Evaluation of the NCEP Climate Forecast System and Its Downscaling for Seasonal Rainfall Prediction over Vietnam

TAN PHAN-VAN AND THANH NGUYEN-XUAN
Department of Meteorology and Climate Change, Vietnam National University–Hanoi University of Science, Hanoi, Vietnam

HIEP VAN NGUYEN
Applied Geophysics Center, Institute of Geophysics, Vietnam Academy of Science and Technology, Hanoi, Vietnam

PATRICK LAUX
Institute of Meteorology and Atmospheric Environmental Research, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany

HA PHAM-THANH
Department of Meteorology and Climate Change, Vietnam National University–Hanoi University of Science, Hanoi, Vietnam

THANH NGO-DUC
Laboratory for Remote Sensing and Modeling of Surface and Atmosphere, University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi, Vietnam

(Manuscript received 11 July 2017, in final form 14 February 2018)

ABSTRACT

This study investigates the ability to apply National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS) products and their downscaling by using the Regional Climate Model version 4.2 (RegCM4.2) on seasonal rainfall forecasts over Vietnam. First, the CFS hindcasts (CFS_Rfc) from 1982 to 2009 are used to assess the ability of the CFS to predict the overall circulation and precipitation patterns at forecast lead times of up to 6 months. Second, the operational CFS forecasts (CFS_Ope) and its RegCM4.2 downscaling (RegCM_CFS) for the period 2012–14 are used to derive seasonal rainfall forecasts over Vietnam. The CFS_Rfc and CFS_Ope are validated against the ECMWF interim reanalysis, the Global Precipitation Climatology Centre (GPCC) analyzed rainfall, and observations from 150 meteorological stations across Vietnam. The results show that the CFS_Rfc can capture the seasonal variability of the Asian monsoon circulation and rainfall distribution. The higher-resolution RegCM_CFS product is advantageous over the raw CFS in specific climatic subregions during the transitional, dry, and rainy seasons, particularly in the northern part of Vietnam in January and in the country’s central highlands during July.

1. Introduction

Seasonal forecasts provide weather information in terms of mean states of the atmosphere on seasonal time scales (from months to some seasons ahead), thus providing crucial information for decision-making across different sectors (e.g., Pal et al. 2013; Siegmund et al. 2015). During the recent decades, the demand for seasonal rainfall forecasts has been increasing, and their usage has been established, for instance in water resources, food security, and coastal zone management (Cottrill et al. 2013).

There are two approaches to making seasonal rainfall forecasts: statistical and dynamical. Statistical approaches are used as a result of limited computing resources and...
are based on the relationship between a predictand (e.g., seasonal accumulated rainfall) and one or more predictor variables. The predictors are usually related to slowly varying terrestrial or atmospheric components, such as sea surface temperature (SST), snow cover, and soil moisture (e.g., Shukla and Mooley 1987; Krishnamurti et al. 2002; Klotzbach and Gray 2003; Sahai et al. 2003; Annamalai et al. 2005; Duffy et al. 2006). The Regional Climate Outlook Forums (RCOFs), established by the World Meteorological Organization (WMO 2016), are based on such statistical approaches. These RCOFs operationally provide the so-called consensus-based seasonal climate outlooks in many parts of the world, aiming to better manage climate-related risks and to support decision-making on the regional scale. The Southeast Asian countries also participate in the Association of Southeast Asian Nations (ASEAN) Climate Outlook Forum (ASEANCOF 2016). Both RCOFs and ASEANCOF provide probabilistic predictions (given as tercile probabilities) of seasonal mean rainfall, surface air temperature, and some other parameters.

Other efforts at making statistical seasonal forecasts for the South Asian domain include the study of Singhrattna et al. (2005), which developed a statistical method for summer monsoon rainfall forecasting over Thailand using large-scale ocean, atmosphere, and land variables as predictors. They found that rainfall is predictable from one to two seasons in advance. Similar approaches have been applied for China (Kim and Kim 2010; Ye et al. 2015), Central and South Asia (Gerlitz et al. 2016), South Korea and Japan (Jo et al. 2012; Hwang et al. 2001), and India (Guhathakurta 2008). In all these regions, rainfall amounts mainly originate from the Asian monsoon system.

Apart from the statistical approach, the capital advantage of the dynamical approach is that the predicted rainfall is physically consistent with other climate variables, such as air temperature, humidity, and solar radiation (Doblas-Reyes et al. 2006). In recent years, atmosphere–ocean coupled global circulation models (AOGCMs), as well as regional climate models (RCMs), have been developed and employed as tools for dynamical (seasonal) prediction. An important advance in this direction is the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS), which became operational in 2004 (Saha et al. 2006, 2014). Several studies have been carried out to evaluate the skill of CFS products in subseasonal-to-seasonal predictions of rainfall and monsoon indices worldwide (e.g., Siegmund et al. 2015; Yuan et al. 2011, 2012; Liu et al. 2013; Pattanaik and Kumar 2010). According to Yuan et al. (2012), the CFS has a high degree of skill in forecasting monsoon indices for about 2 weeks in advance. Pattanaik and Kumar (2010) showed that the CFS has good skill in forecasting the monsoon circulation index over India, with high correlation coefficients between observed and predicted values in March–May.

