Skill of South Asian Precipitation Forecasts in Multiple Seasonal Prediction Systems

Asia Regional Resilience to a Changing Climate (ARRCC)

Work Package 2: Strengthening Climate Information Partnerships – South Asia (SCIPSA)

April 2021

This study has been produced as part of the UK Aid funded Asia Regional Resilience to a Changing Climate (ARRCC) programme which is being delivered in partnership with the Met Office and World Bank. The study was conducted under Work Package 2: Strengthening Climate Information Partnerships – South Asia (SCIPSA) and was produced by Met Office with support from RIMES (Regional Integrated Multi-hazard Early Warning System for Africa and Asia) and representatives from National Hydrological and Meteorological Services involved in the ARRCC programme.

Lead Author

Jessica Stacey - Climate Services Scientist, Met Office

Reviewers and Contributors

Dr Philip Bett – Senior Scientist, Met Office
Francis Colledge – Senior Scientific Consultant, Met Office
Andrew Colman – Senior Climate Scientist, Met Office
Dr Joseph Daron – Science Manager, Met Office
Dr Richard Graham - Science Manager, Met Office
Tamara Janes – Science Manager, Met Office
Dr Richard Levine – Senior Scientist, Met Office
Dr Sreejith Op – Regional Climate Centre Pune, India Meteorology Department
Dr D. S. Pai – Regional Climate Centre Pune, India Meteorology Department
Dr Govindarajalu Srinivasan – RIMES

The author also acknowledges valuable discussion and contributions from:

Dr. Jehangir Ashraf Awan, Pakistan Meteorology Department
Dr Indira Kadel, Department of Hydrology and Meteorology, Nepal
Dr Abdul Mannan, Bangladesh Meteorology Department
Dr Jennifer Pirret, Met Office
Soma Popalzai, Afghanistan Meteorological Department

Skill of South Asia Precipitation Forecasts in Multiple Seasonal Prediction Systems

Contents

Executive Summary

An assessment of 12 dynamical seasonal prediction systems is conducted, assessing their ability to predict South Asian seasonal precipitation during the two key monsoon seasons; southwest (June to September (JJAS)) and northeast (October to November (OND)). This research has been conducted as part of the UK-aid funded Asia Regional Resilience to a Changing Climate (ARRCC) programme within the Strengthening Climate Information Partnerships South Asia (SCIPSA) project. ARRCC SCIPSA has a regional focus, supporting the South Asian Seasonal Climate Outlook Forum (SASCOF) and recognising the important role of Regional Climate Centre (RCC) Pune. SCIPSA also provides national support in four focus countries, primarily Afghanistan, Bangladesh, Nepal and Pakistan. Thus, following the skill assessment of the South Asia region, an in-depth analysis is performed for South Asia with emphasis on each of these four ARRCC focal countries.

The main objective of this study is to inform the model selection process for the seasonal forecast produced at SASCOF. This forum is held prior to each of the above two monsoon seasons (JJAS and OND), bringing together experts from various institutions, including Global Producing Centres (GPCs), RCC Pune and National Meteorological and Hydrological Services (NMHSs), to produce the regional consensus precipitation forecast for the respective seasons. The WMO have defined an initiative to make the process for producing the SASCOF forecast more objective and recommend that skilful dynamical prediction systems (also referred to as dynamical models) that are appropriately calibrated and combined should form a basis for the forecast. Such objective methods will enhance the robustness, reproducibility and traceability of the SASCOF forecast.

Based on various verification metrics computed for the period 1993 to 2016, most of the models are shown to possess positive skill in predicting South Asian precipitation variability, noting considerable spatial differences. The verification metrics used include correlation coefficients, the Relative Operating Characteristic (ROC) scores and reliability diagrams. During the JJAS season, models exhibit moderate to good skill for large swathes of central and northern India and Nepal, with correlations of 0.4 to 0.8 and relative operating characteristic (ROC) scores of 0.6 to 1.0. Correlations are much lower (less than 0.4) for much of the northwest and northeast, for example in Bangladesh. In contrast for the OND season, models possess higher skill in the northwest and far southeast with correlations of 0.4 to 0.8, whereas correlation does not exceed 0.4 for other parts of the region. Improvements in model performance are most imperative for areas where skill is low, but precipitation totals and year-to-year variability remain high, for example in Bangladesh during JJAS.

For both seasons, models are typically more skilful in locations where precipitation variability has a strong relationship with the El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) in the observations. In contrast, models typically have lower skill in areas where weaker relationships exist with ENSO and IOD, or there are very low precipitation amounts, for example in Afghanistan during the JJAS season. Furthermore, the models which simulate a stronger ENSO-precipitation relationship are typically more skilful.

The range in skill between models highlights the importance of using a Multi-Model Ensemble as a basis for the SASCOF regional forecast, rather than any individual model. Whereas at the country-level, there are clearly models that exhibit substantially more skill over others, and careful consideration should therefore be made when selecting models for the seasonal forecast. Additional analysis on the skill of an MME and ways to combine these models at both the regional and national levels is recommended as further work.

1. Introduction

South Asia is the most densely populated geographical region in the world and highly vulnerable to variability in precipitation. Anomalous dry or wet years can lead to widespread adverse impacts on livelihoods and the economy, especially in the agriculture (Ray et al., 2015; Aryal et al., 2020) and water (Srivastava et al., 2020) sectors. However, the intensity and distribution of precipitation possesses considerable spatiotemporal variability across the region. Seasonal forecasts, if skilful, can provide information on how the precipitation may deviate from normal several months ahead, and have the potential to support long-term planning decisions and provide advanced warning of adverse or potentially beneficial climatic conditions, such as droughts and floods (e.g. Golding et al., 2019; Daron et al., 2020). Climatic conditions in South Asia are predicted to become more extreme and unpredictable due to climate change, threatening food security and water availability (IPCC, 2014), and thus accurate seasonal forecasts will become even more imperative in the future.

