Evaluation of the forecast skill of North American Multi-Model Ensemble for monthly and seasonal precipitation forecasts over Iran

Amin Shirvani [1], Willem A. Landman [2], Mathew Barlow [3], Andrew Hoell [4]

[1] Department of Water Engineering, Oceanic and Atmospheric Research Center, College of Agriculture, Shiraz University, Shiraz, Iran

[2] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

[3] Department of Environmental, Earth, and Atmospheric Sciences, University of Massachusetts Lowell, Lowell, Massachusetts, USA

[4] NOAA Physical Sciences Laboratory, Boulder, Colorado, USA

*Correspondence: Amin Shirvani, Department of Water Engineering, Oceanic and Atmospheric Research Center, College of Agriculture, Shiraz University, Shiraz, Iran. Email: am_shirvani@hotmail.com; ashirvani@shirazu.ac.ir

ABSTRACT

North American Multi-Model Ensemble (NMME) precipitation forecast skill over Iran is evaluated using Taylor diagrams and ranked probability skill scores (RPSS) as determined over a 29-year test period (1991–2019). The forecast skill for both monthly (October through June for lead-times of 0.5–3.5 months) and seasonal (October–December [OND], January–March [JFM], and April–June [AMJ] for lead-times of 1.5–3.5 months) timescales is evaluated using six NMME models as well as multi-model ensemble means (MMM). The latest versions of these models for forecasting Iran's precipitation have not been evaluated thus far. The Global Precipitation Climatology Center (GPCC) version 2020 dataset is used to verify the models. Among individual NMME models, Geophysical Fluid Dynamics Laboratory-Seamless System for Prediction and Earth System Research (GFDL-SPEAR) has generally the highest forecast skill. Both Taylor diagrams and RPSS of most of the models have indicated that the highest forecast skill is found for the month of November such that the Pearson correlation for both SPEAR and MMM is statistically significant for all lead-times. For both monthly and seasonal timescales, the temporal Pearson correlation (TPC) between the observed and forecasts of the MMM is higher than the TPC of the individual models. The spatial Pearson correlation (SPC) and normalized centred root mean square error (NCRMSE) of the SPEAR is close to MMM, but the normalized standard deviation (NSD) of the SPEAR is closer to one compared to the MMM for months from November to March and two seasons (OND and JFM seasons). The MMM precipitation forecasts are underestimated over the northern regions and Zagros mountains for JFM and OND seasons for both 1.5- and 2.5-month lead-times. The degree to which the forecast skill of MMM is dependent on the El Niño–Southern Oscillation (ENSO) connections with precipitation over Iran is examined. Significant Spearman correlations between simultaneous observed Niño3.4 index and Iran precipitation are found for OND, but not for JFM and AMJ. The MMM reproduces the observed ENSO teleconnections to the tropical Pacific in OND, consistent with forecast skill in that season. However, the MMM also produces forecast skill in JFM and AMJ when the ENSO influence is marginal, showing that ENSO is not the only source of skill in the models.

Keywords: forecast skill, forecasting, Iran, NMME, precipitation

1 INTRODUCTION

Much of Iran has an arid or semi-arid climate with the driest areas over the central, eastern, and southern parts. Ranked as the fourth most natural disaster-prone country in the world (https://reliefweb.int/sites/reliefweb.int/files/resources/AD_2019_Iran_0.pdf), precipitation variability in Iran has a significant impact on agricultural production and water resources. Droughts have been found to be severe in central and Southwest Asia from 1998 to 2001, causing stresses in agriculture, water resources, animal husbandry, and public health in these areas (Agrawala *et al*., 2001; Barlow *et al*., 2002). These droughts have imposed a direct cost of USD $1605 million, equivalent to 30.3% of the total value added to the cropping sector in Iran (Salami *et al*., 2009). Also, the flooding in 2001 and 2002 that followed the drought affected about 1.5 million people in Iran (http://www.emdat.be/). These natural hazards indicate that monthly and seasonal precipitation forecast is vital for planning in this region.

The El Niño–Southern Oscillation (ENSO) has a large contribution to seasonal forecast skill (Goddard and Dilley, 2005; Balmaseda and Anderson, 2009; Weisheimer *et al*., 2009; Landman *et al*., 2012) such that the seasonal prediction skill level is dependent on the strength of the teleconnection between ENSO and seasonal variability of precipitation (Landman *et al*., 2019). The connection between the ENSO signal and January–March (JFM) precipitation in Iran is weak (Nazemosadat and Cordery, 2000; Nazemosadat and Ghasemi, 2004), but is statistically significant for October–December (OND) rainfall season (Nazemosadat and Cordery, 2000). The high frequencies of above-normal precipitation in September–November (SON) over the Middle East is related to El Niño events (Mason and Goddard, 2001). Alizadeh-Choobari *et al*. (2018) reported that both La Niña and eastern Pacific El Niño events are related to annual precipitation in Iran; however, the relationship between central Pacific El Niño and annual precipitation is not statistically significant. In general, the abovementioned studies as well as other previous analysis (Barlow *et al*., 2002; 2016; Mariotti, 2007; Hoell *et al*., 2017; 2018) have indicated that El Niño increases the probability of above-average Southwest Asia precipitation, while La Niña increases the probability of below-average Southwest Asia precipitation.

The forcing of precipitation variability over Southwest Asia, defined as the nations of Pakistan, Afghanistan, Iran, and Iraq, has been studied based on modes of the Atlantic and tropical and Indo-Pacific oceans (Hoell *et al*., 2015). Hoell *et al*. (2015) reported that November through April precipitation over Southwest Asia is significantly correlated with variations in sea surface temperature (SST), which include Pacific decadal variability (PDV), ENSO, and the long-term change of global SST (GSST). However, precipitation during individual months of this 6-month rainfall season is not related to SST signatures that include PDV, ENSO, and GSST in different combinations (Hoell *et al*., 2015).

Atmospheric general circulation models (GCMs) and ocean–atmosphere coupled models, which have been widely used for seasonal forecasting globally, have improved precipitation forecast skill (e.g., Barnston *et al*., 2003; Tippett *et al*., 2003; Barnston *et al*., 2010; Barnston and Mason, 2011; Landman *et al*., 2012; Doblas-Reyes *et al*., 2013; Landman, 2014; Sachindra *et al*., 2014a; 2014b; Ehsan *et al*., 2017). Such models are also useful for forecasting

precipitation over the Middle East (Tippett *et al*., 2003; 2005) and Iran (Shirvani and Landman, 2016; Najafi *et al*., 2021). The North American Multi-Model Ensemble (NMME) is a source of intraseasonal to interannual forecasts (Kirtman *et al*., 2014) and combines several dynamical climate models. The NMME consists of retrospective forecasts, which span the 1982–2010 period, and operational real-time forecasts with up to 9 months lead-time.

NMME forecast skill has been assessed in many parts of the world, including the United States (e.g., Khajehei *et al*., 2018; Slater *et al*., 2019), China (e.g., Ma *et al*., 2015), India (e.g., Cash *et al*., 2019), Pakistan (e.g., Ehsan *et al*., 2020a), and East Africa (e.g., Shukla *et al*., 2019; Ehsan *et al*., 2021). Progress on high-resolution NMME models and hybrid statistical-dynamical approaches have been discussed in Becker *et al*. (2022). In the present study, NMME retrospective and real-time forecasts are combined to evaluate the skill of six individual models as well as the multi-model ensemble mean (MMM) for both monthly (October through June for lead-times of 0.5–3.5 months) and seasonal (October–December [OND], January–March [JFM], and April–Jun [AMJ] for lead-times of 1.5–3.5 months) forecasts over Iran for the period 1991–2019. The latest versions of the NMME models and skill of these models for forecasting Iran's precipitation has not been evaluated thus far. A substantial ENSO influence has been documented in the previous work (Nazemosadat and Cordery, 2000; Barlow *et al*., 2002; Nazemosadat and Ghasemi, 2004; Mariotti, 2007; Barlow *et al*., 2016; Hoell *et al*., 2017; 2018; Alizadeh-Choobari *et al*., 2018). The focus of the current study is to investigate how well that influence translates to MMM forecast skill for Iran, and whether there is MMM forecast skill that is in addition to the ENSO influence. Additionally, the degree to which the forecast skill of MMM is dependent on the ENSO connections with precipitation over Iran is examined.

2 DATA AND METHODS

2.1 NMME data

NMME hindcasts and forecasts (Kirtman *et al*., 2014) for six models are used in this study. The models, whose details are outlined in Table 1, are the National Centers for Environmental Prediction-Climate Forecast System Version 2 (NCEP-CFSv2), National Aeronautics and Space Administration-Goddard Earth Observing System (GEOS) Subseasonal to Seasonal (S2S) prediction System (NASA-GEOSS2S), Ocean-Land-Atmosphere studies (COLA-CCSM4), Geophysical Fluid Dynamics Laboratory-Seamless System for Prediction and Earth System Research (GFDL-SPEAR), Global Environmental Multiscale-Nucleus for European Modeling of the Ocean (GEM-NEMO), and Canadian Meteorological Centre (CanCM4i). These at the latest model versions included in NMME whose forecasts are provided by the NOAA Climate Prediction Center (https://www.cpc.ncep.noaa.gov/products/NMME/). The NMME data are archived by the International Research Institute for Climate and Society. The GFDL-SPEAR retrospective period is started from 1991 and therefore all NMME models have a common period from 1991 to present. In the present study, the end of retrospective period of each model, which listed in Table 1, is used and both retrospective and real-time forecasts are combined to evaluate the NMME models for the period 1991–2019. In this period, for example, the GFDL-SPEAR contains only retrospective forecasts and the anomalies are computed with respect to the period 1991–2019. All NMME retrospective forecasts are bias-

corrected (making use of the hindcasts) using cross-validation (for details how to make bias correction see Kirtman *et al*. (2014) and Kirtman and Min (2009)).

TABLE 1. Six forecast models used in this work

Name	Model	Acronym used	Ensemble members	Period	Native Atm. Res.	Native Ocn. Res.	Prediction length (months)	Reference
NCEP Climate Forecast System, version 2	NCEPCFSv2	CFSv2	28	Hindcast = 1982–2010 Forecast = 2011–present	GFS T126L64	MOM4L40 0.25° Eq.	0.5–9.5	Saha *et al*. (2014)
Goddard Earth Observing System (GEOS) Sub-seasonal to Seasonal prediction (S2S) system	NASAGEOS S2S	NASA	Hindcast = 4 Forecast = 10	Hindcast = 1981–2017 Forecast = 2017–present	GEOS5 AGCM 0.5° L72	MOM5L40 0.5° Eq.	0.5–8.5	Molod *et al*. (2020)
COLA-RSMASCCSM4	Cola-CCSM4	COLA	10	1982–present	CAM4 0.9 × 1.25° L26	POPL60 0.25° Eq.	0.5–11.5	Kirtman *et al*. (2014)
Geophysical Fluid Dynamics Laboratory-Seamless System for Prediction and Earth System Research	GFDL-SPEAR	SPEAR	Hindcast = 15 Forecast = 30	Hindcast = 1991–2020 Forecast = 2020–present	AM4.0 0.5° L33	MOM6L75 0.3° Eq.	0.5–11.5	Delworth *et al*. (2020)
Canadian Meteorological Centre-CanCM4i	CanCM4i	CanCM4i	10	Hindcast = 1981–2018 Forecast = 2016–present	CanAM4 T63L31	CanOM4 L40.94° Eq.	0.5–11.5	Lin *et al*. (2020)
ECCC-GEM-NEMO	GEMNEMO	GEM-NEMO	10	Hindcast = 1981–2019 Forecast = 2019–present	GEM 256 × 128	NEMO 1 × 1 1/3 Eq.	0.5–11.5	Lin *et al*. (2020)

The forecast lead-times used in this study for monthly forecasts are 0.5, 1.5, 2.5, and 3.5 months, and 1.5, 2.5, and 3.5 for seasonal forecasts. For example, if the initial condition is

started in early January 1991 and the target months are January 1991, February 1991, March 1991, and April 1991 the lead-times are considered 0.5, 1.5, 2.5, and 3.5 months, respectively. Also, if the target season is OND 1991 and the initial conditions are started in early October 1991, September 1991, and August 1991 the lead-times are considered $\frac{0.5+1.5+2.5}{3} = 1.5$, $\frac{1.5+2.5+3.5}{3} = 2.5$ and $\frac{2.5+3.5+4.5}{3} = 3.5$ months, respectively. The model data is available on a 1° horizontal resolution, and the number of ensemble members for each model varies from 4 to 30 (see Table 1), of which the ensemble means are used for all models. When a combination of models or multi-model is created, the forecast skill is generally higher than any single model (Smith *et al*., 2013; Saha *et al*., 2014). In the present study, the multi-model mean (MMM) is created by averaging the ensemble mean of these six models. Forecast SST and 200 hPa geopotential height from the multi-model mean are used to evaluate ENSO teleconnections.

2.2 Observational data

Global Precipitation Climatology Centre (GPCC) version 2020 monthly precipitation data (Schneider *et al*., 2011), which is based on quality-controlled data from 67,200 stations world-wide is used to verify model skill. GPCC has been evaluated with observations data over Iran in previous work (Raziei *et al*., 2011; Hosseini-Moghari *et al*., 2018; Fallah *et al*., 2020). GPCC is the most suitable dataset among four gridded precipitation datasets-Climatic Research Unit (CRU), GPCC, PERSIANN-Climate Data Record (PCDR), and University of Delaware (UDEL) (Hosseini-Moghari *et al*., 2018). Fallah *et al*. (2020) evaluated (a) gauge-interpolated datasets (GPCCv8, CRU TS4.01, PREC/L, and CPC-Unified), (b) multi-source products (PERSIANN-CDR, CHIRPS2.0, MSWEP V2, HydroGFD2.0, and SM2RAIN-CCI), and (c) reanalyses (ERA-Interim, ERA5, CFSR, and JRA-55) with gauge observations over the Karun basin in southwestern Iran. They found that overall the GPCCv8 dataset agrees best with the observations. GPCC dataset in comparison with NCEP-NCAR dataset shows a better agreement with rain gauge observations and GPCC can be used for drought monitoring in Iran (Raziei *et al*., 2011). GPCC is at 0.25° × 0.25° spatial resolution and available from the IRI data library. The GPCC monthly data is re-gridded to 1° to match the resolution of the NMME models.

Iran lies within the western Alpine-Himalayan chains with the Alborz and Zagros mountain ranges. These two mountain ranges play a vital role in the Iranian atmospheric systems, consequently influencing the amount and distribution of precipitation over Iran (Shirvani, 2017). Using the GPCC data, Figure 1 depicts the percentage of the climatological mean of monthly precipitation over the study area, taken here as the region bounded by the rectangle 26°–39°N and 44°–62°E. The wet season spans from October through May, with the largest precipitation totals observed from December through March. The beginning of the water year within the study area is in the late September, and according to Figure 1, about 75% of the annual precipitation falls from October to March. OND and JFM seasons are, respectively, considered as autumn and winter seasons in the previous studies (e.g., Nazemosadat and Cordery, 2000; Nazemosadat and Ghasemi, 2004; Shirvani and Landman, 2016).