With the rapid development of RCMs and the increased supply of computing resources, seasonal rainfall forecasts based on dynamical downscaling have been widely applied (e.g., Pattanaik et al. 2012; Kim et al. 2012). By using nesting strategies, the spatial resolution of RCMs can be increased to allow for the representation of local features, such as land surface characteristics, steep topographical gradients, and physical small-scale processes related to rainfall generation (Olsson et al. 2015). Consequently, dynamical downscaling using RCMs can better represent surface heterogeneity, topography, and small-scale features, which may lead to improved rainfall predictions compared to those of global models. Although reducing the errors in seasonal rainfall forecasts from both global and regional models is a challenge (Chu et al. 2008; Sohn et al. 2012), it was demonstrated that dynamical downscaling of the CFS can improve the skill of seasonal rainfall forecasts in some regions worldwide (e.g., An et al. 2009; Yuan et al. 2012; Siegmund et al. 2015). Yuan et al. (2012) nested the Weather Research and Forecasting (WRF) Model into the CFS to downscale seasonal predictions of winter precipitation during 1982–2008 over continental China. They found that the downscaling reduced the seasonal mean wet bias by 25%–71%. It is worth mentioning here the Multi-RCM Ensemble Downscaling (MRED) experiment, where seven RCMs were used to downscale CFS seasonal winter reforecasts from 1982 to 2003 (De Sales and Xue 2013; Shukla and Lettenmaier 2013; De Haan et al. 2015). Results showed that the skill of MRED forecasts is moderately higher than that of statistical downscaling.

Located in the eastern part of the Indochina peninsula, the mainland of Vietnam can be divided into seven climatic subregions based on rainfall, temperature, sunshine duration, and radiation features (Nguyen and Nguyen 2013), including northwest (R1), northeast (R2), Red River delta (R3), north central (R4), south central (R5), central highland (R6), and south (R7) (Fig. 1). This climate classification has been widely accepted and used by the Vietnamese climate community (e.g., Phan et al. 2009; Van Khiem et al. 2014; Ngo-Duc et al. 2014; Vu-Thanh et al. 2014). Rainfall regimes in Vietnam are mainly controlled by summer and winter monsoons and tropical disturbances, as well as local conditions such as topography (Ngo-Duc et al. 2013; Nguyen-Le et al. 2014; Ngo-Thanh et al. 2017). The annual rainfall regimes in the northern and southern regions of Vietnam are divided into two distinct seasons: the rainy season during summer (from May to October) and the dry season during winter (from November to April of the next year). In contrast, in central Vietnam, the rainy season occurs during fall/winter with maximum rainfall amounts in October and
November (Phan et al. 2009; Matsumoto 1997; Yokoi and Matsumoto 2008; Nguyen-Thi et al. 2012; Nguyen-Le et al. 2015) mainly because of the interaction between the winter monsoon and the Truong Son mountain ranges, which leads to orographic rainfall on the windward coastal central plain of Vietnam. Over the entire country, rainfall during the rainy season (May–September in the northern and southern regions and August–December in the central regions) normally accounts for about 80% of the total annual rainfall (Nguyen et al. 2014). Previous studies also revealed a significant relationship between the rainfall in Vietnam and El Niño–Southern Oscillation (ENSO) and other large-scale processes. For example, it was shown that central Vietnam is drier (wetter) when the SST over the Niño-3.4 region is warmer (colder) (Yen et al. 2011; Vu et al. 2015; Nguyen-Le et al. 2015). Beside the relationship with ENSO, Chen et al. (2012) demonstrated that heavy rainfall events in central Vietnam are also linked with the intensification (weakening) of the low-level easterlies at 15°N and westerlies at 5°N. Ngo-Thanh et al. (2017) showed that the year-to-year variations of the onset dates and the rainfall amount within the rainy season over the central highlands of Vietnam are linked with the preceding winter and spring SSTs in the central-eastern and western Pacific.

In Vietnam, operational seasonal forecasting is currently based on a statistical forecasting system, which has been applied at the Institute of Meteorology, Hydrology and Climate Change (IMHEN) since the early 2000s. The system uses SST anomalies from different ENSO indices (including the Niño-3, Niño-3.4 and Niño-4) as predictors (Nguyen et al. 2003) to determine seasonal rainfall and temperature anomalies at a 3-month lead time. The operational products are published in the official monthly bulletins (in Vietnamese) of IMHEN. This operational system provides useful early warning information on seasonal rainfall anomalies for Vietnam.

Phan et al. (2014, 2015) were the first to analyze dynamical seasonal forecasts based on downsampling of NCEP CFS products using RCMs for Vietnam. In Phan et al. (2014), the authors performed seasonal predictions of temperature by using the Regional Climate Model version 4.2 to downscale the CFS forecasts (RegCM_CFS). The results showed that the skill of the RegCM_CFS in seasonal temperature predictions was significantly improved with bias correction. The same RegCM_CFS downscaling system was used to detect tropical cyclones (TCs) over the northwestern Pacific and Vietnam’s coastal regions (Phan et al. 2015). The results demonstrated that the RegCM_CFS could reasonably reproduce the general distribution of TC counts as well as TC track patterns.
In this study, the seasonal predictability of rainfall across Vietnam is examined, thereby extending previous studies of temperature and tropical cyclones. For this reason, the original and dynamically downscaled CFS forecasts are considered. Thus, this study provides an essential step to establishing an operational seasonal forecast system with lead times of 1–6 months for decision support across the country. Two research questions are addressed: 1) Can the raw CFS rainfall prediction be used for operational seasonal forecasts across Vietnam? 2) Can a better resolution, obtained by dynamical downscaling, improve the seasonal rainfall forecast?

2. Data and methods

a. Data

This study utilizes data from global seasonal forecast models, atmospheric reanalyses, and gridded and gauge precipitation products. More information about the different data sources is given in the following sections.