1.1 Climate of South Asia

South Asia's immense geographical scale and varied topography leads to starkly different regional climates, ranging from arid desert in the northwest, alpine tundra and glaciers in the north, to humid tropical regions in the southwest. The South Asian monsoon is the principle source of precipitation for most of the region. The fundamental driving mechanism of the monsoon circulation is the pressure gradients (ocean to land) established by thermal contrasts between the land mass of Asia and the large extent of ocean to its south. The resulting flow carries huge amounts of moisture from the oceanic region to the land, which drives widespread, and occasional torrential precipitation, as the humid and unstable air is forced to rise over the land. South Asia experiences two distinct monsoon seasons based on the direction of the rain-bearing winds: the southwest (or summer) monsoon during the JJAS period and the northeast (or winter) monsoon during the OND period. Precipitation associated with the southwest monsoon does not reach the far northwest of South Asia, for example Afghanistan, where it remains largely dry during the JJAS season, as shown in Figure 2. Precipitation here is predominantly driven by low pressure systems, known as western disturbances, which originate in the extratropical North Atlantic as well as the Mediterranean and then move eastwards across the northwest of South Asia from October to June.

1.2 Predictability of seasonal precipitation in South Asia

Weather forecasts have seen considerable improvements in recent decades, yet the chaotic nature of the atmosphere still prevents skilful day-to-day forecasts past lead times of two to three weeks (Buizza & Leutbecher, 2015). However, forecasts of the *average* conditions of a

season are possible, and this is due to slowly evolving variations in lower-boundary forcing, for example changes in sea-surface temperatures (SSTs), soil moisture, sea ice and snow cover (Charney & Shukla, 1981). These slowly evolving variations can influence large scale atmospheric processes, and thus provide a source of predictability at long-range timescales. The atmosphere and ocean are most closely coupled in the tropics, and thus, prediction of precipitation on the seasonal timescales here is generally the most skilful (Kumar et al., 2013; Scaife et al., 2019).

It has been well established that seasonal prediction of South Asian precipitation is strongly linked with tropical SST anomalies (Goddard et al., 2001; Kucharski & Abid, 2017), especially in the central and eastern tropical Pacific; a phenomenon known as the El Niño Southern Oscillation (ENSO). Through changes to the Walker circulation (Bjerknes, 1969), during the southwest monsoon, El Niño events (warm phase of ENSO) tend to suppress precipitation and La Niña events (cool phase of ENSO) tend to enhance it (Pant & Parthasarathy, 1981; Rasmusson & Carpenter, 1983; Ju & Slingo, 1995). More recently, irregular oscillations in SSTs in the Indian Ocean, where the western part becomes alternately warmer (positive phase) or colder (negative phase) than the eastern part, known as the Indian Ocean Dipole (IOD; Saji et al., 1999), have also been linked to monsoon precipitation variability (Behera et al., 1999; Kripalani & Kumar, 2004). The IOD and ENSO are closely interconnected; the IOD has been thought to enhance or modulate the influence of ENSO on South Asia precipitation (Ashok et al., 2001). Whilst ENSO and IOD are the main source of predictability for the region (Johnson et al., 2017), many other drivers of variability have been identified, for example, SST and atmospheric patterns in the North Atlantic ocean (Wang et al., 2018; Yadav et al., 2009), springtime snow depth in the Himalayas (Hahn et al., 1976) and aerosols (Ramanathan et al., 2005). On shorter timescales, internal dynamics, mainly the eastward propagating Madden-Julian Oscillation and the northward propagating intraseasonal oscillation, known as the Boreal Summer Intraseasonal Oscillation, also drive active (enhanced precipitation) and break (modulated precipitation) phases with a period of 30 to 90 days (Goswami & Xavier, 2003).

For decades, seasonal forecasts for southwest monsoon precipitation have been produced using statistical methods (Walker, 1924; Gowariker et al., 1989; Rajeevan et al., 2004, 2007; van den Dool, 2006; Pai et al., 2017), but in recent years, advances in dynamical general circulation models (GCMs), have allowed them to become the predominant method for producing seasonal forecasts (for example, Pillai et al., 2018; Scaife et al., 2019). GCMs used for seasonal forecasts are commonly referred to as dynamical models or seasonal prediction systems. Methodologies which include a mixture of statistical and dynamical methods are also being increasingly utilised, as statistical postprocessing of the dynamical predictions is of interest to various users to give finer-scale information (Lang & Wang, 2010; Mohanty et al.,

2019; D. P. Walker et al., 2019). Canonical Correlation Analysis is a commonly used method, and the main technique available from the International Research Institute's Climate Predictability Tool (CPT)[1], which is a regularly used software package to help produce consensus forecasts at Regional Climate Outlook Forums (RCOFs), for example in East Africa (Kipkogei et al., 2017; Colman et al., 2019) and West Africa (Colman et al., 2017).

Numerous studies have assessed the ability of seasonal dynamical models to simulate and predict precipitation associated with the Asian monsoon on different timescales (for example, Ju & Slingo, 1995; Webster et al., 1998; Kim et al., 2012; Ramu et al., 2017; Pillai et al., 2018; Cash et al., 2019; Jain et al., 2019; Köhn-Reich & Bürger, 2019; Mohanty et al., 2019). While improvements have been made in predicting the Asian monsoon large scale flow patterns, especially with the introduction of coupled atmosphere ocean models, providing skilful predictions of seasonal precipitation over South Asia remains a challenge and is an active area of research. Capturing the relationship between precipitation variability and the SST teleconnections, such as ENSO and IOD, are critical for improving the skill of GCMs (Pillai et al., 2018).