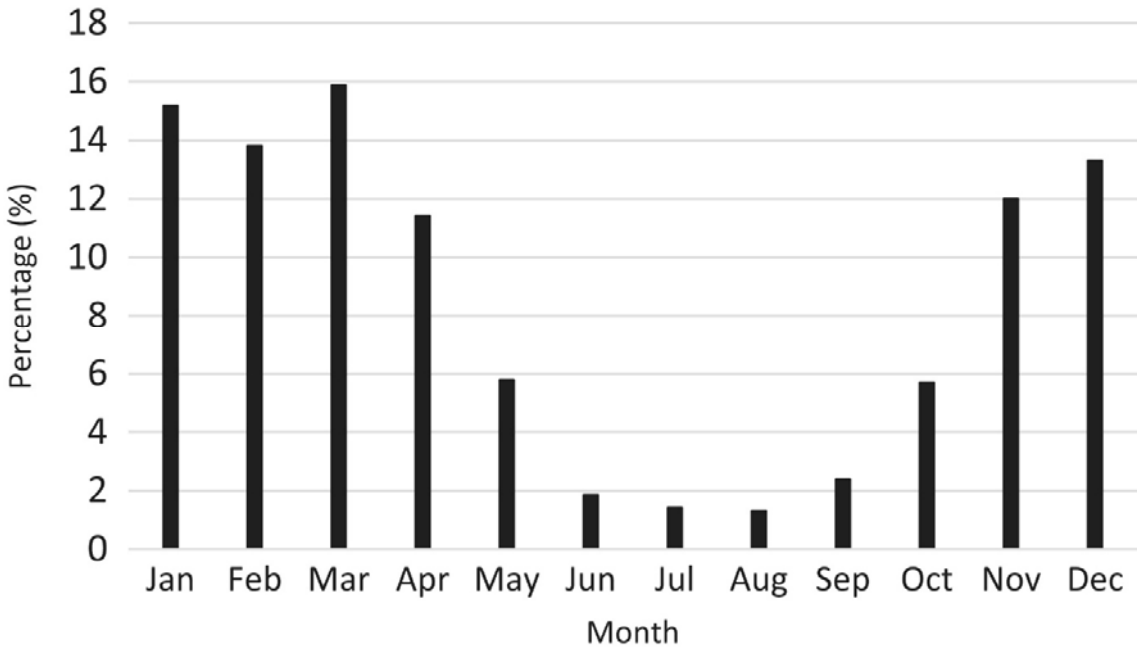

FIGURE 1. The percentage of the climatological mean of GPCC monthly precipitation over the study area for the period 1991–2019

Two hundred hectopascal geopotential height on a fixed 2.5° × 2.5° latitude–longitude grid from the NCEP-National Atmospheric Research (NCAR) reanalysis dataset (Kalnay *et al*., 1996) was used.

2.3 ENSO phase

ENSO phases and the Niño3.4 index, area averaged SST anomalies over 5°S–5°N, 170°–120°W, during 1991–2019 are obtained from the NOAA Climate Prediction Center. These estimates of ENSO characteristics are based on SSTs from the Extended Reconstructed SST Version 5 (ERSST.v5; Huang *et al*., 2017). To assess model performance during strong ENSO phases, La Niña and El Niño years are selected based on the threshold of −/+0.5°C for 3 months running mean Niño3.4 SST anomalies (Table 2).

TABLE 2. La Niña and El Niño years

La Niña	El Niño
1998–1999	1991–1992
1999–2000	1997–1998
2007–2008	2009–2010
2010–2011	2015–2016

2.4 Forecast skill

The forecast skill of the NMME is usually assessed using both deterministic (nonprobabilistic) and probabilistic measures (Kirtman, 2003; Becker *et al*., 2014). In this study, the temporal and spatial Pearson correlation (TPC and SPC; as the deterministic measures) as well as ranked probability skill scores (RPSS; as the probabilistic measure) are used to evaluate the forecast

skills of the NMME over Iran. The spatial Pearson correlation coefficient, normalized standard deviation (NSD), and normalized centred root mean square error (NCRMSE) between the observations and forecasts are calculated and illustrated in a Taylor diagram (Taylor, 2001). Moreover, climatological mean difference (MD) between forecasts and observations are calculated and explained. The teleconnection between regional precipitation and Niño3.4 SST is evaluated using both Pearson and Spearman correlation. Statistical significant of correlation coefficient is evaluated using the Student's *t* test (Wilks, 2011).

The probabilistic nature of ensemble forecast systems requires verification methods based on probabilistic skill scores (Müller *et al*., 2005). The Brier score (BSS) and the ranked probability score (RPS) are widely used to describe the skill of categorical probabilistic forecasts. The RPS, which is an extension of the Brier score to a multi-categories event, is the sum of squared differences between the cumulative forecast and observation categories (Wilks, 2011). A RPSS compares the RPS for the forecast with that of a standard control, or reference, forecasts. The reference forecasts are climatological quantile values of the observed precipitation in this study. The RPSS is zero if RPS for forecast and climatology forecast is equal, indicating no skill. The highest RPSS is 1, indicating perfect skill. The RPSS (Epstein, 1969; Müller *et al*., 2005; Wilks, 2011) is used here to assess NMME forecasts for three equi-probable categories of above-, near-, and below-normal observed precipitation such that the climatological quantile value is 0.33 for each of the three categories. Forecast probabilities of these categories, which are used in computing RPS, are often estimated using the logistic regression model (LRM) (Hagedorn *et al*., 2008; Wilks, 2009). For estimating forecast probabilities, suppose that *p* is the probability forecast and *x* is the single predictor (here, *x* is the NMME forecast), a LRM is expressed as

$$p = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} \tag{1}$$

where β_0 and β_1 are regression parameters (or coefficients). A LRM can be also expressed as a logit function

$$\log\left(\frac{p}{1-p}\right) = \text{logit}(p) = \exp(\beta_0 + \beta_1 x). \tag{2}$$

The unknown parameters of LRM are estimated using an iterative maximum likelihood method (McCullagh and Nelder, 1989; Wilks, 2011). In this study, the generalized linear model (GLM) function within the R software is used for parameters estimation. The cross-validation (Wilks, 2011) ordinal regression model using a one-year-out window is applied for evaluating the selected model.

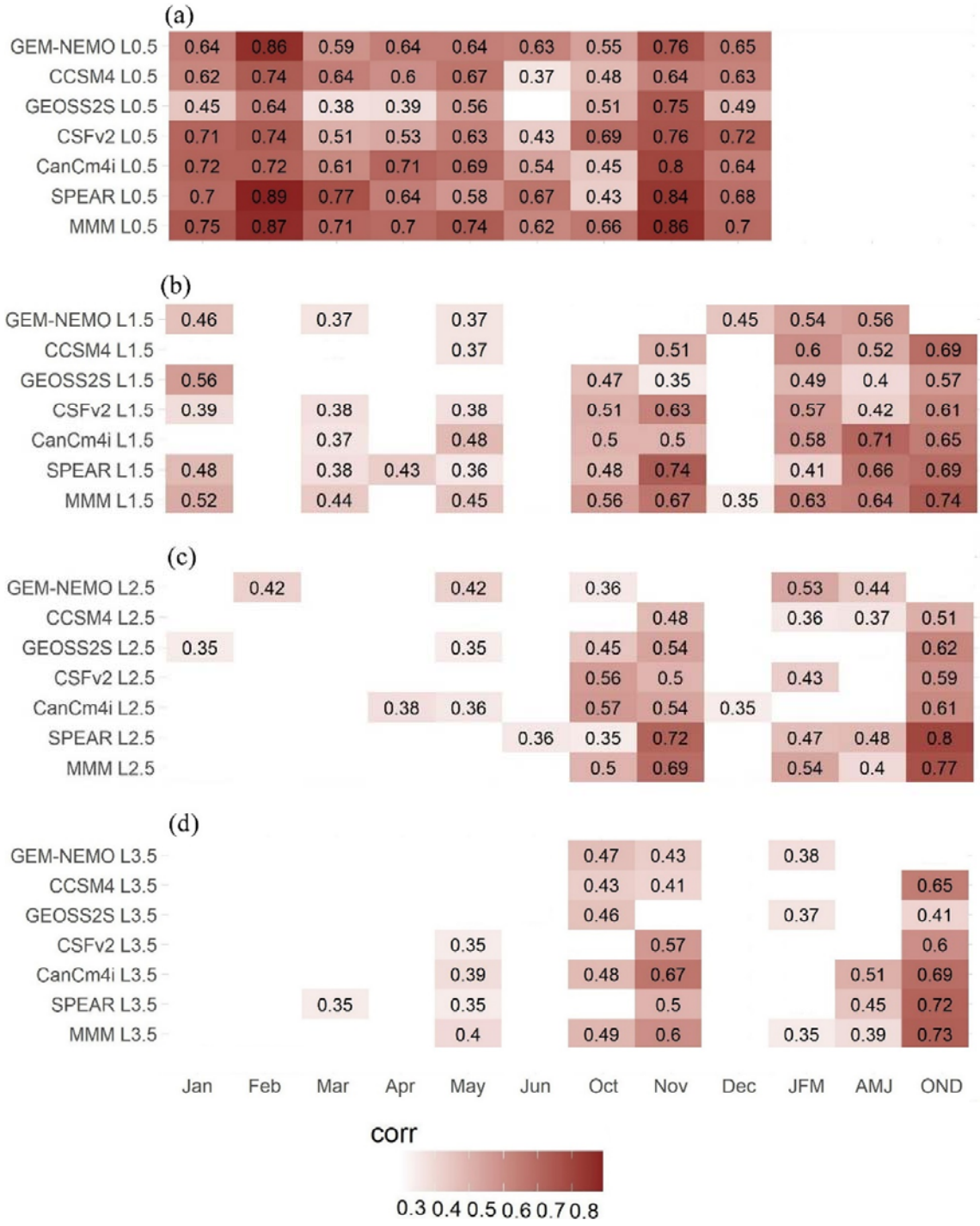

FIGURE 2. The temporal Pearson correlation between time series of the regional averaged observations and forecasts from six NMME models and MMM for monthly and seasonal timescales over the study area. (a) 0.5-month, (b) 1.5-month, (c) 2.5-month, and (d) 3.5-month lead-times. White parts indicate the correlation values do not have statistical significance at the 95% level

3 RESULTS AND DISCUSSION

The time series of the observed and forecasts averaged over the study area are calculated and then the Pearson correlation coefficient between these time series is computed to

provide a temporal Pearson correlation over the whole study area. Figure 2 shows the precipitation forecast skill over Iran in terms of TPC for six NMME models and MMM forecasts for each lead-time and for months (from October to Jun) and for seasons (OND, JFM, and AMJ). The corresponding values of this figure present in Table 3. As expected, higher TPC is generally observed at the shortest lead-times (i.e., 0.5-month lead for monthly and 1.5-month lead for seasonal timescales). SPEAR is generally the best model among single models with the highest TPC. The skill for November forecasts for all models is higher than all other months such that the TPC for this month for both MMM and SPEAR is statistically significant for all lead-times (Figure 2 and Table 3). The TPC of the MMM is statistically significant in January, March, May, October, November, and December at 1.5 month lead-time.

TABLE 3. The temporal Pearson correlation between time series of the regional averaged observations and forecasts from six NMME models and MMM for monthly and seasonal timescales over the study area

Lead	Model	Month									Season		
		Jan	Feb	Mar	Apr	May	Jun	Oct	Nov	Dec	JFM	AMJ	OND
0.5	SPEAR	0.70*	0.89*	0.77*	0.64*	0.58*	0.67*	0.43*	0.84*	0.68*			
	CanCM4i	0.72*	0.72*	0.61*	0.71*	0.69*	0.54*	0.45*	0.80*	0.64*			
	CFSv2	0.71*	0.74*	0.51*	0.53*	0.63*	0.43*	0.69*	0.76*	0.72*			
	GEOSS2S	0.45*	0.64*	0.38*	0.39*	0.56*	0.27	0.51*	0.75*	0.49*			
	CCSM4	0.62*	0.74*	0.64*	0.6*	0.67*	0.37*	0.48*	0.64*	0.63*			
	GEM-NEMO	0.64*	0.86*	0.59*	0.64*	0.64*	0.63*	0.55*	0.76*	0.65*			
	MMM	0.75*	0.87*	0.71*	0.70*	0.74*	0.62*	0.66*	0.86*	0.70*			
1.5	SPEAR	0.48*	0.12	0.38*	0.43*	0.36*	0.11	0.48*	0.74*	0.23	0.41*	0.66*	0.69*
	CanCM4i	0.26	0.30	0.37*	0.29	0.48*	0.15	0.50*	0.50*	0.19	0.58*	0.71*	0.65*
	CFSv2	0.39*	0.26	0.38*	0.17	0.38*	0.00	0.51*	0.63*	0.19	0.57*	0.42*	0.61*
	GEOSS2S	0.56*	0.22	0.03	0.01	0.30	−0.06	0.47*	0.35*	0.15	0.49*	0.40*	0.57*
	CCSM4	0.27	0.15	0.33	0.29	0.37*	0.20	0.19	0.51*	0.20	0.60*	0.52*	0.69*
	GEM-NEMO	0.46*	0.31	0.37*	0.34	0.37*	0.12	0.15	0.25	0.45*	0.54*	0.56*	0.22
	MMM	0.52*	0.30	0.44*	0.34	0.45*	0.10	0.56*	0.67*	0.35*	0.63*	0.64*	0.74*
2.5	SPEAR	−0.08	0.27	0.32	0.27	0.27	0.36*	0.35*	0.72*	0.24	0.47*	0.48*	0.80*
	CanCM4i	0.18	0.04	0.11	0.38*	0.36*	0.20	0.57*	0.54*	0.35*	0.32	0.33	0.61*
	CFSv2	0.23	0.01	0.10	0.02	0.15	−0.06	0.56*	0.50*	0.00	0.43*	0.20	0.59*
	GEOSS2S	0.35*	0.00	0.21	−0.01	0.35*	0.24	0.45*	0.54*	−0.06	0.30	0.23	0.62*
	CCSM4	−0.02	0.15	0.29	−0.13	0.18	0.32	0.05	0.48*	0.11	0.36*	0.37*	0.51*

	GEM-NEMO	0.06	0.42*	0.12	0.06	0.42*	0.04	0.36*	0.30	0.15	0.53*	0.44*	0.26
	MMM	0.13	0.26	0.33	0.19	0.34	0.27	0.50*	0.69*	0.26	0.54*	0.40*	0.77*
3.5	SPEAR	−0.10	0.09	0.35*	0.20	0.35*	0.17	0.34	0.50*	0.09	0.08	0.45*	0.72*
	CanCM4i	0.21	0.07	0.19	0.32	0.39*	0.18	0.48*	0.67*	0.11	0.29	0.51*	0.69*
	CFSv2	0.25	−0.16	0.08	0.08	0.35*	−0.04	0.25	0.57*	0.04	0.14	0.15	0.60*
	GEOSS2S	0.22	0.09	0.00	0.03	0.05	0.01	0.46*	−0.04	0.18	0.37*	0.14	0.41*
	CCSM4	0.15	−0.25	0.20	−0.17	0.32	0.06	0.43*	0.41*	0.29	0.09	0.09	0.65*
	GEM-NEMO	0.14	0.01	0.15	0.01	0.11	0.30	0.47*	0.43*	0.08	0.38*	0.20	0.31
	MMM	0.25	−0.06	0.30	0.20	0.40*	0.11	0.49*	0.60*	0.28	0.35*	0.39*	0.73*

Note: Asterisk (*) indicates significant correlations at 5% significance level.