1) NCEP CFS

The CFS is a fully coupled atmosphere–ocean–land–sea ice seasonal forecast system, which became an operational global forecast system at NCEP in 2004. Details of this system are provided in Saha et al. (2006, 2014). CFS version 2 (CFSv2) is the current version of the CFS, which provides three product types, namely the CFS hindcasts or retrospective forecasts, the CFS operational forecasts, and the CFS Reanalysis (CFSR). In this study, the CFS hindcast and the CFS operational forecast products are used.

The retrospective 9-month CFSv2 data (hereafter referred to as CFS_Ric) with a horizontal resolution of 1.0° and lead time of up to 9 months for the period 1982–2009 (28 years; Saha et al. 2006, 2014; Siegmund et al. 2015) are available from the NCEP website. The data are provided every 5 days, starting from the first day of the month (i.e., the 1st, 6th, 11th, 16th, 21th, and 26th), with four four cycles per day (i.e., 0000, 0600, 1200, and 1800 UTC). One should note that the four-times-daily CFSv2 hindcast data at standard pressure levels that can be used as boundary forcing for an RCM have only become available in the last couple of years, so it would have been impossible to attempt a CFS reforecast downscaling experiment prior to that. Because of the current available computing resources, in this study, only CFSv2 hindcast data with lead times up to 6 months are used.

The CFS operational forecast data are available from early 2011 with three different lead times of 9 months, 3 months, and 45 days. For the 9-month lead time, the CFS is run every day with four cycles initialized at 0000, 0600, 1200, and 1800 UTC (Saha et al. 2014; Siegmund et al. 2015). Because of limitations associated with computing resources, only the CFSv2 operational forecast data (hereafter referred to as CFS_Ope) from the 0000 UTC cycle with lead times of up to 6 months for every 7 days (four runs per month) are used. The data period comprises 3 years from January 2012 to December 2014. The data are provided at a temporal resolution of 6h and a horizontal resolution of 1.0° on standard pressure levels. The same dataset is also used as the initial and boundary conditions for dynamical downscaling (see section 2b). For evaluation purposes, the daily and monthly values are aggregated from 6-hourly data and then interpolated to appropriate grid cells and station locations.

2) ERA INTERIM REANALYSIS

The ERA interim reanalysis is a global atmospheric reanalysis dataset developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), which covers from January 1989 onward and is continuously updated in near–real time (Dee et al. 2011). The data are provided at different time and horizontal resolutions for various variables on standard pressure levels. In this study, the ERA-Interim data for relative humidity and wind fields on the 925- and 850-hPa pressure levels for the period 1982–2009 are used to validate the CFS hindcast and forecast products.

3) THE GPCC DATASET

The Global Precipitation Climatology Centre (GPCC) dataset is a monthly gridded inland rainfall dataset interpolated from rain gauges worldwide for the period from 1901 to the present. GPCC is provided by the German Weather Service (DWD). In this study, the GPCC rainfall dataset (version 6.0) with a resolution of 0.5° × 0.5° (Schneider et al. 2014) was used to verify the CFS rainfall hindcasts.

4) OBSERVED STATION RAINFALL DATA

In this study, daily rainfall data from 151 meteorological stations distributed across Vietnam (Fig. 1) were collected for the periods 1982–2009 and 2012–2014 and are used for the verification of the rainfall hindcasts and operational forecasts, respectively. The data were provided by the Vietnam National Hydrometeorological Service.

b. Evaluation methods

To investigate the performance of the NCEP CFS and the RegCM-based dynamical downscaling of seasonal rainfall forecasts over Vietnam, the following three experiments were carried out in this study.

In this experiment, the ability of the CFS_Rfc to capture the atmospheric circulation over Southeast Asia and Vietnam was first evaluated by comparing wind and RH fields of the CFS_Rfc with those of the ERA-Interim dataset. Seasonal rainfall predictions of CFS_Rfc monthly values were then verified against station observations and the GPCC data.

In addition, the probability-based tercile forecasts of the CFS_Rfc are also verified. For this purpose, the 33rd and 66th percentiles of the monthly data were computed as in Phan et al. (2014) for the climatic subregions (Fig. 1) using station observations (hereinafter referred to as q33o and q66o) and the CFS_Rfc data (hereinafter referred to as q33m and q66m) for the period 1982–2009. Three categories of below normal (B), normal (N), and above normal (A) were assigned if the CFS_Rfc (observed) monthly rainfall amount was less than q33m (q33o), in between q33m (q33o) and q66m (q66o), or greater than q66m (q66o). The probability of each category in the forecast was then defined by the ratios of \( P(B) = n_B/n \), \( P(N) = n_N/n \), and \( P(A) = n_A/n \), where \( n_B \), \( n_N \), and \( n_A \) are the numbers of samples satisfying categories B, N, and A, respectively, and \( n \) is the total sample size. The maximum of the three probabilities \( \{ P(B), P(N), P(A) \} \) determines the category to which the forecast is assigned; that is, a category is forecasted to have occurred if its probability is the maximum of \( \{ P(B), P(N), P(A) \} \).

2) Evaluation of the CFS Operational Forecasts (2012–14)

To investigate the performance of the CFS_Ope, the spatial distribution of CFS_Ope rainfall was compared to that of the observations. Box plots of the observed data and CFS_Ope for all six lead times were also used. The evaluations are identical for both the CFS_Ope and the downscaling for a fair temporal comparison between the two datasets.