For further reading, a comprehensive review of the climate of South Asia, the climate drivers and predictability of precipitation variability can be found in Stacey et al. (2019)[2].

1.3 National and Regional Seasonal Forecasts for South Asia

Supported by the World Meteorological Organisation (WMO), many locations around the world hold a Regional Climate Outlook Forum[3] (RCOF) to produce a consensus seasonal forecast for the upcoming season (Ogallo et al., 2008). In South Asia, this is known as the South Asian Seasonal Climate Outlook Forum (SASCOF). SASCOFs are held three times a year; once for each of the two monsoon seasons and one for the winter season (DJF). The forum is held in April for the JJAS forecast period (lead time of greater than 1 month), in September for the OND forecast period (lead time less than 1 month), and in November for the DJF season (lead time less than one month). The Indian Meteorological Department Pune (IMD Pune) are the WMO-designated Regional Climate Centre for the region and have been leading and co-ordinating the preparation and issuing of seasonal consensus forecasts for South Asia since 2010. The regional forecast is then refined by the National Meteorological and Hydrological Services (NMHS) in each country to produce a seasonal forecast specific to their country and

[1] https://iri.columbia.edu/our-expertise/climate/tools/cpt/

[2] https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/business/international/scipsa_review_seasonal_forecasting_south_asia_final.pdf

[3] https://public.wmo.int/en/our-mandate/climate/regional-climate-outlook-products

tailored to the requirements of different sector users, such as those in agriculture. An example of the forecast issued as part of the 2020 seasonal consensus outlook statement[4] is provided in Figure 1.

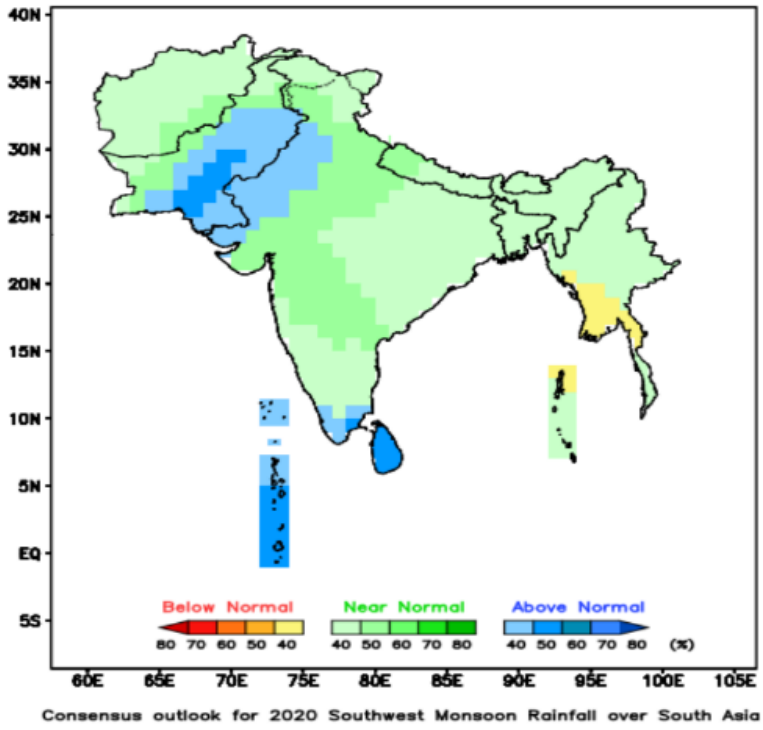

Figure 1. The regional consensus tercile 2020 JJAS precipitation forecast produced by IMD at the SASCOF in April 2020. Source: RCC, Pune SASCOF-17 statement.

Following a global review of the RCOFs, the WMO defined an initiative to move towards a more objective-based forecasting process (WMO, 2020). Since there are multiple models available on which to base the seasonal forecast, the WMO first recommend selecting a set of candidate models based on their regional skill and performance. The selected models can then be combined to produce an average of the forecast inputs, known as a Multi-Model Ensemble (MME). Statistical evidence shows that multi-model ensembles are generally a better predictor of observed climate than any single model over a long period of time (see the SPECS (2016) review for more on this topic).

For South Asia, there are limited studies that assess the skill and reliability of the seasonal forecast models used within the SASCOF process for the entire region. Furthermore, most skill assessments (e.g. Rajeevan et al., 2012; Jain et al., 2019) use the forecast with the most up to date initialisation month (i.e. May for the JJAS period), but to make it relatable to SASCOF we assess predictions for JJAS made in April. Holding the forum further in advance

[4] http://rcc.imdpune.gov.in/SASCOF17/final_SASCOF17_consensus_statement_JJAS_2020.pdf

of the season of interest allows more time for planning, and thus model skill assessment on this timescale is essential.

1.4 Outline of study

In this study, the skill and reliability of 12 dynamical seasonal prediction systems have been analysed in order to quantify each model's ability to capture year-to-year precipitation variability in South Asia. The aim is for this assessment to inform the model selection process when producing the seasonal forecast at the SASCOF, supporting an objective forecast methodology in line with WMO recommendations. Therefore, in this report, our focus is on the regional seasonal forecast for precipitation during the JJAS and OND seasons consistent with the SASCOF forecasts.