Figure 2 and Table 3 indicate that the TPC for seasonal timescales is generally higher than for monthly timescales such that the MMM and SPEAR produces a high and significant TPC for all lead-times (except 3.5-month lead-time for JFM SPEAR). The TPC for OND precipitation forecasts for all models (except GEM-NEMO) is higher than JFM and AMJ precipitation forecasts. Overall, for both monthly and seasonal timescales the TPC of the MMM is higher than the TPC of the individual models. The TPC of the SPEAR model is the most similar to the TPC of the MMM. Therefore, from a temporal perspective, the MMM shows the largest Pearson correlation.

Figure 3 shows that the SPEAR model produces a high and significant SPC, a low NCRMSE and a low difference between normalized standard deviation (NDS) and the value of one for January, February, March, November, and December for both 0.5- and 1.5-month lead-times. For example, SPC, NCRMSE, and NDS of the SPEAR model for January at 0.5-month lead-time are 0.79, 0.62, and 0.97, and those of the MMM are 0.77, 0.64, and 0.63, respectively. The estimated SPC are statistically significant at the 95% level for all months and seasons. Figure 3 indicates that the MMM performs better than the individual model in April, May, and June. However, the SPEAR model produces a better NSD in January, February, March, and November (Figure 3). Also, Taylor diagrams (Figure 4) indicate that the SPEAR model performs better than the other models in JFM and OND seasons. From a spatial perspective, the combination of six NMME models does not show higher skill, particularly normalized standard deviation, than SPEAR model for OND and JFM season over Iran. These results are similar to the results of DJF precipitation multimodel ensemble forecasts over the central Southwest Asia (Ehsan *et al*., 2020b). However, the MMM produces a better skill for AMJ rainfall. In general, Taylor diagrams (Figures 3 and 4) indicate that the SPC and NCRMSE of the SPEAR model is close to MMM, but the NSD of the SPEAR model is closer to one compared to the MMM for months from November to March and two seasons (OND and JFM). When the average of NMME models is used, smoothing of the MMM forecasts is done and the variance of MMM is reduced and then the standard deviation ratio of MMM to the observations become less than one.

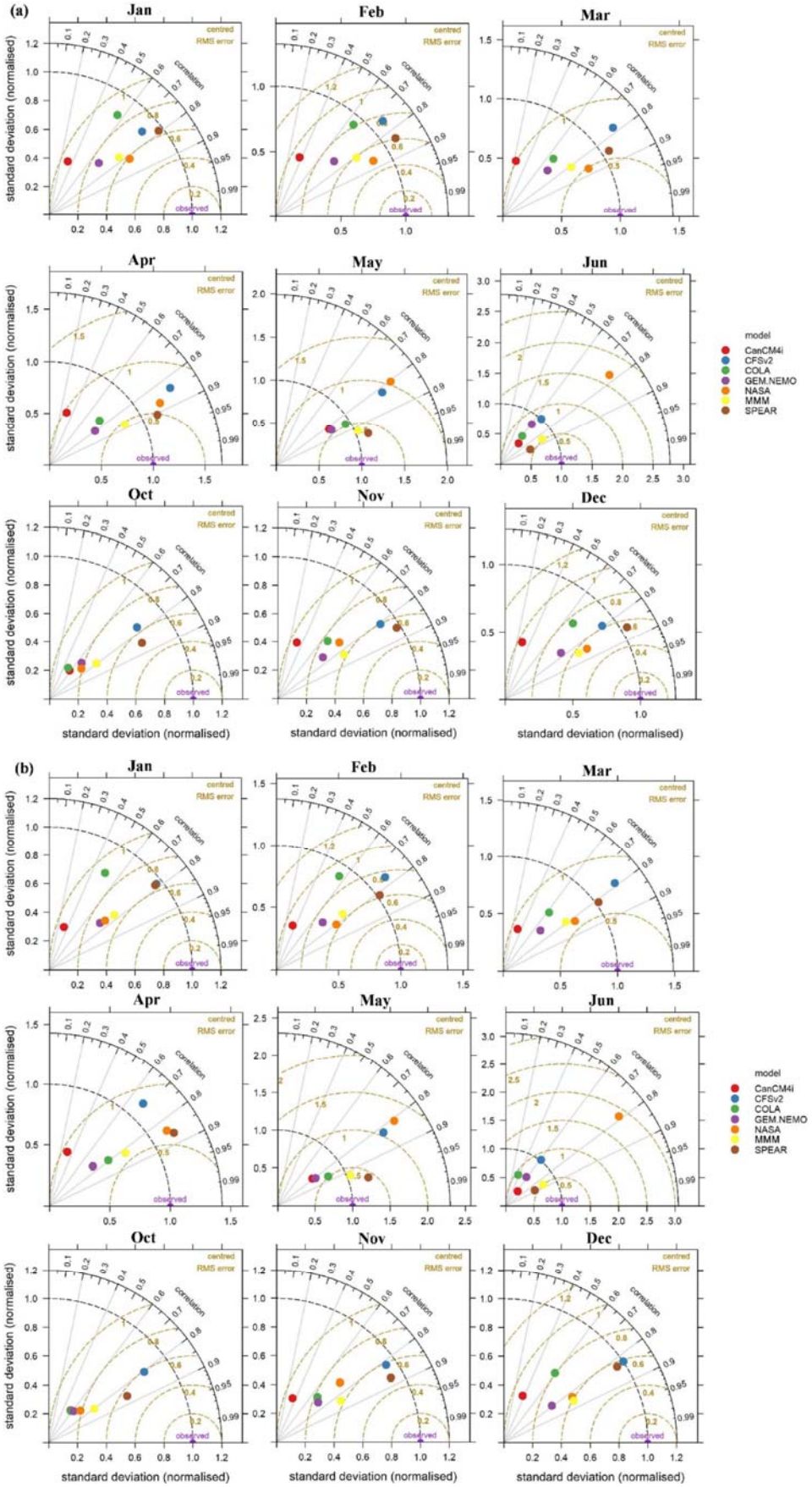
(a)
Jan
Feb
Mar
Apr
May
Jun
Oct
Nov
Dec
centred
RMS error
correlation
observed
standard deviation (normalised)
model
CanCM4i
CFSv2
COLA
GEM.NEMO
NASA
MMM
SPEAR
(b)

FIGURE 3. Taylor diagrams for climatological monthly precipitation over Iran between the observations and NMME models for 0.5-month (a) and 1.5-month (b) lead-times during 1991–2019

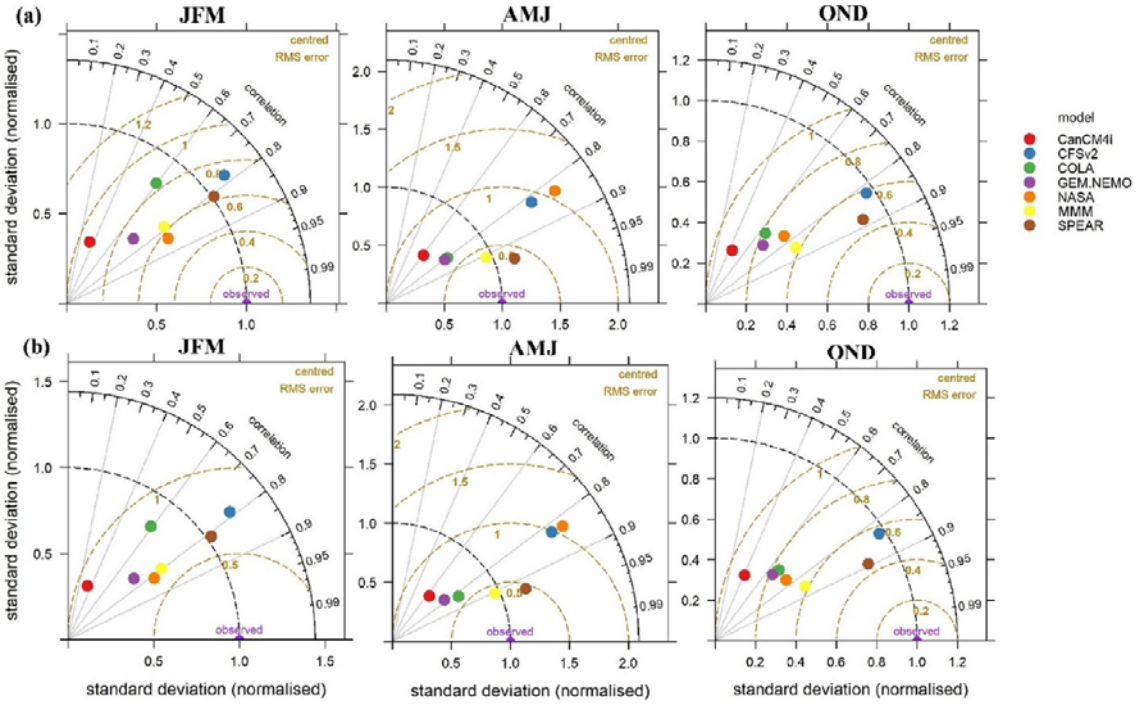

FIGURE 4. Taylor diagrams for climatological seasonal precipitation over Iran between the observations and NMME models for 1.5-month (a) and 2.5-month (b) lead-times during 1991–2019

In order to provide further information on different parts of the study area, the TPC and RPSS between time series of the observed and MMM forecasts are calculated for monthly and seasonal precipitation at each gridpoint. The TPC map for monthly forecasts at 0.5-month lead-time is plotted in Figure 5. Figure 5a indicates that there is a significant correlation between observed and forecast precipitation in most areas of the study area for all months at a 0.5-month lead-time. High correlation skill is obtained over the Zagros ranges for most months. The forecast skill is low over the northwest of Iran in January, February, and October (Figure 5a). In March, May, June, and November, the TPC is nonsignificant over the coast of the Caspian Sea (Figure 5a). The TPC is significant over the whole study area in December at 0.5-month lead-time. A significant Pearson correlation is obtained over some parts of the study area at 1.5-month lead-time for January, March, May, October, and November (Figure 5b). However, there is low skill over most parts of the domain for February, April, June and December for this lead-time.

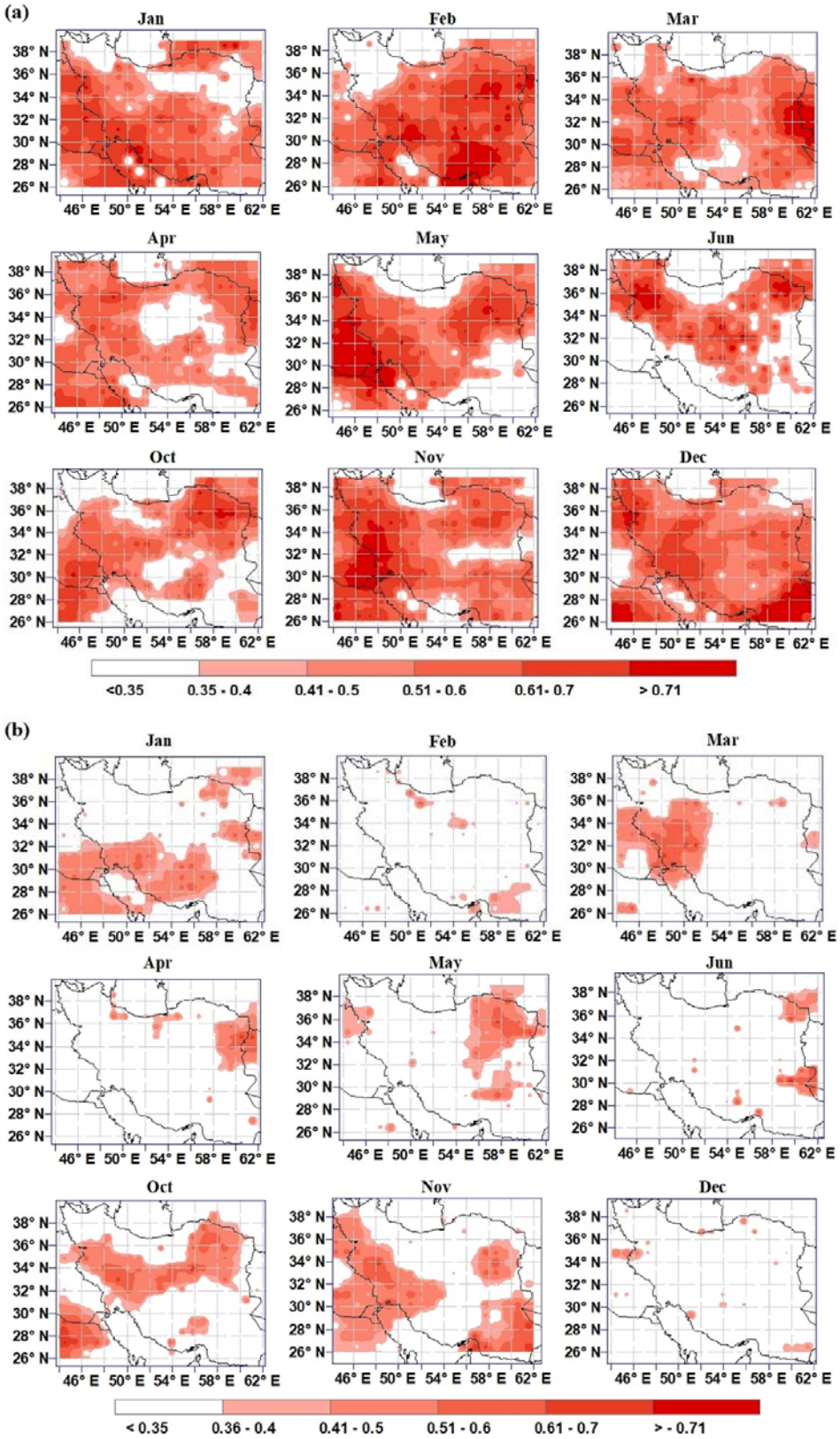

FIGURE 5. The temporal Pearson correlation map for multi-model mean (MMM) monthly forecasts for 0.5-month (a) and 1.5-month (b) lead-times. Areas where the correlation values do not have local statistical significance at the 95% level are masked out