3) Evaluation of the Downscaled RegCM_CFS Operational Forecasts (2012–14)

To get a finer resolution of rainfall forecasts over Vietnam, the Regional Climate Model version 4.2, a hydrostatic and compressible limited-area model (Giorgi and Anyah 2012), was employed to downscale the CFS_Ope (hereinafter referred to as RegCM_CFS). The model was run from January 2012 to December 2014, initialized at 0000 UTC every 7 days (four runs per month) and for lead times of up to 6 months. The configuration of the RegCM_CFS is similar to that used in Phan et al. (2014), which includes the Biosphere–Atmosphere Transfer Scheme (Giorgi et al. 1993), the Community Climate Model version 3 radiation scheme (Giorgi and Mearns 1999), and Grell convective schemes (Grell 1993). The model is run with 18 vertical \( \sigma \) levels and in a 36-km horizontal resolution. The domain size is 145 \( \times \) 131 grid points in the east–west and north–south directions, respectively. One should also note that no nudging technique was employed for the interior RCM domain. Because the RegCM_CFS was initialized 4 times per month, the monthly rainfall forecasts were computed as the average of these four runs (for each lead time).

To evaluate the forecast performance of NCEP CFS rainfall and its downscaling across Vietnam, the model rainfall was interpolated to station locations using the inverse distance weighting method. For the verification of monthly forecasts, statistical criteria including relative mean error (RME), relative mean absolute error (RMAE), and correlation coefficient (CORR) were calculated. To evaluate if the CORR is significantly different from zero, a two-tailed \( t \) test was also applied.

To measure the accuracy of the tercile forecasts, the proportion correction (PC), Heidke skill score (HSS), and Peirce skill score (PSS) (Jolliffe and Stephenson 2003) were examined. The scores are defined as follows:

- proportion correction,

\[
PC = \frac{\sum_{i=1}^{3} n_{ij}}{n}; \tag{1}
\]

- Heidke skill score,

\[
HSS = \frac{n \sum_{i=1}^{3} n_{ij} - \sum_{i=1}^{3} n_{ij} \sum_{j=1}^{3} n_{ij}}{n^2 - \sum_{i=1}^{3} n_{ij} \sum_{j=1}^{3} n_{ij}}; \tag{2}
\]

- Peirce skill score,

\[
PSS = \frac{n \sum_{i=1}^{3} n_{ij} - \sum_{i=1}^{3} n_{ij} \sum_{j=1}^{3} n_{ij}}{n^2 - \sum_{i=1}^{3} n_{ij} \sum_{j=1}^{3} n_{ij}}; \tag{3}
\]

where \( i = 1, 2, 3 \) correspond to the B, N, and A categories of the model forecast (observation); \( n_{ij} \) is the number of samples where the model forecast is category \( i \) and the observation is category \( j \); \( n_{ij} = \sum_{j=1}^{3} n_{ij} \) is the total number of samples where the model forecast is category \( i \), and \( n_{ij} = \sum_{i=1}^{3} n_{ij} \) is the total number of samples where the observation is category \( j \); and \( n \) is the total number of samples. While the PC [0; 1] shows the
fraction of the correct forecasts, both the HSS and PSS are in the range of \([-1; 1]\) and show the accuracy of the forecast in predicting the correct category. An HSS (PSS) equal to 0 or negative indicates that the forecast has no skill.

Furthermore, to compare the performance of RegCM_CFS and CFS rainfall forecasts, an added value (AV) of the RegCM_CFS over the CFS was used. The AV is defined as

\[
AV = \frac{|(\text{CFS} - \text{OBS}) - (\text{RCM} - \text{OBS})|}{\text{Max}(|\text{CFS} - \text{OBS}|, |\text{RCM} - \text{OBS}|)}.
\]  

where RCM and CFS, respectively, are the monthly mean of the RegCM_CFS and CFS_Ope rainfall for the period 2012–14; OBS is the corresponding observed rainfall. From (4), the numerator is positive if the CFS error is greater than the RCM error, and vice versa. Because the numerator values largely vary depending on

FIG. 2. Wind vectors (m s\(^{-1}\); vector) and relative humidity (\%; shaded) at 925 hPa during DJF, time averaged for the period 1982–2009 (a) for ERA-Interim and for CFS_Rfc for (b) 1-, (c) 3-, and (d) 6-month lead times (Lt).
the rainfall amount in each month, the denominator, which is always positive, is used for “standardizing” the AV so that the AV values have a range of [−1; 1]. The AV thus can be used to measure the improvement of RegCM_CFS over CFS_Ope on forecasting observed rainfall. A positive value of the AV means that the RegCM_CFS forecast rainfall is better than that of the CFS_Ope. It should be noted that there are several other ways to define the added value of RCM downscaling (i.e., Di Luca et al. 2013; De Haan et al. 2015).

3. Results

a. Performance of the CFS hindcasts

1) CIRCULATION PATTERNS

The average CFS_Rfc (1-, 3-, and 6-month lead times) and ERA-Interim wind and RH fields at 925 hPa during winter (DJF) and summer (JJA) during the period 1982–2009 are shown in Figs. 2 and 3, respectively. Overall, the CFS_Rfc captures the summer
and winter monsoon circulation and moisture fields well across the region.

During DJF, northeasterly winds prevail over all of the Indochina peninsula and Vietnam. The northern and north-central parts of Vietnam experience a cold and dry period due to the cold surges from the north. The ERA-Interim dataset shows strong northeasterly winds flowing from southern China through the South China Sea (SCS) and bringing moist air with an RH of more than 70% into the mainland of Vietnam, and weak winds with dry air prevail in the western part of Vietnam (Fig. 2a). In general, the CFS_Rfc captures fairly well the observed wind speed and direction, as well as the RH, at all lead times (Figs. 2b–d). Only marginal
differences in wind speed and direction between different lead times are observed. Regardless of some underestimations of RH over the northern and south-central parts of Vietnam at lead times of 3 and 6 months (Figs. 2c,d), the overall pattern of RH for the whole of Vietnam is well captured in CFS_Rfc (Fig. 2).