This research has been conducted as part of the UK-aid funded Asia Regional Resilience to a Changing Climate (ARRCC) programme under the Strengthening Climate Information Partnerships South Asia (SCIPSA) project. The purpose of the SCIPSA project is to bring together regional and national climate information providers, end-users and researchers to strengthen seasonal forecasting activities and advice services to vital sectors in the region. Whilst the ARRCC programme is regional in nature, it primarily focusses on supporting four focal countries: Afghanistan, Bangladesh, Nepal and Pakistan. Thus, this skill assessment covers the South Asia region, and a more in-depth skill analysis for each of the four ARRCC focal countries will follow.

The structure of this report will be as follows: Section 2 introduces the observational and model data used for this study and outline the methods used in calculating the verification metrics. Section 3 presents and discusses the results of the skill assessment and also explore the influence of the two main drivers of South Asian precipitation: ENSO and IOD. Finally, section 4, discusses and summarises the results, with suggestions for further work.

2. Data and Methods

The data and methods used to assess the performance of the seasonal forecast models are described in this section. There are various methods for assessing the performance of a dynamical seasonal forecast model, and this study focuses on those recommended by the World Meteorological Organization (WMO, 2019).

2.1 Observations

Observations have been taken from the Climate Hazards Group InfraRed Precipitation with Station dataset (CHIRPS) (Funk et al., 2015). The CHIRPS dataset is based on precipitation estimates derived from high-resolution satellite imagery, blended with station rain-gauge data to create a near real-time gridded daily precipitation time series from 1981 to present, downloaded as monthly aggregates. The dataset covers the tropics and sub-tropics (50°S to 50°N) for land points only and has been downloaded from the IRI Climate Data Library[5] at a resolution of 1.0° x 1.0° to align with model resolutions (it is also available at 0.05° x 0.05°). CHIRPS v2.0 was chosen as it covers the region and time period of interest, and is commonly used in the SASCOF forecast production process.

2.2 Seasonal prediction systems

The observational data has been compared with corresponding forecasts from 12 seasonal dynamical coupled prediction systems; information on each of these systems can be found in Table 1. The prediction systems, which will also be referred to as models throughout this document, were chosen based on the availability of the data at the time of analysis and previous use in the region. All data has been downloaded from the IRI Climate Data Library[5] at a resolution of 1.0° x 1.0°, and re-gridded to the CHIRPS dataset to ensure a consistent grid (original misaligned by 0.5°). A common hindcast period of 1993 to 2016 has been taken between all models. Since the purpose of this assessment is to inform the SASCOF process (see section 1.4 Outline of study), the target period and initialisation months have been chosen to align with those used in SASCOF. For the southwest monsoon, hindcasts are for the forecast period JJAS and initialised in April (two-month lead time). For the northeast monsoon, hindcasts are for the forecast period OND and initialised in September (one-month lead time). An ensemble mean has been taken from all available ensembles from each model; note that, as shown in

[5] http://iridl.ldeo.columbia.edu/

Table 1, the number of ensemble members varies greatly between models.

Table 1. List of dynamical seasonal prediction systems included in the skill assessment and their configurations (as of Feb 2021)

System name	**Centre / Country**	**Atmosphere horizontal resolution (at equator)/vertical levels**	**Ocean model name resolution /vertical levels**	**Hindcast Ensemble size**	**Hindcast Period**	**Reference**
CanCM4i (cmci)	Canadian Meteorological Center	Grid point: 1.41° E-W; 0.94° N-S (155km) 40 levels	CanOM4 1.41° E-W; 0.94° N-S 40 levels	10	1981 - 2018	von Salzen et al. (2013)
CMCC SPSv3	Italy	1° x 1°, 110km, 46 levels	NEMO3.4, 25km 50 levels	40	1993 - 2016	Sanna et al. (2017)
COLA-ccsm4	NCAR	Grid point: 0.9° E-W; 1.25° N-S (100km) 26 levels	POP2 1°, 0.3 near equator 60 levels	10	1982 - 2020	https://www.cesm.ucar.edu/models/ccsm4.0/
DWD GCFS2p0	Offenbach, Germany	1° x 1° (70km at around 50°N), 95 levels	0.4°, 40 levels	30	1993 - 2016	Fröhlich et al. (2021)

GEM-NEMO	Canada	Grid point: 1.4° E-W; 1.4° N-S (155km) 79 levels	NEMO v3.6 1°; 0.33° near equator 50 levels	10	1981 - 2018	IRI data library documentation[6]
GEOSS2S	NASA	0.5° x 0.5°, 40 levels	MOM5, 0.5°, 40 levels	4	1981 - 2016	Vernieres et al. (2012)
GFDL-A	NOAA	Cubed sphere: (50km E-W; 50km N-S) 32 levels	MOM5 1°, 0.33° near equator 50 levels	12	1980 - 2020	Kirtman et al. (2014)
GFDL-B	NOAA	Cubed sphere: (50km E-W; 50km N-S) 32 levels	MOM5 1°, 0.33° near equator 50 levels	12	1980 - 2020	Kirtman et al. (2014)
GloSea-5	Met Office, UK	Grid point: 0.83° E-W; 0.56° N-S (90km) 85 levels	NEMO v3.4 0.25°, 75 levels	28	1993 - 2017	Maclachlan et al. (2015)
Meteo-France 7	Meteo-France, France	~50km, 91 levels	NEMO v3.6, ORCA 1° grid, 75 levels	25	1993 - 2016	https://www.wmolc.org/contents2/index/Toulouse

[6] https://iridl.ldeo.columbia.edu/documentation/Models/NMME/CanSIPSv2/technote_cansips-v2_20190703_e.pdf

NCEP CFS2	NOAA, USA	~100km, 64 levels	GFDL MOM3, 1/3° x 1° in tropics, 40 levels	24	1982 - 2020	Saha et al. (2014)
SEAS-5	ECMWF	~36km, 91 levels	NEMO v3.4 ORCA 0.25° grid, 75 levels	25	1981 - 2016	Johnson et al (2019)

2.3 Pearson's correlation

To display the spatial variations of deterministic skill (e.g. Figure 4), the Pearson's correlation coefficient, as shown in Equation 1: Pearson's correlation coefficient., is calculated at each grid-point by comparing the total seasonal precipitation between the model hindcast ensemble mean *x* and observations *y* at each time step *i*. All references to 'correlation' and/or 'r' throughout this document use the Pearson's method.