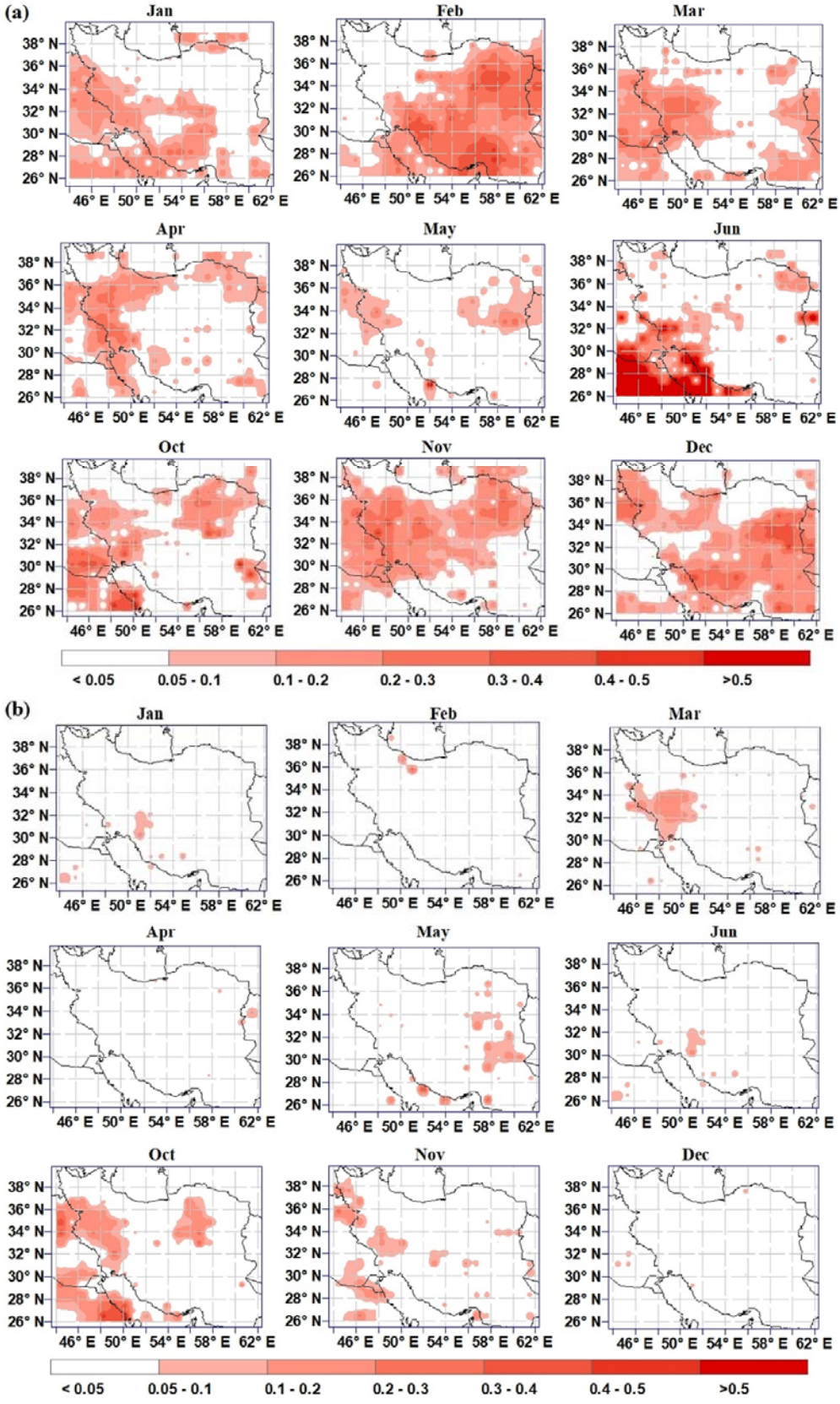

FIGURE 6. The spatial distribution of the RPSS for MMM monthly forecasts for 0.5-month (a) and 1.5-month (b) lead-times

Figure 6a shows the spatial skill of MMM forecasts for monthly forecasts for 0.5-month lead-time as measured by the RPSS. This figure shows that the RPSS is high at 0.5-month lead-time for February, November, and December in comparison with the other months. The lowest RPSS is over the northern parts of Iran. The RPSS indicates skilful forecast over southeast and east of Iran in February at 0.5-month lead-time. The spatial RPSS also indicates a low forecast skill over the larger part of the study area for most months at 1.5-month lead-time (Figure 6b). Overall, the results of the RPSS are consistent with that of the correlation maps. Figure 7 shows the spatial distribution of the mean difference between MMM forecasts and observed for monthly timescale at 0.5-month lead-time. This figure indicates that the MMM precipitation forecasts underestimate the precipitation totals of January, February, March, October, November, and December over the northern regions and Zagros mountains where the precipitation amount is high. The MMM precipitation forecasts are overestimated in April, May, and June over the larger part of the study area. Also, these forecasts are overestimated over the central regions during all months.

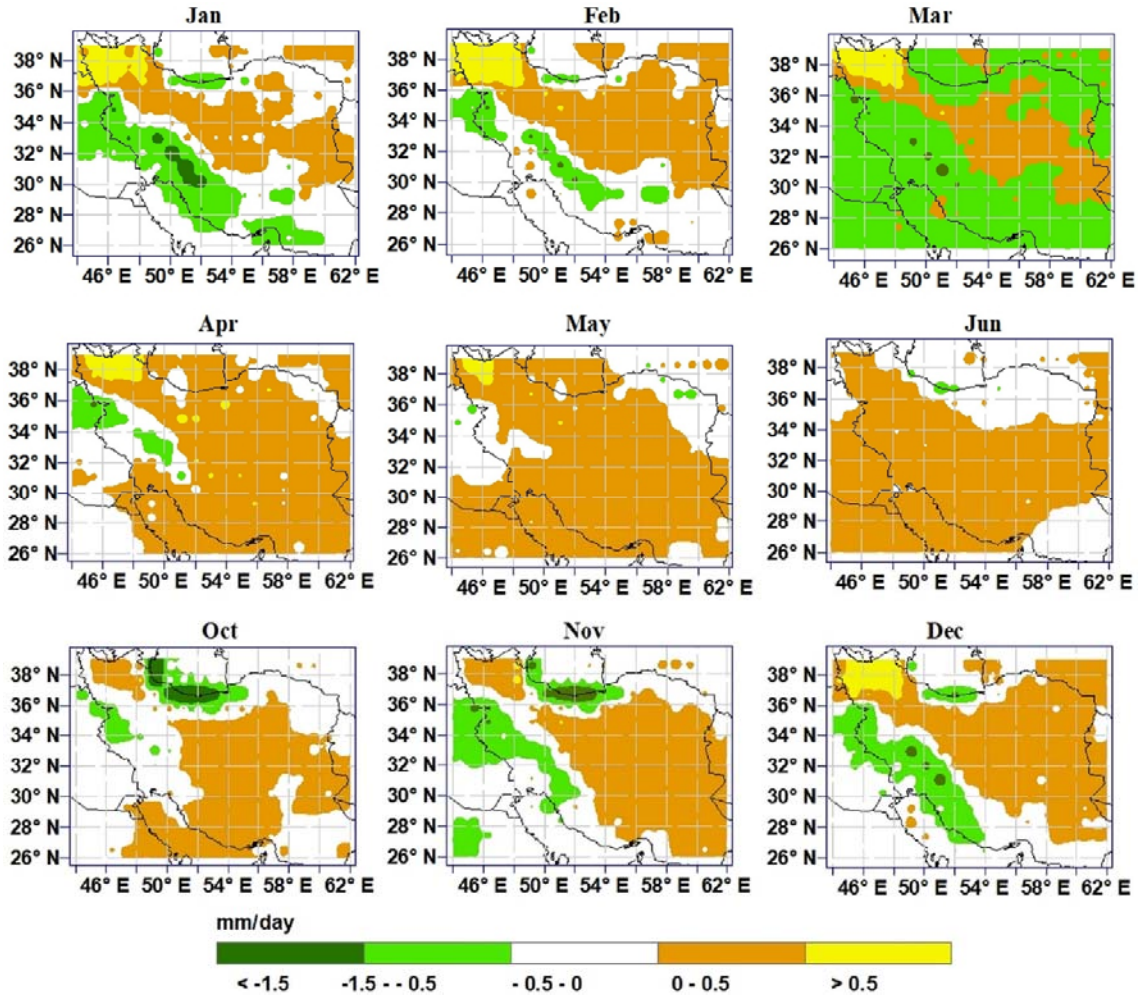

FIGURE 7. The spatial distribution of the mean difference (MMM Forecasts − Observed) for monthly timescale for 0.5-month lead-time

Figure 8 indicates that TPC is significant over the western, southern, and northeastern areas of Iran for JFM precipitation at a 1.5-month lead-time, and TPC is significant over some areas

in the western and northeastern at a 2.5-month lead-time. Figures 8 and 9 indicate that the skill for OND precipitation forecasts is higher than JFM and AMJ precipitation. These figures also indicate that the skill for AMJ forecasts is lower than JFM and OND precipitation. Both Taylor diagram and RPSS indicate that the seasonal precipitation forecasts have a higher skill in OND. This result is consistent with previous works (Shirvani and Landman, 2016; Najafi *et al*., 2021). The spatial pattern of the MD for seasonal timescales indicates that the MMM precipitation forecasts for both 1.5- and 2.5-month lead-times are underestimated over the northern regions and Zagros mountains for JFM and OND (Figure 10). The spatial pattern of the MD for AMJ precipitation is different from the other seasons such that the MMM precipitation forecasts for both 1.5- and 2.5-month lead-times are similar and overestimated for AMJ over the larger part of the study area.

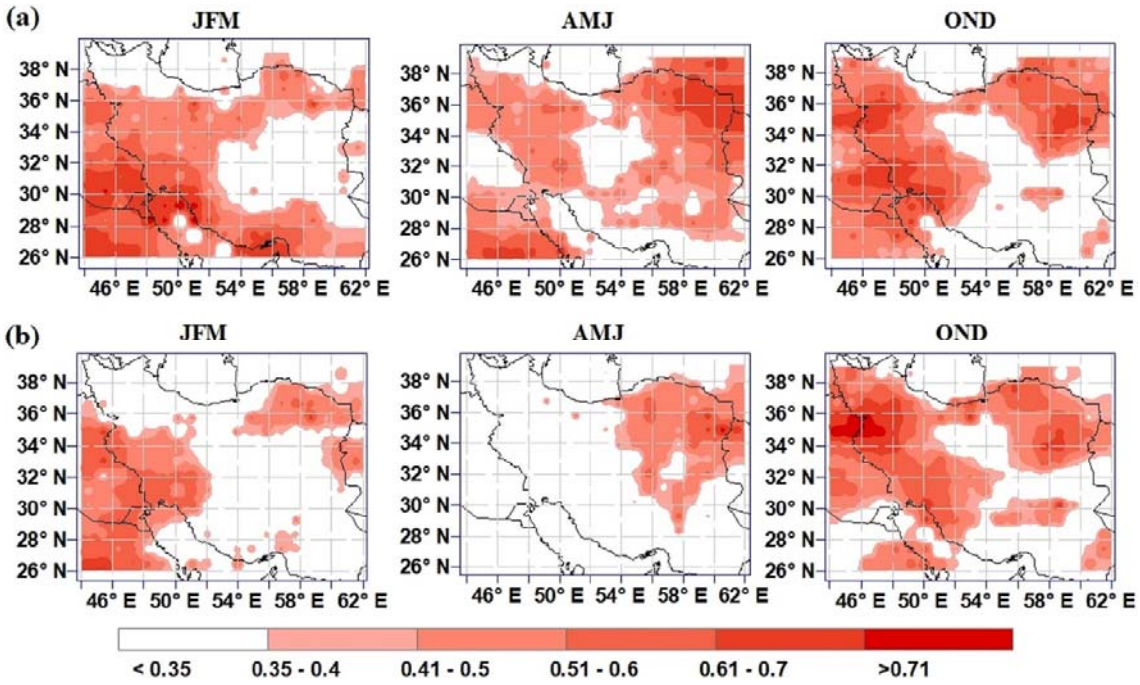

FIGURE 8. The temporal Pearson correlation map for seasonal MMM forecasts for 1.5-month (a) and 2.5-month (b) lead-times. Areas where the correlation values do not have local statistical significance at the 95% level are masked out

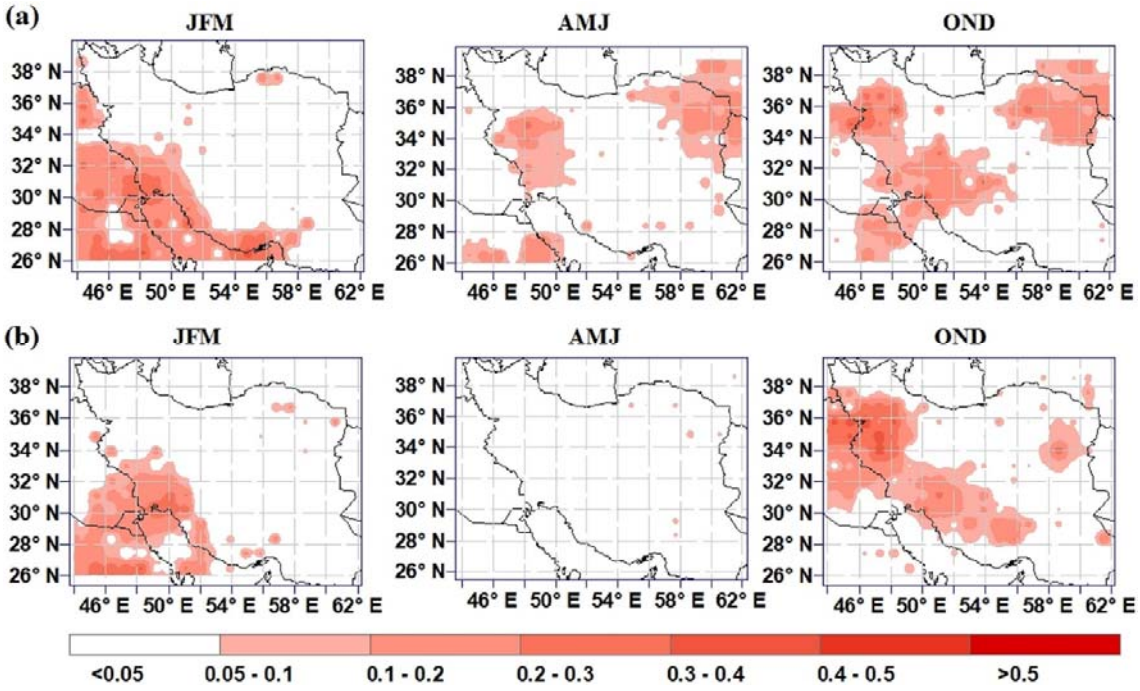

FIGURE 9. The spatial distribution of the RPSS for MMM seasonal forecasts for 1.5-month (a) and 2.5-month (b) lead-times