During JJA, the ERA-Interim results show southwesterly winds, which bring moisture-laden air from the Bay of Bengal and dominate the Indochina peninsula with strong wind speeds in western Vietnam. The southwesterly winds are weakened over the SCS and in the leeside of the mountain range in central Vietnam.

Fig. 5. (left) RME (%) and (right) RMAE (%) of the CFS_Rfc rainfall compared to GPCC for (top)–(bottom) January, April, July, and October over the period 1982–2009 for 1-, 3-, and 6-month lead times.
FIG. 6. (a) Correlation coefficients between (left)–(right) ONI and GPCC rainfall and CFS_Rfc rainfall for 1-, 3-, and 6-month lead times for (top)–(bottom) January, April, July, and October over the period 1983–2009. Grid cells with dots show $P$ values below 0.05.
Fig. 6b. As in (a), but for correlation coefficients between CFS_Rfc rainfall and GPCC rainfall.
Vietnam (Fig. 3a). The above features of winds and moisture in the summer are well represented by the CFS_Rfc at all lead times (Figs. 3b,d). CFS_Rfc also captures well the winds and RH at 850 hPa (figures not shown).

2) RAINFALL

Figure 4 shows monthly rainfall from observations (station observations and GPCC), as well as the CFS_Rfc products averaged over the period 1982–2009 for different lead times in January, April, July, and October. It may be noted that the GPCC rainfall distribution is consistent with that of the station data. Specifically, the GPCC data show local rainfall minima and maxima over central Vietnam during April–July and in October, respectively. The consistency of the GPCC with the observed data implies that the regularly gridded GPCC product can be used to verify CFS rainfall.

Figure 4 also shows that the spatial rainfall distribution of the CFS_Rfc across Vietnam is in good agreement with that of the GPCC. A large rainfall band in Laos, over the windward side of the Truong Son mountain range, is also well captured by the CFS_Rfc. Over Vietnam, the CFS_Rfc captures the local rainfall maxima (minima) in central Vietnam in January and October (April and July). The most significant difference in reproducing seasonal rainfall occurs in April and October over the R4 and R3 regions (Fig. 1), where the CFS_Rfc largely underestimates rainfall. It is well known that April and October are transition months between the winter (summer) and summer (winter) seasons in Vietnam. A potential reason for the underestimation of rainfall during these months is that the model does not well represent the weather patterns in which rainfall is caused by the complex interactions between cold surges from the north and tropical disturbances, such as the ITCZ and tropical cyclones.

Figure 5 presents the RME and the RMAE results of the CFS_Rfc rainfall compared to the GPCC data, in the left and right panels, respectively. It can be seen that the absolute values of RME are relatively small (RME in the range of \( \pm 40\% \)). The CFS_Rfc underestimates the values over most parts of the Indochina peninsula in April, July, and October with RME values in the range from \(-40\%\) to \(-10\%\). On the other hand, rainfall is significantly overestimated in the northern part of Vietnam and also partly in Laos during January and over South China, Laos, and Thailand in April and July (Fig. 5, left). Over Vietnam, the CFS_Rfc rainfall tends to be overestimated during the dry season and underestimated during the wet season. The spatial distribution of the RMAE is also consistent with that of the RME (Fig. 5, right). The RMAE mostly ranges between 20% and 60% in April, July, and October and between 80%...
and 120% in January. RMAE values are smallest in July when the observed monthly mean of the rainfall amount ranges around 250–350 mm over the north and south parts and about 100–150 mm in central Vietnam.

To investigate whether the CFS rainfall is modulated by a large-scale process such as ENSO, the CORR between the observed, the GPCC, the CFS_Rfc monthly rainfall, and several selected ENSO climate indices including the bivariate ENSO time series (BEST), the multivariate ENSO index (MEI), the east-central tropical Pacific SST index (Niño-3.4), the Oceanic Niño index (ONI), and the Southern Oscillation index (SOI) (see https://www.esrl.noaa.gov/psd/data/climateindices/list/ for definitions) were calculated and analyzed. The results obtained (not shown) indicated that positive rainfall anomalies usually occurred during La Niña conditions while negative anomalies occurred during El Niño events, especially in the R5–R7 subregions. In other words, mostly drier (wetter) conditions were observed during El Niño (La Niña) events, in which the effect of El Niño was more pronounced than that of La Niña. Over the R4–R7 subregions, 70%–80% of El Niño cases (60%–80% of La Niña cases) led to drier (wetter) conditions. Meanwhile, in the R1–R3 subregions, the differences between drier and wetter frequencies in both El Niño and La Niña were small. Figure 6a shows the CORRs between the ONI and the GPCC (hereafter called CORR_ONI-GPCC) and between the ONI and the CFS_Rfc (hereafter called CORR_ONI-CFS) monthly rainfall forecasts at different lead times for January, April, July, and October (1983–2009). It can be noted that, while CORR_ONI-GPCC is almost significantly negative (except in July) over the southern part of Vietnam, it is either slightly positive or relatively small over the northern part; meaning that, rainfall over R4–R7 in Vietnam is more associated with ENSO compared with that over R1–R3. The patterns of CORR_ONI-CFS are similar to those of CORR_ONI-GPCC, but their values are much larger (from 0.4 to 0.8 in January and October over the northern part, and from −0.4 to −0.8 in January, April, and October over the southern part). A comparison between Figs. 5 and 6a suggests that the CFS rainfall forecast errors (RME and RMAE) are larger over the areas where the differences between CORR_ONI-CFS and CORR_ONI-GPCC are large. The small differences between CORR_ONI-CFS and CORR_ONI-GPCC that occurred in July over the entire region led to small forecast errors, respectively. In other words, the increase of the CFS_Rfc rainfall forecast errors seems to be partly related to a pseudoenhancement of the correlation between the CFS forecasts and the ENSO events.