Equation 1: Pearson's correlation coefficient.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2(y_i - \bar{y})^2}}$$

Stippling (i.e. dots) on maps highlights areas of statistical significance at the 95% two-tailed confidence interval. The correlations shown in the bar plots (e.g. Figure 6) for the whole region and each of the ARRCC focal countries have been calculated between the spatial mean precipitation from each of the models and the spatial mean precipitation from the observations.

Pearson's correlation has been chosen for its common usage and simplicity. For robustness, these have been compared to correlation maps using Kendall's tau (not shown), which relies on ranking and is less susceptible to extremes as it weights each year equally. The sets of results are spatially similar and therefore Pearson's correlation is deemed acceptable for this purpose.

2.4 ROC maps and diagrams

A commonly used verification metric to assess the skill of probabilistic seasonal forecasts is known as the relative operating characteristic (ROC) (Marzban, 2004; Mason, 1982). ROC scores quantify the skill in terms of whether a specific forecast "event" occurs. In this study, events are classified as precipitation totals falling into one of three tercile categories: above, near and below normal. The event is considered to be forecast if the forecast probability for that event exceeds a specific threshold. The corresponding observation is then checked to determine which of the following four categories the forecast falls into, as shown in Table 2.

Table 2: ROC score categories

	Event is observed	**Event is not observed**
Event is forecast	Hit	False Alarm
Event is not forecast	Miss	Correct Negative

Equation 2: Hit rate and false alarm rate used to produce a ROC curve

$$Hit\ Rate = \frac{Total\ hits}{Total\ observed\ occurences\ (hits\ plus\ misses)}$$

$$False\ Alarm\ Rate = \frac{Total\ false\ alarms}{Total\ non\ occurences\ (false\ alarms + correct\ negatives)}$$

ROC diagrams display the "hit rate" against the "false alarm rate", using the definitions of Equation 2. For each tercile category, these two rates are plotted against one another for each forecast probability threshold forming a ROC curve. A skilful forecast will have a ROC curve which bows towards the top-left of the plot as it maximizes the hit rate and minimizes the false alarm rate, whereas for a forecast with no skill, the curve will follow the diagonal line as the hit and false alarm rates would be equal. The overall ROC skill score can be calculated by quantifying the area under this curve, with a score of 1 for a perfect forecast and 0.5 for a forecast with no skill (no better than climatology i.e. always predicting a probability of 33% for each tercile category). ROC maps show the ROC score calculated at each grid-point for each of the three terciles. ROC skill maps have been generated for the South Asia region (e.g. Figure 5) as well as ROC curves for South Asia (e.g. Figure A3) and each of the ARRCC focal countries (e.g. Figure A4-8).

2.5 Reliability diagrams

Reliability diagrams (Hamill, 1997; Hartmann et al., 2002) are an important tool for deciding whether forecast probabilities for an event are a true reflection of the chance of the event occurring. Reliability diagrams (e.g. Figure A4-A8) match up the forecast probabilities for each tercile category, with the observed frequency of each tercile category given the forecast probability. A reliable forecast would give similar values for both forecast probability and observed frequency, and thus the points would be close to the diagonal (y=x). For example, say we collected the cases in our hindcast period where the seasonal forecast probability of above-normal precipitation was 40%, then we would count, for these cases, how many times out of those cases the precipitation actually fell into the above-normal tercile; if this was 40% of the time, the forecast would possess "perfect reliability". Reliability diagrams reveal systematic biases in forecast probabilities and may be used to correct such biases. For example, if a model forecasts the probability of a wetter than normal season to be 70%, but according to the reliability diagram it is usually only observed 60% of the time with this forecast probability, then we can instead issue a forecast probability of 60%.

Reliability diagrams can possess substantial noise due to small samples, especially when the verification period or spatial domain is not large. To reduce sampling issues, probability bins of 0.2 have been used in this study. Sharpness diagrams (see section 2.6 Sharpness diagrams) should be used alongside reliability diagrams to check that there are enough forecasts in each bin.

2.6 Sharpness diagrams

The histograms located beneath the reliability diagrams show, for each event, the total number of hindcasts within each probability bin over the reference period and at all grid-points, and unlike ROC and reliability diagrams, observations are not considered. Known as “sharpness” or “frequency” diagrams, these histograms have two purposes. First, they ensure there is sufficient data in all bins when assessing the forecast reliability to allow statistically significant interpretation. Second, they provide information on forecast “sharpness”, that is, when the hindcasts populate a range of probability bins, including bins away from the climatological norm (i.e. 33% for all terciles). A set of forecasts with low sharpness will rarely deviate from climatology – the frequency histogram will peak at the climate frequency with few hindcasts in higher or lower probability bins. In contrast, the greater the tendency for higher populations of the low and high probability bins, the greater the sharpness of the forecasts. Note that a forecast system based on random selection of event probabilities will be sharp. For forecasts to be skilful, they must possess sharpness and reliability, with the reliability curve lying within a margin of the diagonal (Graham et al., 2005).

Further information on ROC curves, reliability diagrams and other skill scores or metrics used in this study can be found in Wilks (2011).

2.7 Domain choice

This analysis covers the South Asia region, defined as 5-40°N 60-110°E. Due to the lack of observational data, the Maldives has not been included in this study.