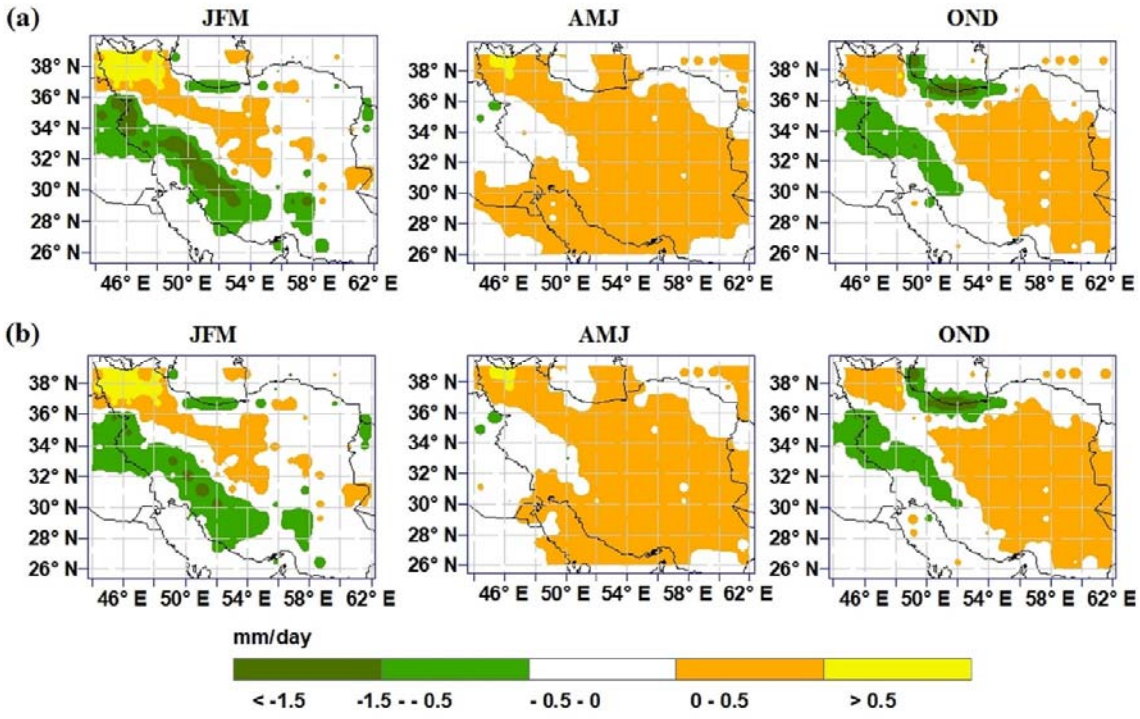

FIGURE 10. The spatial distribution of the mean difference (MMM Forecasts − Observed) for seasonal timescale for 1.5-month (a) and 2.5-month (b) lead-times

In order to determine whether or not the MMM's forecast skill is dependent on ENSO connections with precipitation, the TPC between simultaneous Niño3.4 SST and seasonal precipitation for observations and MMM forecasts for 1.5-month lead-time over the study area is computed and shown

in Figure 11. This figure indicates that there is a positive and significant Pearson correlation between Niño3.4 SST and precipitation for both observation and forecast in OND over the larger part of Iran. For this season there is a positive significant Pearson correlation between observed Niño3.4 SST and precipitation over the western, southwestern, northwestern, and northeastern areas of Iran. However, the linear association between observed Niño3.4 SST and precipitation is not significant over most areas for the JFM and AMJ seasons (Figure 11a). That is significant over most areas for MMM forecasts for AMJ season (Figure 11b). Therefore, most areas of Iran are not linearly correlated with Niño3.4 SST during JFM, which is one of the main rainy seasons. Moreover, Niño3.4 SST–precipitation teleconnection during AMJ is not correctly generated by MMM. These results have indicated that although the response of JFM precipitation to ENSO events is weak, the skill of MMM precipitation forecasts is acceptable for this season at 1.5-month lead-time over most parts of Iran, as well as at 2.5-month lead-time for some parts of Iran (Figure 8). These results suggest that the skill of MMM is not restricted to ENSO phases as has been found for other regions, for example, southern Africa (Landman and Beraki, 2012).

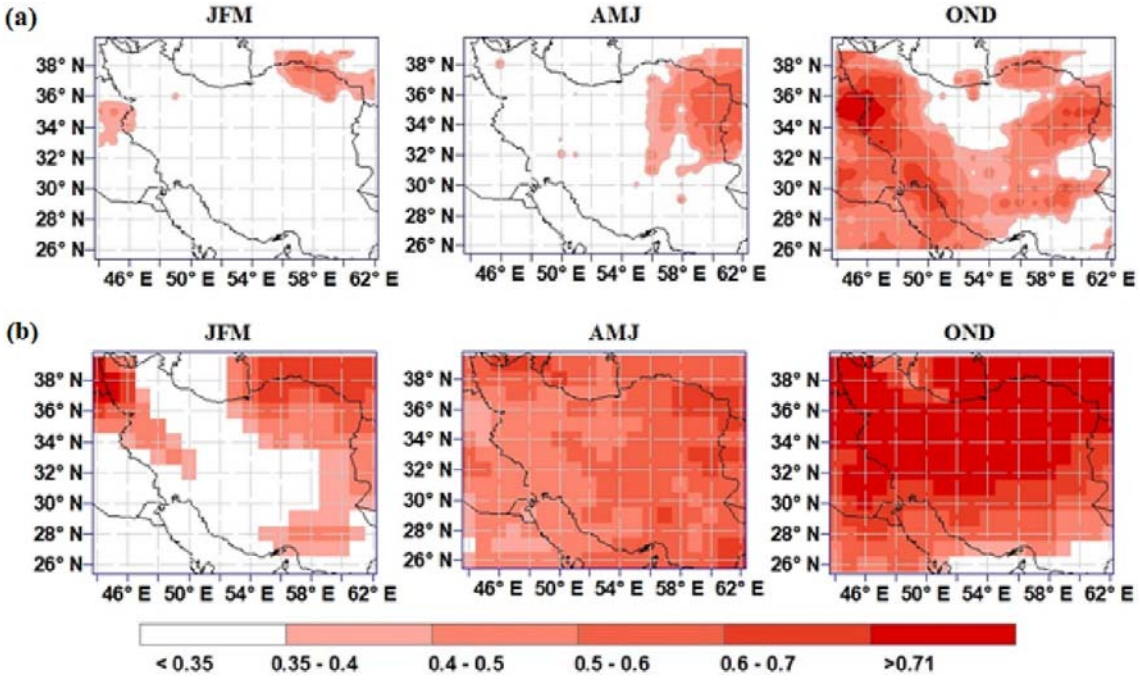

FIGURE 11. Niño3.4SST–precipitation Pearson correlation maps for observation (a) and NMM forecast at lead 1.5-month (b) for JFM, AMJ, and OND seasons. Colour parts indicate significant areas at the 95% level

The regional averaged of the observed seasonal precipitation and MMM precipitation forecast (at 1.5 and 2.5-month lead-times) are standardized and then its time series plotted, and scatter plot and heatmap of the correlations are presented in Figures 12-14, respectively. These figures also contain seasonal Niño3.4 SST anomalies. Figure 12a (Figure 12b) shows that both below and above-average observed JFM (AMJ) precipitation are occurred during El Niño years, indicating that the JFM (AMJ) precipitation variability is not closely aligned with ENSO events. However, the below (above)-average observed and forecasted precipitation for OND season are occurred during La Niña (El Niño) years (Figure 12c). There is a moderate (weak) uphill pattern between the observations with forecast at 1.5 and 2.5 month lead-times (Niño3.4 SST) for JFM season (Figure 13a,b). For the OND season, Figure 13c indicates a moderate uphill pattern between the observations with forecast at 1.5 and 2.5 month lead-

times, and Niño3.4 SST. The heatmap of the Spearman and Pearson correlations for regional observations, MMM forecast at 0.5 and 1.5 month lead-times, and Niño3.4 SST anomalies plotted in Figure 14. Neither the Spearman nor the Pearson correlations are statistically significant between observed precipitation and between seasons, indicating that the seasonal precipitation pattern is different between seasons. The Spearman and Pearson correlation coefficients between the observed JFM precipitation with forecasts at 1.5 and 2.5 month lead-times are statistically significant at the 0.05 significance level, while these correlations between the observed JFM precipitation and Niño3.4 SST anomalies are not significant (Figure 14). These correlations between the observed AMJ precipitation with forecasts at 1.5 month lead-time are statistically significant at the 0.05 significance level. However, the Spearman correlation between the observed AMJ precipitation with 2.5 month lead-time forecasts and Niño3.4 SST anomalies are not statistically significant at 0.05 significance level. Figure 14 shows that both the Spearman and Pearson correlations between OND observed precipitation with forecast at 0.5 and 1.5 month lead-times, and Niño3.4 SST anomalies are strong and statistically significant at the 0.05 significance level. These results which are consistent with previous studies (Nazemosadat and Cordery, 2000; Nazemosadat and Ghasemi, 2004) indicate that the connection between the ENSO signal and OND precipitation in Iran is strong, but is weak for the JFM season. Also, the regional precipitation over Southwest Asia is not strongly influenced by ENSO in the individual months of January, February, and March (Hoell *et al*., 2015).

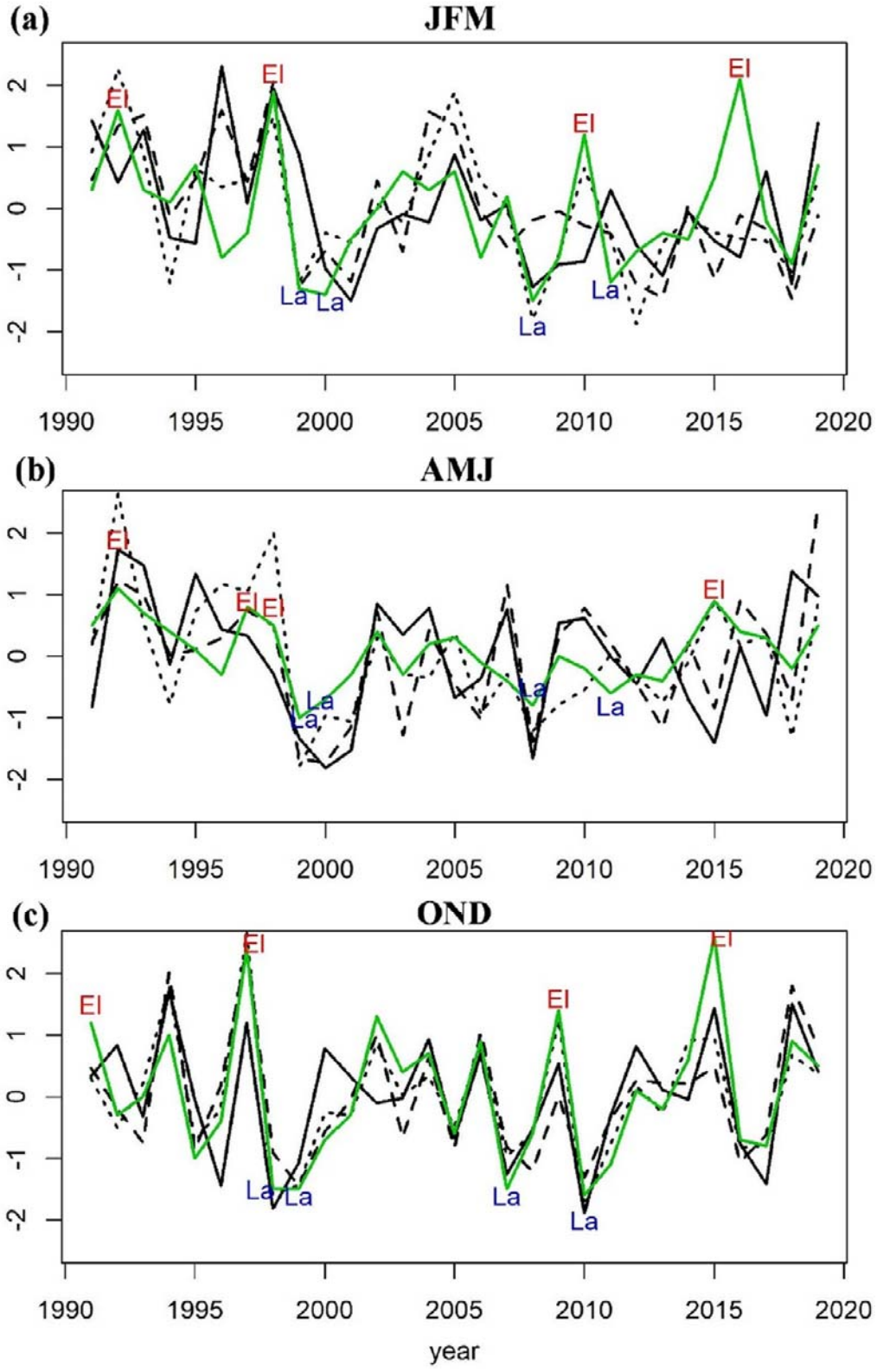

FIGURE 12. Time series plot of the standardized regional (area-averaged) observations (black line), MMM precipitation forecast at leads 1.5 (black dot) and 2.5 months (black dash), and Niño3.4 SST anomalies (green line) for JFM (a), AMJ (b), and OND (c) seasons. The "EL" refers to the strongest El Niño and "La" refers to the strongest La Niña during the study period