The CORRs between the GPCC and CFS_Rfc monthly rainfall forecasts for January, April, July, and
October (1982–2009) at different lead times are shown in Fig. 6b. With the exception of some areas at some lead times at which the correlation coefficients are moderate to high (0.4–0.8), such as in southern Vietnam in April and northwestern Vietnam in January, the correlation coefficients are low (~0.2–0.4) or even negative generally. With 1-month lead time, the correlation coefficients range from 0.4 to 0.6 in R1, R2, and R4–R7 during January and in R4–R7 in April and October. The correlations slightly decrease when increasing the forecast lead times. The lowest values of CORR (from ~0.2 to ~0.4) are found at 6-month lead times in R1–R3 in April. In R7 and R4 the correlation coefficients do not satisfy the 0.05 significant level threshold.

From Figs. 5, 6a, and 6b, it may be noted that the distributions of RME, RMAE, and CORR exhibit rather similar
patterns at different forecast lead times, suggesting the possibility of applying the CFS seasonal rainfall forecasts with the longest lead times (i.e., 6 months).

Station-based RME and RMAE values between the CFS_Rfc and observed rainfall for each climatic sub-region of Vietnam are shown in Figs. 7a and 7b. It can be
seen that the CFS_Rfc overestimates rainfall in January and February and underestimates rainfall for most other months of the year (Fig. 7a). Meanwhile, the RMAE is relatively small (less than 30%) during the wet summer months (May–September), especially in the R1, R4, R6, and R7 subregions, and remarkably increases during the dry winter months (October–February; Fig. 7b). The RMAE is about 30% during the summer months and even less than 20% in R1, R6, and R7, while it ranges between 60% and about 80% during the winter months. Only in a few cases is the RMAE approximately 100%. Notably, both the RME and the RMAE remain nearly constant with different lead times.

In general, the CFS_Rfc can reasonably well capture the observed rainfall during the rainy season over all subregions (RMAE; 20%–30%) and fails to reproduce rainfall during the dry season. Since the rainy season contributes about 80% to the total annual rainfall amount (Nguyen et al. 2014), the higher skill of the CFS_Rfc in rainfall forecasts for rainy seasons is very useful for supporting early warning systems in Vietnam.

Figure 8 presents the box plots of the annual cycles of observed and CFS_Rfc rainfall at up to 6-month lead times over the period of 1982–2009 for each climatic subregion. In all subregions CFS_Rfc is able to represent the annual variation features of the observed rainfall. The CFS_Rfc captures the rainy season in the northern and southern subregions and the transition from the rainy season to the dry season during the fall/winter in the R4 and R5 subregions. In R4 and R5 CFS_Rfc tends to overestimate (underestimate) rainfall during the dry (wet) seasons. In other subregions, the CFS_Rfc tends to underestimate rainfall. Compared to the observed data, the interquantile range (IQR) in the CFS_Rfc is noticeably smaller than that of the observations, especially during the wet months. The IQR values of the observed rainfall during the wet months are about 100 mm and even greater than 200 mm in the R4 and R5 subregions. Meanwhile, the IQRs of the CFS_Rfc range mostly between 50 and 100 mm during the wet months and are quite small during the dry months (Fig. 8). The differences in IQRs between CFS_Rfc and the observed rainfall could be explained by the coarse horizontal resolution of the CFS model, which results in smoothing out the regional heterogeneity and thus being impossible to explicitly simulate the physical mechanisms of local rainfall generation such as convective processes, topography-induced rainfall, etc. Consequently, local variations of rainfall cannot be reproduced in detail.

Figure 9 shows the PC scores of tercile forecasts (i.e., three categories of B, N, and A; see section 2b) computed from observed and CFS_Rfc data at up to 6-month lead times over the period 1982–2009 for each climatic subregion (R1–R7). The PC score for a random forecast is 0.33; thus, the forecast information is only useful if its PC score is larger than 0.33. It can be seen that the CFS_Rfc has no skill in subregions R1 and R2 and for almost all months in R3. The model
has some skill during wintertime in the R4 and R5 subregions and in the summertime in the R6 subregion. As shown in Fig. 8, the IQRs of the simulated CFS_Rfc are too small compared to the observed examples, especially in subregions R1 and R2. This might increase the number of missed cases in the N category and/or false alarm cases in categories A and B, causing a decrease of the PC.

Similar features of the model skills can be seen more clearly in Figs. 10 and 11, which suggest that the HSS and PSS values are only noticeably positive in wintertime in the R4 and R5 subregions and in the summertime in the R6 subregion. There is almost no skill for the R1, R2, and R3 subregions throughout the year.

