in the NCEP Climate Forecast System version 2. *Journal of Climate* 26: 5358-5378.
- Yang, S., Zhang, Z., Kousky, V.E., Higgins, R.W., Yoo, S.H., Liang, J. & Fan, Y. 2008. Simulations and seasonal prediction of the Asian summer monsoon in the NCEP climate forecast system. *Journal of Climate*. 21: 3755-3775.
- Yuan, X., Wood, E.F., Luo, L. & Pan, M. 2011. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophysical Research Letters* 38: L13405.

Centre for Earth Sciences and Environment
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor Darul Ehsan
Malaysia

*Corresponding author; email: juneng@ukm.edu.my

Received: 1 April 2019

Accepted: 15 August 2019