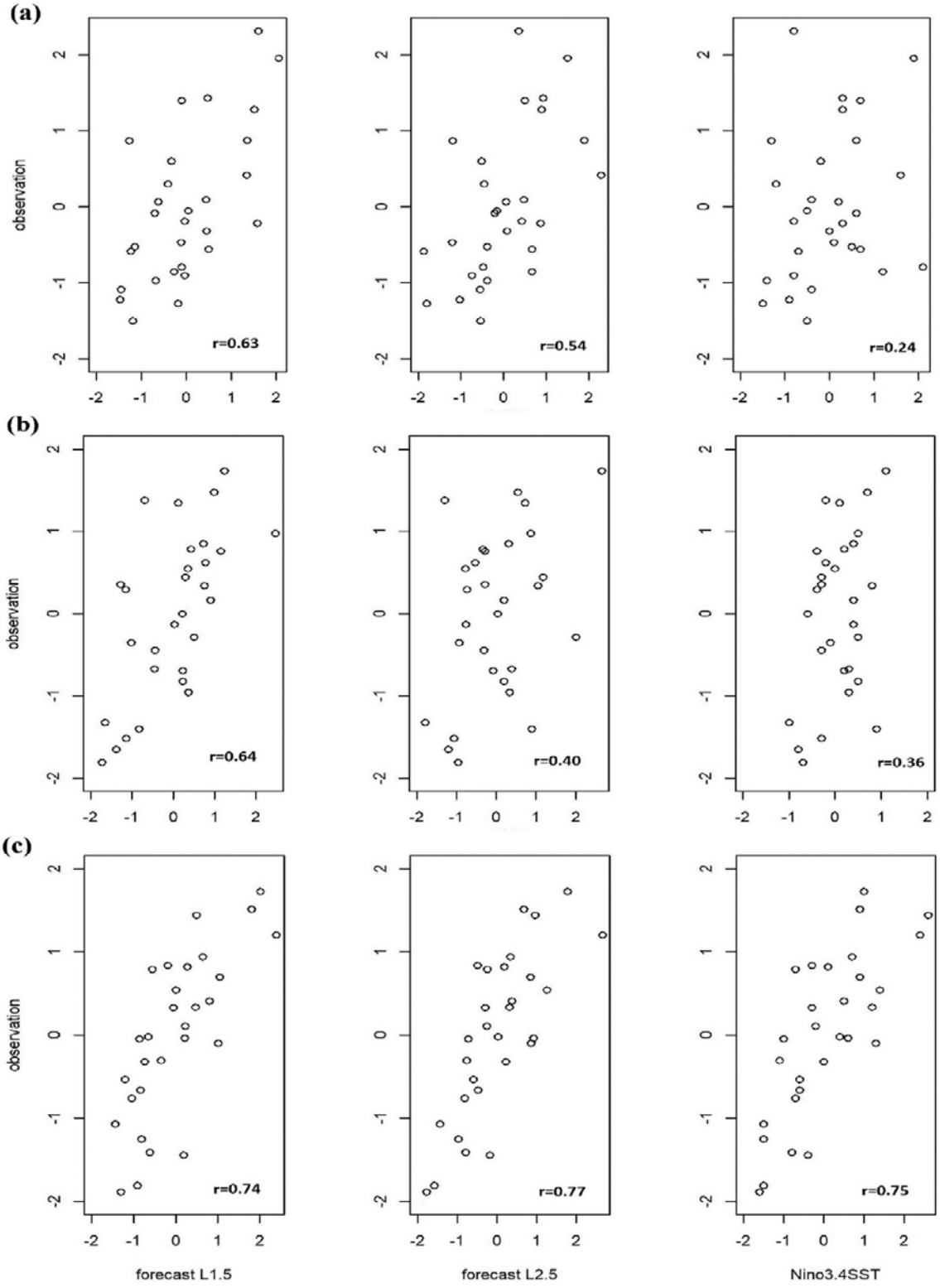

FIGURE 13. The scatter plot between the standardized regional (area-averaged) observations and MMM precipitation forecast at lead-time 1.5 (left), forecast at lead 2.5 months (middle), and Niño3.4 SST (right) for JFM (a), AMJ (b), and OND (c) seasons. The corresponding Pearson correlation (*r*) coefficients present in the scatter plot

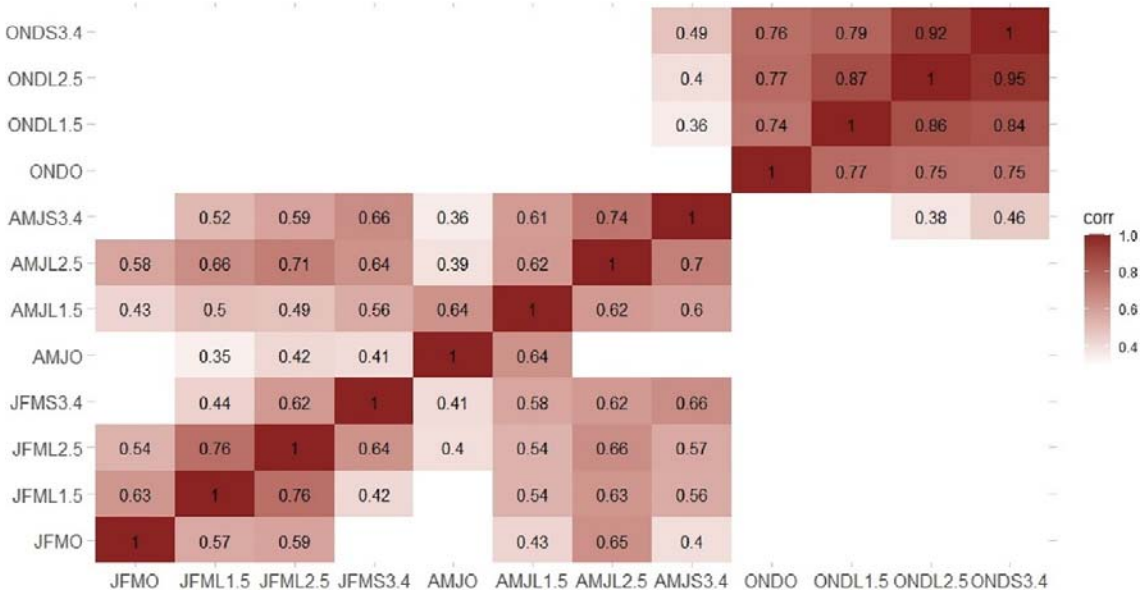

FIGURE 14. Heatmap of the Spearman (below the main diagonal) and Pearson (above the main diagonal) correlations for regional observations, MMM forecast at leads 1.5 and 2.5 months, and Niño3.4 SST anomalies. The O, L1.5, L2.5, and S3.4 after the name of seasons refer to observations, MMM forecast at leads 1.5 and 2.5 months, and Niño3.4 SST, respectively. White boxes indicate the correlation values do not have statistical significance at the 95% level

However, for the OND, JFM, and AMJ precipitation seasons, the area-averaged MMM precipitation forecast at 1.5 and 2.5-month lead-times are statistically significant. Therefore, although the connection between the ENSO signal with JFM and AMJ precipitation in Iran is weak, significant skill levels can be obtained from MMM forecasts at 1.5 and 2.5-month lead-times. The forecast skill level is closely related to ENSO and so decreases notably in JFM and AMJ along with the ENSO signal. Figure 15 shows the ENSO teleconnection for JFM, AMJ, and OND precipitation in terms of 200 hPa heights for observed and MMM forecast data at 1.5-month lead-time, respectively. The spatial pattern of geopotential height at 200 hPa–Niño3.4 SST anomalies correlation for observed data is similar to MMM forecast data in OND season, while there is some different in JFM and AMJ seasons. While the seasonal change in ENSO influence on precipitation has been shown previously (Hoell *et al*., 2017), the circulation analysis here shows that the change in precipitation influence is associated with the degree to which the circulation anomalies in southern Asia are able to penetrate to the west. While a full analysis of the dynamics underlying the change in ENSO teleconnection is beyond the scope of this analysis, both the structure of the Asian jet and the location and intensity of the tropical diabatic heating are important to the propagation of the ENSO signal into the region (Lau et al., 2012; Barlow *et al*., 2016; 2021) and we hypothesize that the seasonal evolution in both factors are playing an important role in the precipitation changes. Figures 17, 16 and 18 show the Pearson correlation between simultaneous global SST anomalies and regional precipitation for observed, MMM and SPEAR forecast data at 1.5-month lead-time, respectively. The spatial pattern of observed global SST anomalies-regional precipitation correlation is similar to MMM and SPEAR models in OND season, while there is some different in JFM and AMJ seasons. Figures 16c and 18c suggested that teleconnection between global SST anomalies and regional precipitation is generated by models in OND season. Therefore, these models are able to capture the physical mechanism generating OND precipitation over

Iran (based only on the SST). The similar result has been discussed in the recent study (Acharya *et al*., 2021) that concluded that the physical mechanism generating rainfall over Ethiopia during JJAS season is captured by the GCM as well. The influence of SSTs from different basins, such as the Black and Mediterranean Seas, is possible sources of the non-ENSO predictability for JFM precipitation over Iran (Figure 16a).

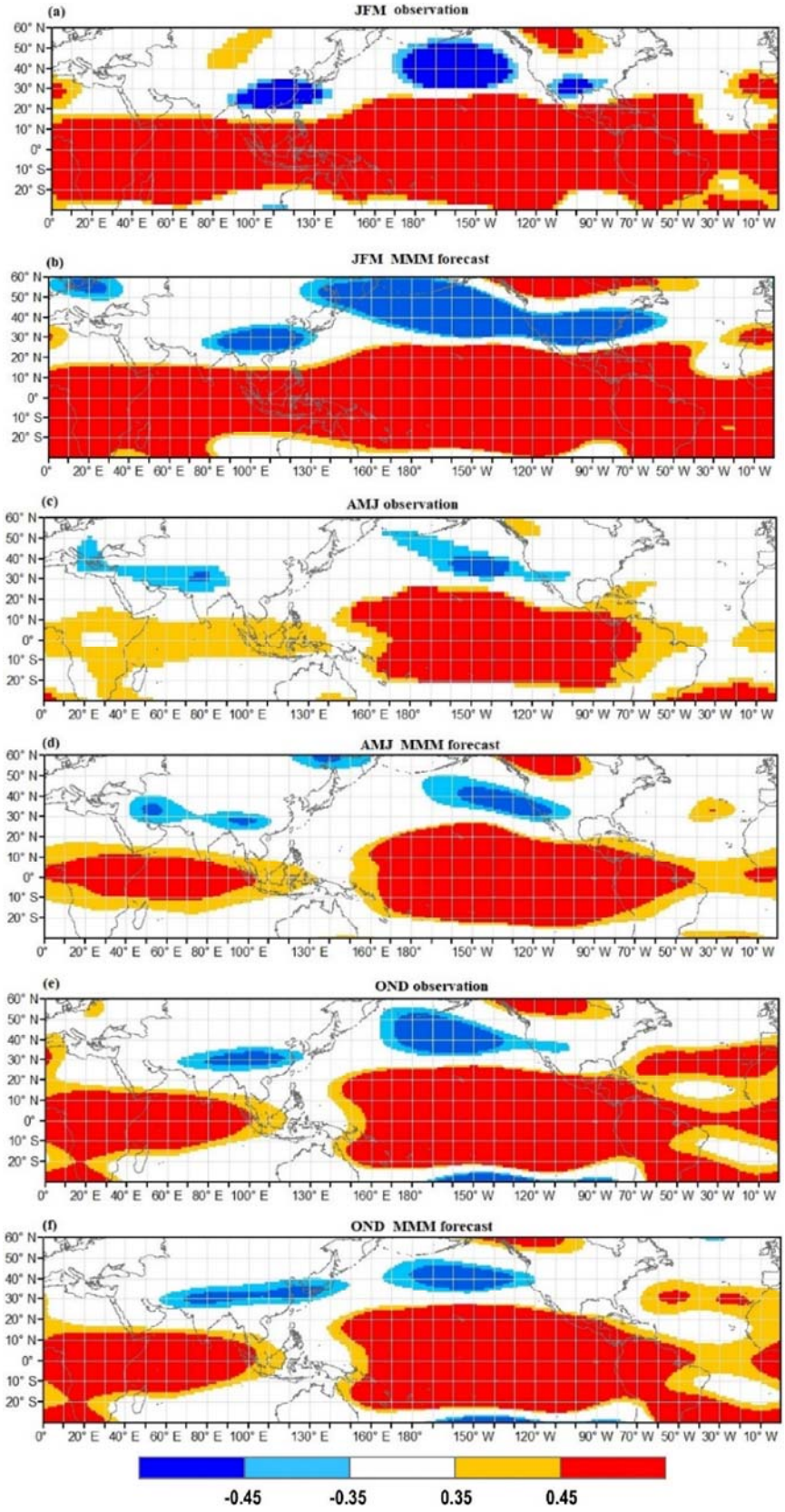

FIGURE 15. Geopotential height at 200 hPa–Niño3.4 SST anomalies correlation maps for observed (a, c, e) and MMM forecast at lead 1.5-month (b, d, f) data. Colour parts indicate significant areas at the 95% level

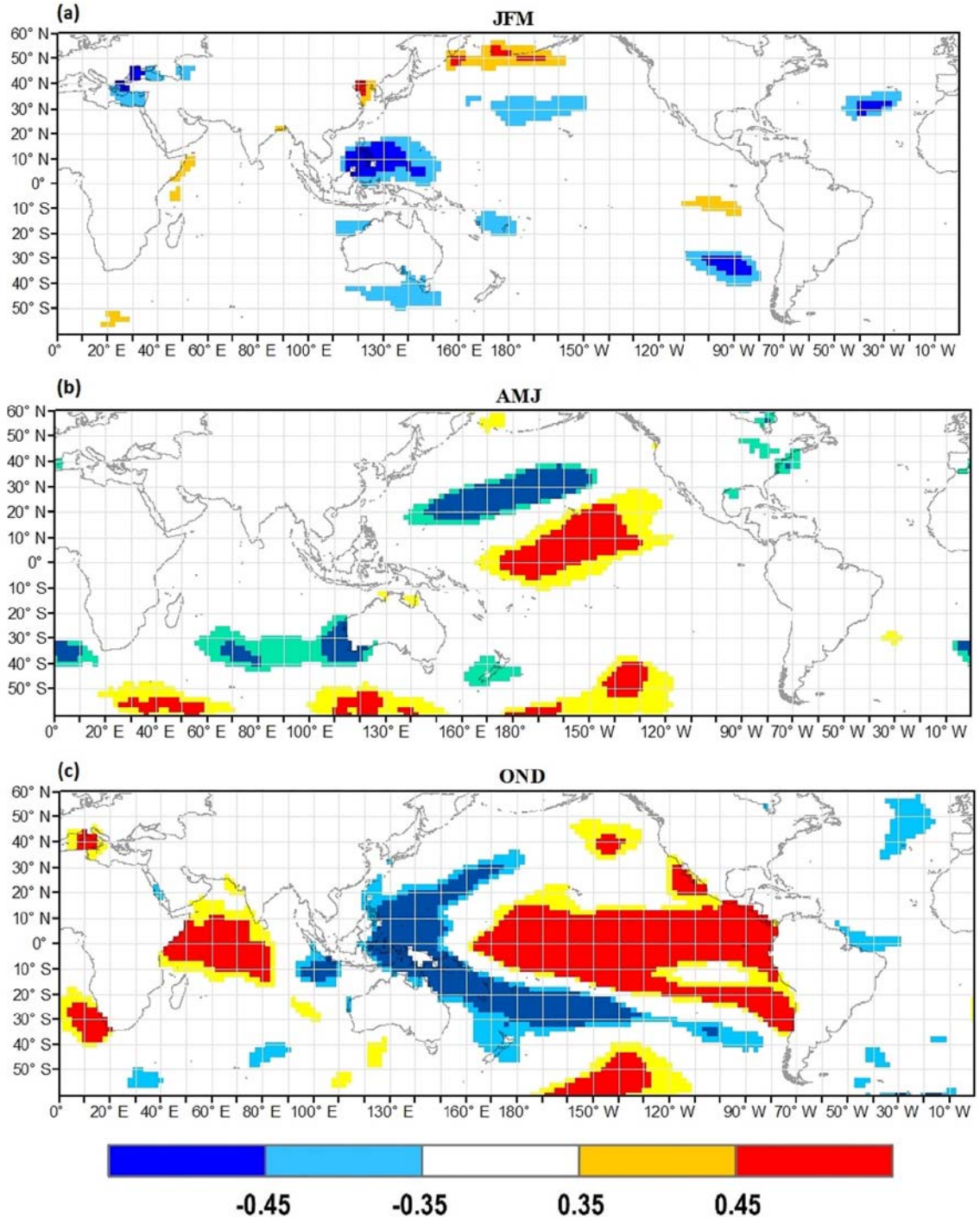

FIGURE 16. Observed global SST anomalies–regional precipitation correlation maps for JFM (a), AMJ (b), and OND (c) seasons. Colour parts indicate significant areas at the 95% level