The HSS and PSS skill scores in R4–R7 are much higher compared to those in R1–R3 suggesting that the CFS is advantageous in tercile forecasts over subregions where rainfall is more strongly associated with ENSO (not shown).

b. Performance of the CFS_Ope seasonal rainfall forecast

In this section, the suitability of the CFS_Ope for seasonal rainfall forecasting across Vietnam is examined. Figure 12 shows the observed monthly rainfall and the CFS_Ope gridded rainfall forecasts during January, April, July, and October at 1–6-month lead times. Generally, there is good agreement among the rainfall patterns between the CFS_Ope and observed data. The

---

**Fig. 12.** Monthly rainfall (mm) for (top)-(bottom) January, April, July, and October averaged over the forecast period 2012–14 for (left)-(right) stations and CFS_Ope for 1–6-month lead times.
CFS_Ope captures the location of the rainfall maxima in January and October in central Vietnam and the locations of these maxima in July in the northern and southern parts of Vietnam. The CFS_Ope is also able to capture the rainfall minima in July along the central coast and in October in northern Vietnam (Fig. 12). However, due to the low resolution of the CFS_Ope, the rainfall minima at the central coast of Vietnam seem to be unrealistic compared to the observations. There are no significant differences in the CFS_Ope rainfall among the different lead times. Similar to the CFS_Rfc (see section 3a), the CFS_Ope also underestimates (overestimates) the rainfall during the rainy (dry) season (Fig. 12).

c. Performance of the RegCM_CFS operational seasonal rainfall forecast

1) DOWNSCALED CFS OPERATIONAL RAINFALL FORECAST

The operational seasonal rainfall forecast for Vietnam created by downscaling the CFS_Ope (i.e., RegCM_CFS) for the period 2012–14 is shown in Fig. 13. The RegCM_CFS well captures the observed rainfall patterns over Vietnam with local maxima (minima) occurring in central Vietnam in January and October (July), and drier (wetter) conditions prevailing over the country in January (July). In general, the spatial
rainfall distribution of the RegCM_CFS (Fig. 13) is in good agreement with that of the CFS_Ope (Fig. 12). Moreover, the higher resolution of the RegCM_CFS allows it to represent local rainfall features more realistically and in more spatial detail compared to the forecasts of CFS_Ope. Unlike the CFS_Ope, the RegCM_CFS rainfall forecasts vary significantly with increasing lead times. This “model drift signal” could be due to the fact that the nudging technique was not applied to the interior domain of RegCM in this study; consequently, the large-scale atmospheric circulation of the driven model (CFS) was not retained.

Figures 14a and 14b present box plots of the annual cycles of the CFS_Ope and RegCM_CFS rainfall compared to the observations. In general, both the CFS_Ope and the RegCM_CFS reasonably capture the seasonal rainfall variability over the subregions. However, in comparison to the RegCM_CFS, the CFS_Ope failed in capturing the IQR. During the wet months, values of the CFS_Ope’s IQR are much less than those of the observations. Figure 14b shows that the RegCM_CFS well reproduces the monthly rainfall during the dry season (November–April) in the R3 and R7 subregions, while rainfall is slightly underestimated in R1 and R2. The model output significantly overestimates the rainfall over the R4, R5, and R6 subregions during the dry months from December to March. Although the RegCM_CFS has some biases in monthly rainfall seasonal forecasts, it reasonably reproduces the observed annual rainfall cycle over seven climatic subregions of Vietnam.

2) ADDED VALUE OF THE REGCM_CFS OVER THE CFS_OPE RAINFALL FORECASTS

One of the problems with the CFS products for both the hindcasts (CFS_Rfc) and the operational mode (CFS_Ope) is that their IQRs are reduced compared to the observations (Figs. 8 and 14a). The RegCM_CFS downscaling has some advantages in reproducing this quantity. More specifically, during the rainy season and in the R4 and R5 subregions, the IQR of the RegCM_CFS (Fig. 14b) is much closer to the observed IQR in comparison with that of the CFS_Ope (Fig. 14a). This implies that with a finer resolution, the RegCM_CFS is able to better capture local rainfall maxima than the CFS_Ope does.

Another advantage of RegCM_CFS is that it would provide a more detailed forecast of rainfall over the seven climatic subregions of Vietnam because of its higher resolution compared with that of the CFS_Ope. As mentioned previously, Vietnam has complex terrain; therefore, with the horizontal resolution of 1.0°, the CFS_Ope forecasts could not well capture the detailed spatial distribution of rainfall over Vietnam (Fig. 13). Figure 15a presents the spatial distribution of the station-based AV of the RegCM_CFS over the CFS_Ope rainfall forecasts, averaged for the period 2012–14, in January, April, July, and October and for lead times of up to 6 months. It can be seen that, except over central Vietnam in January and northern Vietnam in October, the AVs are positive at many stations in all months and all lead times, indicating the improvement of the RegCM_CFS compared to the CFS_Ope. More particularly, in January across northern Vietnam, the positive AVs occur at almost all stations (Fig. 15a, top), showing the ability of RegCM_CFS to capture exclusive local weather features in winter that are dominated by the penetration of the cold surges from the north (Fig. 2) and the interaction of the cold surges with complex terrain over the R1–R4 subregions. In the rainy months, the downscaling can also improve the rainfall forecasts over the entire country. In contrast, negative AVs occur at most of the stations in the northern region in October, indicating the weakness of the RegCM_CFS in capturing the rainfall weather patterns during the transition from summer to winter. The RegCM_CFS also fails to improve the rainfall over the central region, including the central highlands and southern Vietnam in January (Fig. 15a). The spatial distribution of AVs in Fig. 15a is partly consistent with that of CORR in Fig. 6b, suggesting that the RegCM_CFS downscaling can improve the rainfall forecast in places with better correlation between the CFS and observations. Furthermore, the RegCM_CFS downscaling can overcome the CFS_Rfc forecast errors caused by the pseudoenhancement of the correlation between the CFS forecasts and the ENSO events [see section 3a(2)].