As mentioned in section 1.4 Outline of study, more in-depth analyses have been provided for the ARRCC project focal countries using the following domains (a visual outline of each domain can be found in Figure A1):

- Afghanistan: 28.5-39.5°N, 60-76°E
- Bangladesh: 20-28.5°N, 84.5-95°E
- Nepal: 25-32°N, 79-90°E
- Pakistan:
    - North: 32-39°N, 68-79.5°E
    - South: 23-32°N, 60-74°E

In consultation with the country NMHSs, rectangular boxes surrounding the countries have been defined, rather than taking the area inside the borders only, for the following reasons. First, the climate is not bounded by country borders; it is important to assess the skill of models in capturing the large-scale climate features which affect the aforementioned countries by using larger domain boxes. Second, there are places where the precipitation falling in surrounding areas is also of interest, for example over mountains (of which there are ample in the South Asian region) where precipitation runs into the rivers flowing into the countries of interest; this is important for Bangladesh in particular. Also note that Pakistan has been divided into north and south domains due to its diverse climatology.

2.8 Teleconnection indices

In section 3.3 Drivers of South Asian precipitation variability, the relationship between South Asia precipitation and the most well-known climate drivers, ENSO and IOD, are investigated using the two different SST anomaly indices in Table 3. The Pearson's correlation has been used to calculate the simultaneous correlation between each of these indices and observed and model precipitation. Each of the indices has been averaged for the concurrent period as used for the precipitation, and hence, no time lag has been used in this study.

Table 3. List of teleconnection indices used in this study

Index	**Region**	**Definition**	**Data source**
ONI	5°N-5°S, 120°-170°W	The Oceanic Niño Index (ONI) is NOAA's primary indicator for monitoring ENSO over the Niño 3.4 region. A positive (negative) index over 5-months typically signifies an El Niño (La Niña) event.	NOAA Climate Prediction Center[7]
DMI	Difference between 50°E–70°E, 10°S–10°N and 90°E–110°E, 10°S-equator	The Indian Ocean Dipole Mode Index (IODMI) captures SST anomalies between the west and southeast of the tropical Indian Ocean.	NOAA Physical Science Laboratory[8]

[7] https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php
[8] https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmi.had.long.data

3. Results

3.1 Skill of models in predicting precipitation from June to September

In this section, we analyse various metrics to assess the skill and reliability of seasonal precipitation forecasts for the 12 seasonal prediction systems, first focused on the southwest monsoon season from June to September (JJAS). This is the wettest season for most of the region, with many areas receiving over 80% of their annual rainfall (Figure 2 and Figure 3). However, the far northwest, namely Afghanistan, remains predominantly dry during this season. Care should be taken when analysing areas with low precipitation amounts, as precipitation totals can often be more difficult for observations to capture, i.e. showers or more sporadic precipitation can be more easily missed by a rain gauge than an area of widespread dynamic precipitation. This observational uncertainty can potentially make skill results less reliable. Note that other factors can also contribute to observational uncertainty, for example if there are less gauges covering an area.

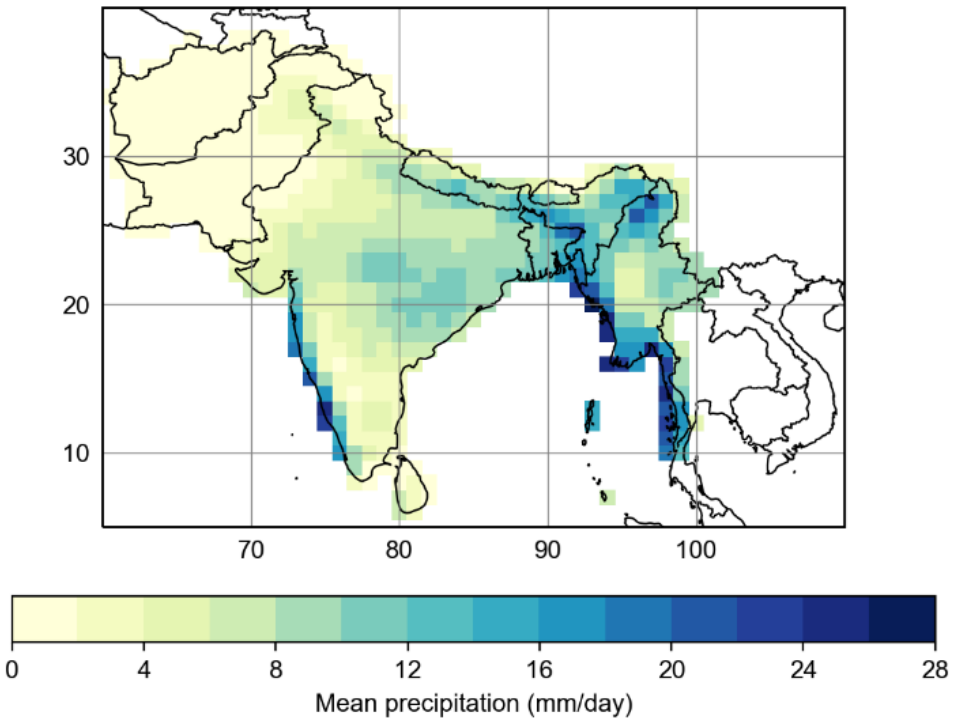

Figure 2. Mean precipitation across South Asia in mm/day for the JJAS season from 1993 to 2016 according to the CHIRPS observation dataset