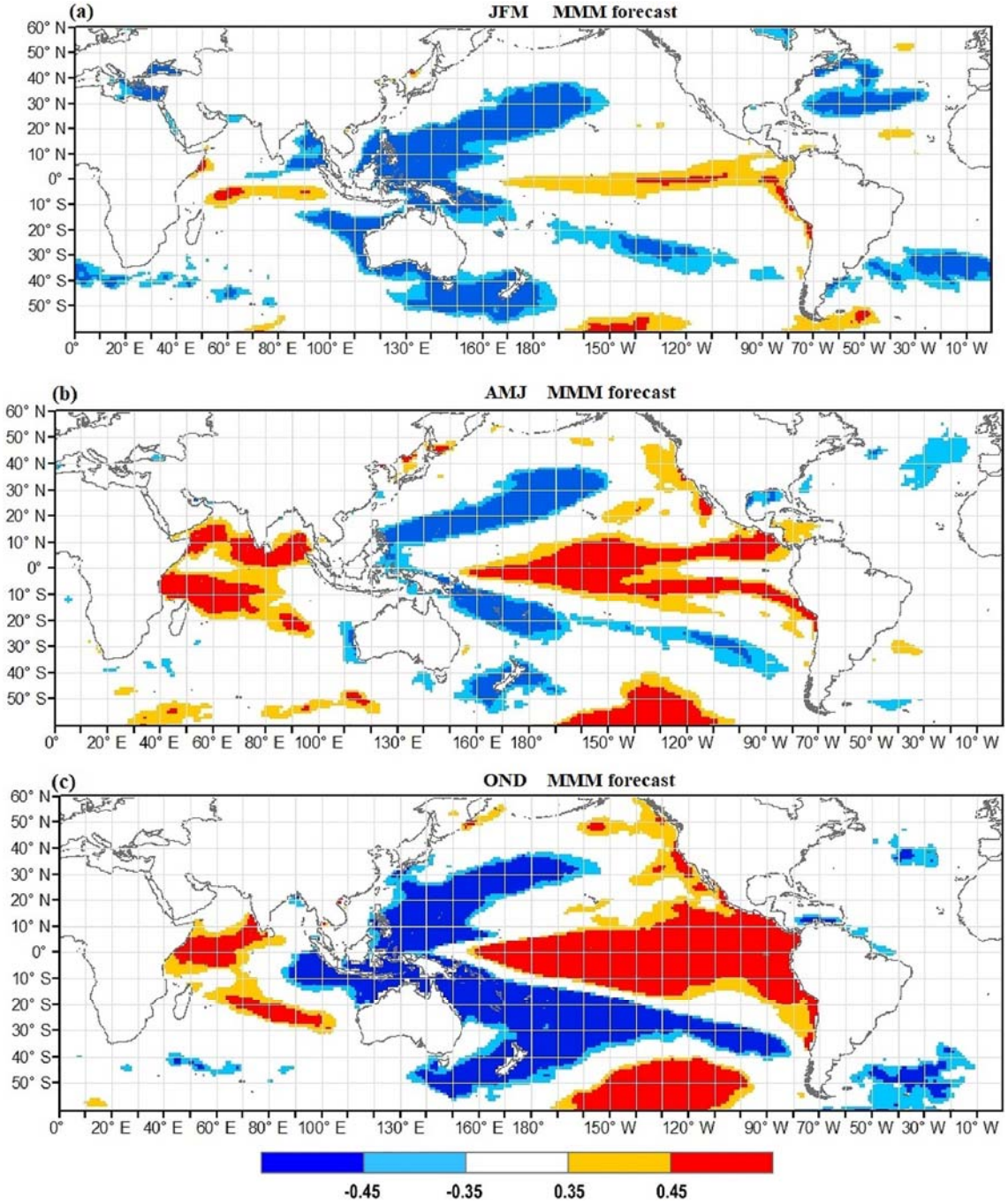

FIGURE 17. Global SST anomalies–regional precipitation correlation map for MMM forecast for 1.5-month lead-time for JFM (a), AMJ (b), and OND (c) seasons. Colour parts indicate significant areas at the 95% level

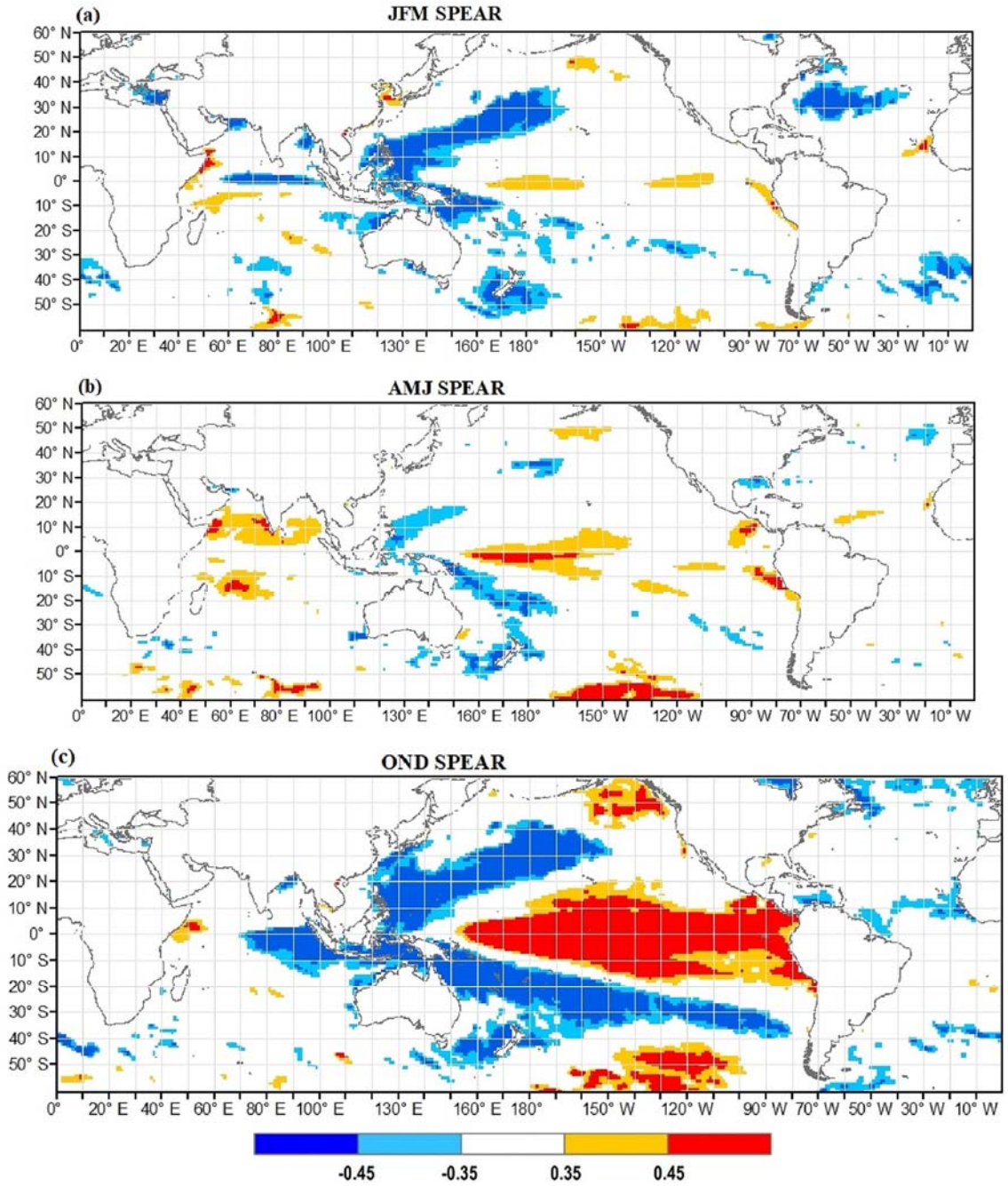

FIGURE 18. Global SST anomalies–regional precipitation correlation map for forecasts from SPEAR model for 1.5-month lead-time for JFM (a), AMJ (b), and OND (c) seasons. Colour parts indicate significant areas at the 95% level

4 CONCLUSIONS

A deterministic and probabilistic investigation of the NMME monthly and seasonal precipitation forecasts is conducted to assess the skill of the NMME precipitation forecasts over Iran during its wet season. The NMME precipitation data from six models (CFSv2, NASA, COLA, SPEAR, CanCM4i, and GEM-NEMO) as well as a multi-model mean forecasts are evaluated for monthly (October–June) at leads of 0.5, 1.5, 2.5, and 3.5 months and seasonal

(OND, JFM, and AMJ) at leads of 1.5, 2.5, and 3.5 months based on the period 1991–2019 (results summarized in Figures 2-4). Forecasts at the 0.5-month (1.5-month) lead-times are effectively forecasts for the month (season) in which the initialization of the NMME model(s) occurs. Although the 0.5 (1.5)-month lead-times is for the same month (season) in which the global model was initialized, the initialization happens early in the month (season) and so such forecasts may still be useful to forecast users. Using Taylor diagrams, for most months (October–March) and lead-times, the highest skills are provided by the SPEAR model. For most models, the month with highest forecast skill is November. For the SPEAR and the MMM, the forecast skill is statistically significant for all lead-times.

The MMM does not show many advantages over the SPEAR in forecasting OND and JFM precipitation. However, the MMM produces the highest overall deterministic and probabilistic forecast skill in April, May, and June months and the AMJ season. A low RPSS and nonsignificant temporal Pearson correlation are observed over northern Iran for most months. The MMM precipitation forecasts are underestimated over the northern regions and Zagros mountains where the amount of precipitation is high for JFM and OND for both 1.5- and 2.5-month lead-times. However, the MMM precipitation forecasts are overestimated over most regions in the AMJ season. Statistically significant temporal and spatial Pearson correlation is found for the SPEAR and MMM forecasts for all seasons and all leads (except 3.5-month lead for JFM SPEAR) and the seasonal precipitation forecasts have the highest skill in the OND season. On the other hand, the temporal Pearson correlation between the area-averaged time series of the observed precipitation and Niño3.4 SST is only significant for the OND and not for the JFM and AMJ seasons. Therefore, although the connection between ENSO events with JFM and AMJ precipitation over Iran is marginal, there is, nonetheless, significant skill in those seasons in the MMM forecasts. That is, the MMM forecasts have skill beyond ENSO. This result is consistent with the results by Landman and Beraki (2012) for South Africa and Shirvani and Landman (2016) for Iran.

This analysis raises some interesting dynamical questions in terms of what causes the changes in regional ENSO influence between the first and second halves of the cold season, and what is the source of predictability in the second half when the ENSO influence is negligible. Barlow *et al*. (2021) have suggested that the structure of the Asian jet can play an important role in modulating the ENSO influence, so analysis of the seasonal changes in the jet structure may prove useful. Possible sources of the non-ENSO predictability are not yet clear, but the presence of model forecast skill suggest that some targeted modelling experiments to separate out the influence of SSTs from different basins, as well as the role of regional factors such as soil moisture may prove fruitful.

ACKNOWLEDGEMENTS

The authors greatly appreciate the anonymous reviewers for their suggestions which led to improvements in the manuscript. The authors acknowledge the NOAA Climate Prediction Center and IRI for supporting the data. The NMME data are obtained from http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/. The GPCC version 2020 data are obtained from https://iridl.ldeo.columbia.edu/SOURCES/.WCRP/.GCOS/.GPCC/. The Niño3.4 data are obtained from https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php. The

reanalysis data are obtained from http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis.html.

REFERENCES

Acharya, N., Ehsan, M.A., Admasu, A., Teshome, A. and Hall, K.J.C. (2021) On the next generation (NextGen) seasonal prediction system to enhance climate services over Ethiopia. *Climate Services*, 24, 100272. https://doi.org/10.1016/j.cliser.2021.100272.

Agrawala, S., Barlow, M., Cullen, H. and Lyon, B. (2001) The drought and humanitarian crisis in central and Southwest Asia: a climate perspective. Palisades, NY: The International Research Institute for Climate Prediction. IRI special report number: 01-11. https://doi.org/10.7916/D8NZ8FHQ.

Alizadeh-Choobari, O., Adibi, P. and Irannejad, P. (2018) Impact of the El Niño–Southern Oscillation on the climate of Iran using ERA-Interim data. *Climate Dynamics*, 51(7–8), 2897–2911. https://doi.org/10.1007/s00382-017-4055-5.

Balmaseda, M. and Anderson, D. (2009) Impact of initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters*, 36, L01701. https://doi.org/10.1029/2008GL035561.

Barlow, M., Cullen, H. and Lyon, B. (2002) Drought in central and Southwest Asia: La Niña, the warm pool, and Indian Ocean precipitation. *Journal of Climate*, 15, 697–700. https://doi.org/10.1175/1520-0442(2002)015<0697:DICASA>2.0.CO;2.

Barlow, M., Hoell, A. and Agel, L. (2021) An evaluation of CMIP6 historical simulations of the cold season teleconnection between tropical Indo-Pacific sea surface temperatures and precipitation in Southwest Asia, the coastal Middle East, and northern Pakistan and India. *Journal of Climate*, 34, 6905–6926. https://doi.org/10.1175/JCLI-D-19-1026.1.

Barlow, M., Zaitchik, B., Paz, S., Black, E., Evans, J. and Hoell, A. (2016) A review of drought in the Middle East and Southwest Asia. *Journal of Climate*, 29, 8547–8574. https://doi.org/10.1175/JCLI-D-13-00692.1.

Barnston, A.G., Li, S., Mason, S.J., Dewitt, D.G., Goddard, L. and Gong, X. (2010) Verification of the first 11 years of IRI's seasonal climate forecasts. *Journal of Applied Meteorology and Climatology*, 49, 493–520. https://doi.org/10.1175/2009JAMC2325.1.

Barnston, A.G. and Mason, S.J. (2011) Evaluation of IRI's seasonal climate forecasts for the extreme 15% tails. *Weather and Forecasting*, 26, 545–554. https://doi.org/10.1175/WAF-D-10-05009.1.

Barnston, A.G., Mason, S.J., Goddard, L., DeWitt, D.G. and Zebiak, S.E. (2003) Multimodel ensembling in seasonal climate forecasting at IRI. *Bulletin of the American Meteorological Society*, 84, 1783–1796. https://doi.org/10.1175/BAMS-84-12-1783.

Becker, E., den Dool, H.V. and Zhang, Q. (2014) Predictability and forecast skill in NMME. *Journal of Climate*, 27, 5891–5906. https://doi.org/10.1175/JCLI-D-13-00597.1.