The areal means of the percentages of stations with positive AV values (MPAV) for each climatic subregion are shown in Fig. 15b. On average, MPAV is about 35%–40% for the entire country. However, it varies remarkably for different lead times and subregions. In the northern part (i.e., R1–R4), the MPAV ranges between 40% and 70% during almost all months (except for October and November), with the highest values (60%–80%) occurring in subregions R1–R3 during DJF. In the southern part (i.e., R5–R7), MPAV varies remarkably from month to month, with the highest values (40%–60%) during the rainy season and the lowest values (below 20%) in January, February, March, and June. It is notable that MPAV is quite high over complex terrain areas during the rainy season, such as from September to November in the south-central region (~40%–60%), and in July.
August, October, and November in the central highlands (~50%–80%).

In comparison with the original CFS forecast, the RegCM_CFS improves the rainfall forecasts in the R5 and R6 subregions in the rainy months in summertime, suggesting that the downscaled CFS can better reproduce the summer monsoon rainfall patterns over the complex terrain areas of Vietnam. Over the northern parts of Vietnam, that is, the R1, R2, and R3 subclimatic regions, the RegCM_CFS significantly improves the rainfall forecasts for almost all months (Figs. 15a,b). At these subregions, the original CFS forecast has almost no skill (Figs. 9–11). So, it can be inferred that the original CFS products can provide useful information for seasonal rainfall forecasts in the regions (R4–R7) where the rainfall patterns are significantly associated with large-scale climatic indices/drivers such as ENSO. In the other regions (R1–R3), the original CFS has less or no skill and the RegCM_CFS downscaling shows its advantages in providing improvements and useful information for seasonal rainfall forecasting.

**FIG. 14.** (a) Box plots of the annual cycles of the observed and CFS_Ope data for 1–6-month lead times over the period 2012–14 for R1–R7.
4. Conclusions

In this study, the CFS_Rfc for the period 1982–2009 was compared with observed and reanalysis data to evaluate its performance across Vietnam. Moreover, three years (2012–2014) of seasonal forecasts of the CFS_Ope and its downscaling using the RegCM4.2 (RegCM_CFS) were also validated against the observations to examine their ability in seasonal rainfall forecasting over the country.

The results from the CFS_Rfc showed that the model was able to capture the monsoon circulation over Southeast Asia and, more specifically, across Vietnam. The model could reproduce the overall rainfall patterns during both summer and winter. The forecast errors of the CFS_Rfc did not differ remarkably with different lead times.

The performance of the CFS_Ope in seasonal rainfall forecasting over Vietnam was generally consistent with that of the CFS_Rfc. The CFS_Ope captured the overall rainfall patterns with local rainfall maxima in January and October in central Vietnam and in July in the northern parts of Vietnam. Similar to the CFS_Rfc,
FIG. 15. (a) AVs of the RegCM_CFS over CFS_Ope for the seasonal rainfall forecasts for Vietnam for (left)–(right) 1–6-month lead times for (top)–(bottom) January, April, July, and October.
the CFS_Ope underestimated (overestimated) rainfall during the rainy (dry) season, which could be partly related to the coarse resolution of the CFS model.

Currently, the ASEAN Climate Outlook Forum provides probabilistic predictions (given as terciles) of seasonal mean rainfall, surface air temperature, and some other parameters, but the spatial and temporal resolutions of the data are still too coarse to satisfy the requirements of seasonal predictions for operational applications in Vietnam. In fact, for early warning and preparedness, water resource management, and planning applications, seasonal predictions at a monthly time resolution and at the provincial level are required. Although the current operational statistical seasonal forecast system in Vietnam can provide some useful information, it still cannot meet the demands of decision-makers. Furthermore, because of the weak relationship between rainfall and the analyzed climate indices, it is difficult to improve the monthly rainfall predictability of this statistical system.

The consistency between the CFS_Ope and the CFS_Rfc thus suggests that the CFS operational monthly rainfall forecasts can provide useful information for operational applications in Vietnam. On the other hand, although the CFS has some skill in seasonal forecasts of monthly rainfall amounts, the forecast errors are still large with values of RMAE that are between 40% and 60%. This should, therefore, be carefully considered when using forecast information of rainfall amounts.

In comparison with the CFS_Ope, the RegCM_CFS provided spatially more detailed information on rainfall over Vietnam. The problem of a noticeably smaller IQR of CFS rainfall could be partially solved in some subregions by downscaling. Because of its finer resolution, the RegCM_CFS was able to better capture the local-scale features of rainfall, especially over the complex terrain areas of Vietnam. Consequently, the rainfall forecasts were improved in some specific climatic subregions over Vietnam during the transitional,

**FIG. 15b.** Percentage of stations with positive AVs (among all stations) for each lead time from 1 to 6 months in R1–R7 and areal means over all six lead times for R1–R7.
dry, and rainy seasons, such as the northern part in January, the central highlands and southern Vietnam in July, and the central part in October. In general, with a higher resolution, the RegCM-CFS better produced rainfall forecasts over Vietnam compared to the CFS-Ope. The result motivates a next study to examine the added values of convective-permitting modeling over this area. Other avenues of exploration such as evaluating the uncertainties from the forcing and land surface models, as well as the effectiveness of statistical/bias correction and/or statistical dynamical downscaling, are also of great interest for future work.

Acknowledgments. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant 105.06-2014.44. The authors sincerely thank three anonymous reviewers, as well as Prof. A. H. Fink and Dr. R. van der Linden from Karlsruhe Institute of Technology (KIT), Germany, for their comments and suggestions, which helped us to improve the quality of our manuscript. HPT would like to thank the German Federal Ministry of Education and Research (BMBF) for support in the context of the LUCCI project (Grant 01LL0908C). The REMOSAT support from the French consortium of the University of Science and Technology of Hanoi (USTH) is also acknowledged.

REFERENCES