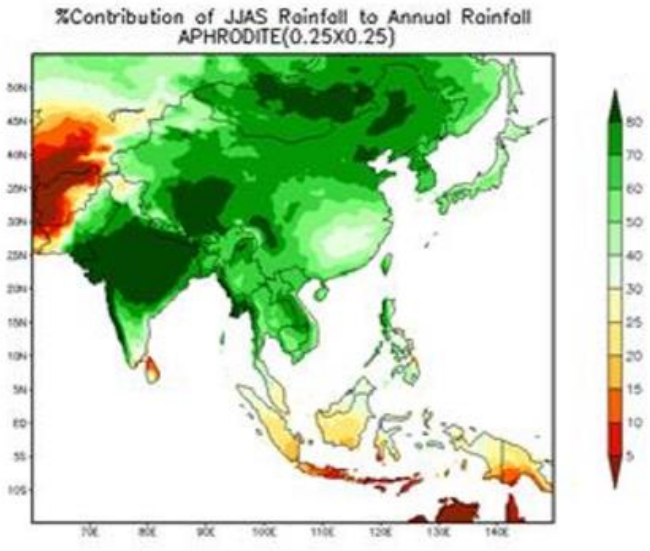

Figure 3. Percentage contribution of total annual precipitation during the JJAS season. Source: MMS and RIMES (2017). Data source: APHRODITE (0.25 x 0.25) data set (Yatagai et al., 2012)

3.1.1 Results for the South Asia Region in JJAS

The spatial plots of correlation and ROC scores for the South Asia region and JJAS season (Figure 4 and Figure 5) display considerable spatial variation between models, although there are some similarities. Compared to typical seasonal forecast performance, precipitation is well captured by almost all models for large swathes of India, especially in central and northern areas, and for much of Nepal, especially in the west, with correlations of 0.4 to 0.8 and significant at 5% level as well as ROC scores of 0.6 to 1.0 in places. Skill is lower in the far south of India and Sri Lanka, with correlation below 0.4 in most models, except for CFS where correlations exceed 0.4. The northwest of the region sees some areas of correlation exceeding 0.4 and significant, particularly in central Pakistan and north Afghanistan. In northwest India and northeast Pakistan, nearly all models exhibit an area of low correlation of below 0.4, and even negative and significant in some models. Correlation is variable in the east of the region over Bangladesh, Bhutan and northeast India, with weak correlation (below 0.4) and low ROC scores (below 0.6). Although in parts of Myanmar, mainly in the south, some models have pockets of moderate correlation (0.4 to 0.6).

Note that the ROC scores for the middle tercile can be found in appendix A2, and exhibit consistently lower scores than the other terciles for both seasons. This can be explained because in an above or below normal year, there are likely to be physical climate drivers (e.g. El Niño) giving a signal for conditions to be different from normal, whereas in the absence of climate forcing unpredictable internal variability is likely to dominate.