Becker, E.J., Kirtman, B.P., L'Heureux, M., Muñoz, Á.G. and Pegion, K. (2022) A decade of the North American Multimodel Ensemble (NMME): research, application, and future directions. *Bulletin of the American Meteorological Society*, 103(3), E973–E995. https://doi.org/10.1175/BAMS-D-20-0327.1.

Cash, B.A., Manganello, J.V. and Kinter, J.L. (2019) Evaluation of NMME temperature and precipitation bias and forecast skill for South Asia. *Climate Dynamics*, 53, 7363–7380. https://doi.org/10.1007/s00382-017-3841-4.

Delworth, T.L., Cooke, W.F., Adcroft, A., Bushuk, M., Chen, J.H., Dunne, K.A., Ginoux, P., Gudgel, R., Hallberg, R.W., Harris, L., Harrison, M.J., Johnson, N., Kapnick, S.B., Lin, S.J., Lu, F., Malyshev, S., Milly, P.C., Murakami, H., Naik, V., Pascale, S., Paynter, D., Rosati, A., Schwarzkopf, M.D., Shevliakova, E., Underwood, S., Wittenberg, A.T., Xiang, B., Yang, X., Zeng, F., Zhang, H., Zhang, L. and Zhao, M. (2020) SPEAR: the next generation GFDL modeling system for seasonal to multidecadal prediction and projection. *Journal of Advances in Modeling Earth Systems*, 12(3), e2019MS001895. https://doi.org/10.1029/2019MS001895.

Doblas-Reyes, F.J., García-Serrano, J., Lienert, F., Biescas, A.P. and Rodrigues, L.R.L. (2013) Seasonal climate predictability and forecasting: status and prospects. *Wiley Interdisciplinary Reviews: Climate Change*, 4, 245–268. https://doi.org/10.1002/wcc.217.

Ehsan, M.A., Kucharski, F. and Almazroui, M. (2020b) Potential predictability of boreal winter precipitation over central Southwest Asia in the North American Multi-Model Ensemble. *Climate Dynamics*, 54, 473–490. https://doi.org/10.1007/s00382-019-05009-3.

Ehsan, M.A., Tippett, M.K., Almazroui, M., Ismail, M., Yousef, A., Kucharski, F., Omar, M., Hussein, M. and Alkhalaf, A.A. (2017) Skill and predictability in multimodel ensemble forecasts for Northern Hemisphere regions with dominant winter precipitation. *Climate Dynamics*, 48(9–10), 3309–3324.

Ehsan, M.A., Tippett, M.K., Kucharski, F., Almazroui, M. and Ismail, M. (2020a) Predicting peak summer monsoon precipitation over Pakistan in ECMWF SEAS5 and North American Multimodel Ensemble. *International Journal of Climatology*, 40(13), 5556–5573. https://doi.org/10.1002/joc.6535.

Ehsan, M.A., Tippett, M.K., Robertson, A.W., Almazroui, M., Ismail, M., Dinku, T., Acharya, N., Siebert, A., Seid Ahmed, J. and Teshome, A. (2021) Seasonal predictability of Ethiopian Kiremt rainfall and forecast skill of ECMWF's SEAS5 model. *Climate Dynamics*, 57, 3075–3091. https://doi.org/10.1007/s00382-021-05855-0.

Epstein, E.S. (1969) A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology*, 8, 985–987. https://doi.org/10.1175/1520-0450(1969)008<0985:ASSFPF>2.0.CO;2.

Fallah, A., Rakhshandehroo, G.R., Berg, P. and Orth, R. (2020) Evaluation of precipitation datasets against local observations in southwestern Iran. *International Journal of Climatology*, 40(9), 4102–4116. https://doi.org/10.1002/joc.6445.

Goddard, L. and Dilley, M. (2005) El Niño: catastrophe or opportunity. *Journal of Climate*, 18(2), 651–665. https://doi.org/10.1175/JCLI-3277.1.

Hagedorn, R., Hamill, T.M. and Whitaker, J.S. (2008) Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: precipitation. *Monthly Weather Review*, 136, 2608–2619. https://doi.org/10.1175/2007MWR2410.1.

Hoell, A., Barlow, M., Cannon, F. and Xu, T. (2017) Oceanic origins of historical Southwest Asia precipitation during the boreal cold season. *Journal of Climate*, 30, 2885–2903. https://doi.org/10.1175/JCLI-D-16-0519.1.

Hoell, A., Barlow, M., Xu, T. and Zhang, T. (2018) Cold season Southwest Asia precipitation sensitivity to El Niño–Southern Oscillation events. *Journal of Climate*, 31, 4463–4482. https://doi.org/10.1175/JCLI-D-17-0456.1.

Hoell, A., Funk, C. and Barlow, M. (2015) The forcing of southwestern Asia teleconnections by low-frequency sea surface temperature variability during boreal winter. *Journal of Climate*, 28, 1511–1526. https://doi.org/10.1175/JCLI-D-14-00344.1.

Hosseini-Moghari, S.M., Araghinejad, S. and Ebrahimi, K. (2018) Spatio-temporal evaluation of global gridded precipitation datasets across Iran. *Hydrological Sciences Journal*, 63(11), 1669–1688. https://doi.org/10.1080/02626667.2018.1524986.

Huang, B., Thorne, P.W., Banzon, V.F., Boyer, T., Chepurin, G., Lawrimore, J.H., Menne, M.J., Smith, T.M., Vose, R.S. and Zhang, H.M. (2017) Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *Journal of Climate*, 30(20), 8179–8205. https://doi.org/10.1175/JCLI-D-16-0836.1.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R. and Joseph, D. (1996) 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437–470.

Khajehei, S., Ahmadalipour, A. and Moradkhani, H. (2018) An effective post-processing of the North American Multi-Model Ensemble (NMME) precipitation forecasts over the continental US. *Climate Dynamics*, 95, 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1.

Kirtman, B.P. (2003) The COLA anomaly coupled model: ensemble ENSO prediction. *Monthly Weather Review*, 131(10), 2324–2341. https://doi.org/10.1175/1520-0493(2003)131<2324:TCACME>2.0.CO;2.

Kirtman, B.P. and Min, D. (2009) Multimodel ensemble ENSO prediction with CCSM and CFS. *Monthly Weather Review*, 137(9), 2908–2930. https://doi.org/10.1175/2009MWR2672.1.

Kirtman, B.P., Min, D., Infanti, J.M., Kinter, J.L., III, Paolino, D.A., Zhang, Q., van den Dool, H., Saha, S., Mendez, M.P., Becker, E., Peng, P., Tripp, P., Huang, J., de Witt, D.G., Tippett, M.K., Barnston, A.G., Li, S., Rosati, A., Schubert, S.D., Rienecker, M., Suarez, M., Li, Z.E., Marshak, J., Lim, Y.K., Tribbia, J., Pegion, K., Merryfield, W.J., Denis, B. and Wood, E.F. (2014) The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society*, 95, 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1.

Landman, W.A. (2014) How the International Research Institute for Climate and Society has contributed towards seasonal climate forecast modelling and operations in South Africa. *Earth Perspectives*, 1, 22. https://doi.org/10.1186/2194-6434-1-22.

Landman, W.A., Barnston, A.G., Vogel, C. and Savy, J. (2019) Use of El Niño–Southern Oscillation related seasonal precipitation predictability in developing regions for potential societal benefit. *International Journal of Climatology*, 39, 5327–5337. https://doi.org/10.1002/joc.6157.

Landman, W.A. and Beraki, A. (2012) Multi-model forecast skill for mid-summer rainfall over southern Africa. *International Journal of Climatology*, 32, 303–314. https://doi.org/10.1002/joc.2273.

Landman, W.A., Dewitt, D., Lee, D.E., Beraki, A. and Lötter, D. (2012) Seasonal rainfall prediction skill over South Africa: one- versus two-tiered forecasting systems. *Weather and Forecasting*, 27, 489–501. https://doi.org/10.1175/WAF-D-11-00078.1.

Lau, W.K.M., Waliser, D.E. and Barlow, M. (2012) Africa and West Asia. In: Intraseasonal Variability in the Atmosphere-Ocean Climate System. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-13914-7_13.

Lin, H., Merryfield, W.J., Muncaster, R., Smith, G.C., Markovic, M., Dupont, F., Roy, F., Lemieux, J.F., Dirkson, A., Kharin, V.V., Lee, W.S., Charron, M. and Erfani, A. (2020) The Canadian Seasonal To Interannual Prediction System version 2 (CanSIPSv2). *Weather and Forecasting*, 35(4), 1317–1343. https://doi.org/10.1175/WAF-D-19-0259.1.

Ma, F., Yuan, X. and Ye, A. (2015) Seasonal drought predictability and forecast skill over China. *Journal of Geophysical Research: Atmospheres*, 120, 8264–8275. https://doi.org/10.1002/2015JD023185.

Mariotti, A. (2007) How ENSO impacts precipitation in southwest central Asia. *Geophysical Research Letters*, 34(16), L16706. https://doi.org/10.1029/2007GL030078.

Mason, S.J. and Goddard, L. (2001) Probabilistic precipitation anomalies associated with ENSO. *Bulletin of the American Meteorological Society*, 82, 619–638. https://doi.org/10.1175/1520-0477(2001)082<0619:PPAAWE>2.3.CO;2.

McCullagh, P. and Nelder, J.A. (1989) Generalized Linear Models, Vol. 37. London, UK: CRC Press.

Molod, A., Akella, S., Andrews, L., Arnold, N., Barahona, D., Borovikov, A., Cullather, R., Chang, Y., Hackert, E., Kovach, R. and Randal, K. (2020) GEOS S2S Version 3: The New NASA/GMAO High Resolution Seasonal Prediction System. In American Meteorological Society (AMS) Annual Meeting (No. GSFC-E-DAA-TN77310).

Müller, W.A., Appenzeller, C., Doblas-Reyes, F.J. and Liniger, M.A. (2005) A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. *Journal of Climate*, 18, 1513–1523. https://doi.org/10.1175/JCLI3361.1.

Najafi, H., Robertson, A.W., Massah Bavani, A.R., Irannejad, P., Wanders, N. and Wood, E.F. (2021) Improved multi-model ensemble forecasts of Iran's precipitation and temperature using a hybrid dynamical-statistical approach during fall and winter seasons. *International Journal of Climatology*, 41(12), 5698–5725. https://doi.org/10.1002/joc.7148.

Nazemosadat, M.J. and Cordery, I. (2000) On the relationships between ENSO and autumn rainfall in Iran. *International Journal of Climatology*, 20, 47–61.

Nazemosadat, M.J. and Ghasemi, A.R. (2004) Quantifying the ENSO-related shifts in the intensity and probability of drought and wet periods in Iran. *Journal of Climate*, 17, 4005–4018. https://doi.org/10.1175/1520-0442(2004)017<4005:QTESIT>2.0.CO;2.

Raziei, T., Bordi, I. and Pereira, L.S. (2011) An application of GPCC and NCEP/NCAR datasets for drought variability analysis in Iran. *Water Resources Management*, 25(4), 1075–1086. https://doi.org/10.1007/s11269-010-9657-1.

Sachindra, D.A., Huang, F., Barton, A. and Perera, B.J.C. (2014a) Statistical downscaling of general circulation model outputs to precipitation—part 1: calibration and validation. *International Journal of Climatology*, 34(11), 3264–3281. https://doi.org/10.1002/joc.3914.

Sachindra, D.A., Huang, F., Barton, A. and Perera, B.J.C. (2014b) Statistical downscaling of general circulation model outputs to precipitation—part 2: bias-correction and future projections. *International Journal of Climatology*, 34(11), 3282–3303. https://doi.org/10.1002/joc.3915.

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. and Becker, E. (2014) The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1.

Salami, H., Shahnooshi, N. and Thomson, K.J. (2009) The economic impacts of drought on the economy of Iran: an integration of linear programming and macroeconometric modelling approaches. *Ecological Economics*, 68, 1032–1039. https://doi.org/10.1016/j.ecolecon.2008.12.003.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B. and Ziese, M. (2011) GPCC full data reanalysis version 6.0 at 0.5: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. Deutscher, Wetterdienst: Global Precipitation Climatology Centre. https://doi.org/10.5676/DWD_GPCCFd_M_V6_05010.

Shirvani, A. (2017) Change in annual precipitation in the northwest of Iran. *Meteorological Applications*, 24, 211–218. https://doi.org/10.1002/met.1619.

Shirvani, A. and Landman, W.A. (2016) Seasonal precipitation forecast skill over Iran. *International Journal of Climatology*, 36, 1887–1900. https://doi.org/10.1002/joc.4467.

Shukla, S., Roberts, J., Hoell, A., Funk, C.C., Robertson, F. and Kirtman, B. (2019) Assessing North American Multimodel Ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa. *Climate Dynamics*, 53, 7411–7427. https://doi.org/10.1007/s00382-016-3296-z.

Slater, L.J., Villarini, G. and Bradley, A.A. (2019) Evaluation of the skill of North American Multi-Model Ensemble (NMME) global climate models in predicting average and extreme precipitation and temperature over the continental USA. *Climate Dynamics*, 53, 7381–7396. https://doi.org/10.1007/s00382-016-3286-1.

Smith, D.M., Scaife, A.A., Boer, G.J., Caian, M., Doblas-Reyes, F.J., Guemas, V., Hawkins, E., Hazeleger, W., Hermanson, L., Ho, C.K., Ishii, M., Kharin, V., Kimoto, M., Kirtman, B., Lean, J., Matei, D., Merryfield, W.J., Müller, W.A., Pohlmann, H., Rosati, A., Wouters, B. and Wyser, K. (2013) Real-time multi-model decadal climate predictions. *Climate Dynamics*, 41, 2875–2888. https://doi.org/10.1007/s00382-012-1600-0.

Taylor, K.E. (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192.

Tippett, M.K., Barlow, M. and Lyon, B. (2003) Statistical correction of central Southwest Asia winter precipitation simulations. *International Journal of Climatology*, 23, 1421–1433. https://doi.org/10.1002/joc.947.

Tippett, M.K., Goddard, L. and Barnston, A.G. (2005) Statistical-dynamical seasonal forecasts of central-southwest Asian winter precipitation. *Journal of Climate*, 18, 1831–1843. https://doi.org/10.1175/JCLI3371.1.

Weisheimer, A., Doblas-Reyes, F.J., Palmer, T.N., Alessandri, A., Arribas, A., Déqué, M., Keenlyside, N., MacVean, M., Navarra, A. and Rogel, P. (2009) ENSEMBLES: a new multi-model ensemble for seasonal-to-annual predictions—skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophysical Research Letters*, 36, L21711. https://doi.org/10.1029/2009GL040896.

Wilks, D.S. (2009) Extending logistic regression to provide full-probability-distribution MOS forecasts. *Meteorological Applications*, 16, 361–368. https://doi.org/10.1002/met.134.

Wilks, D.S. (2011) Statistical Methods in the Atmospheric Sciences. San Diego, CA: Academic Press.