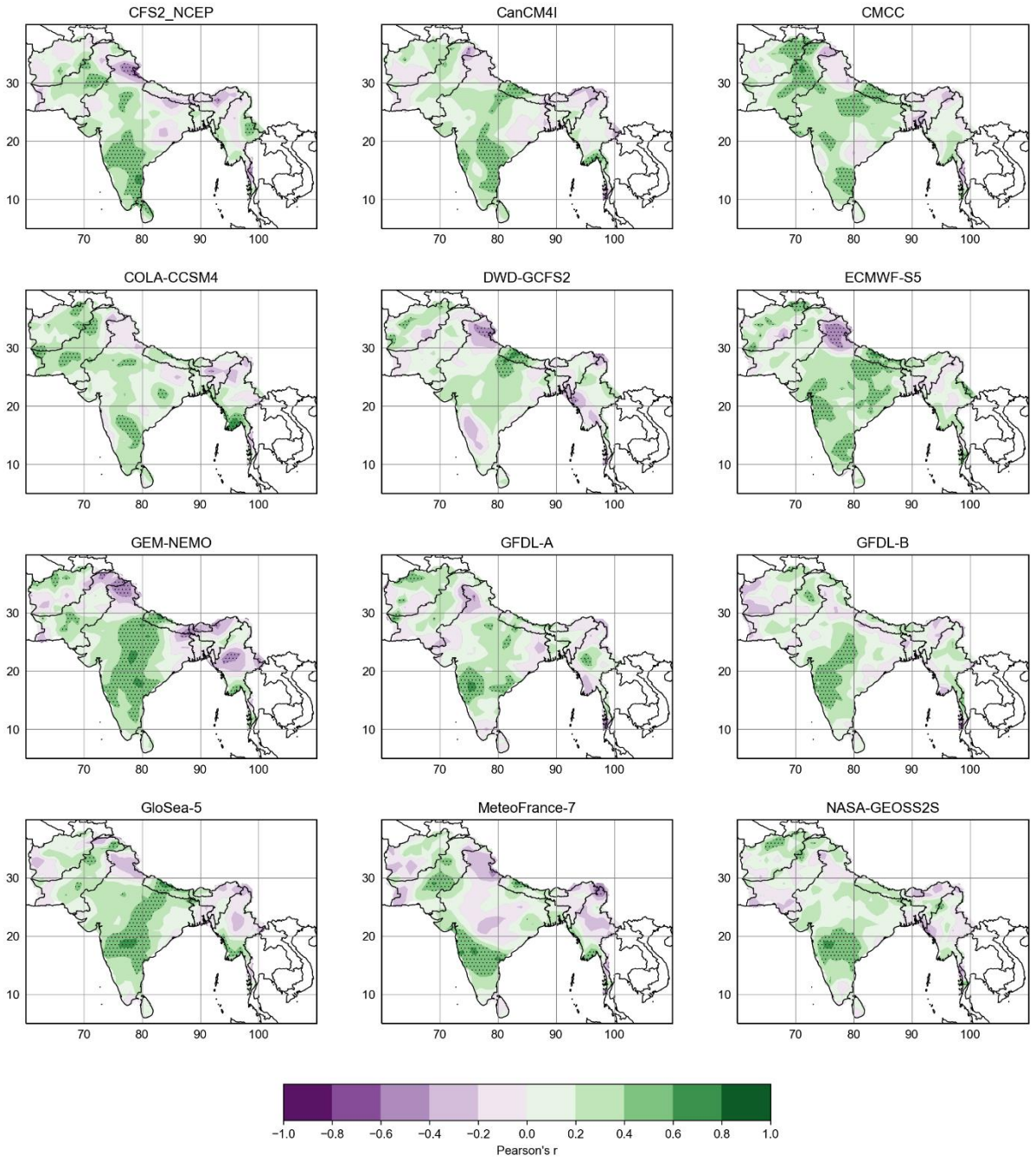

Figure 4. Pearson's correlation between precipitation in CHIRPS observations and the 12 different seasonal prediction systems listed for the JJAS season from 1993 to 2016. Stippling marks statistical significance at the 95% confidence level.

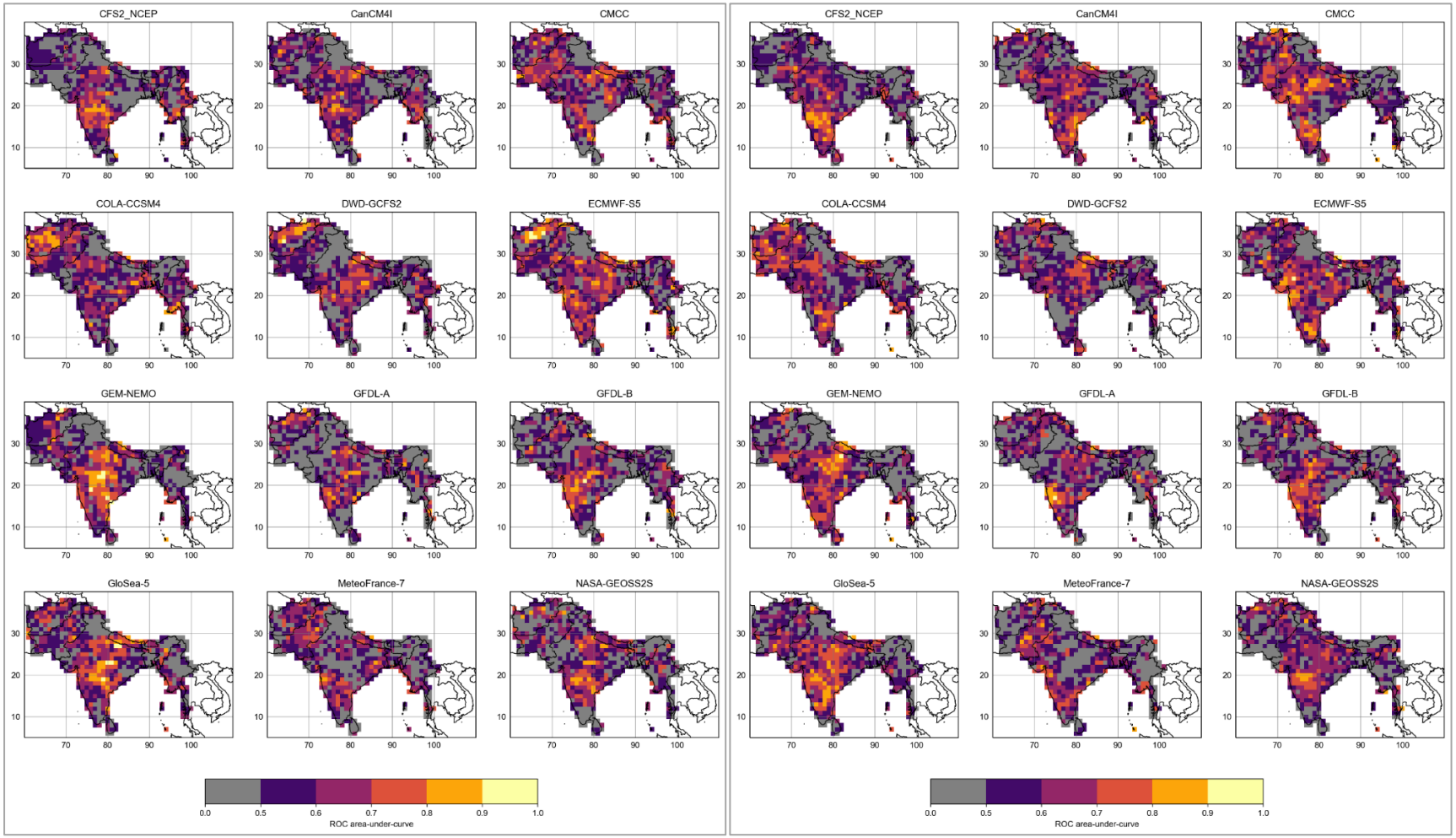

Figure 5: Maps showing the area under the ROC curve for the below-normal tercile (left) and above-normal tercile (right) for each model for JJAS precipitation. Values greater than 0.5 indicate a skilful forecast.

Variation in correlation is also high between models. Based on the analysis here, no single model is the most or least skilful everywhere; they all appear to have strengths and weaknesses in different areas. Figure 4 (top left panel) shows the correlations between the spatially averaged precipitation in observations and each of models over the entire South Asia region. CMCC, GEM-NEMO and ECMWF-S5 are the top three performers, whereas MeteoFrance-7 performs least well. However, the performance of models varies greatly depending on location shown by the country-specific bar plots Figure 6.

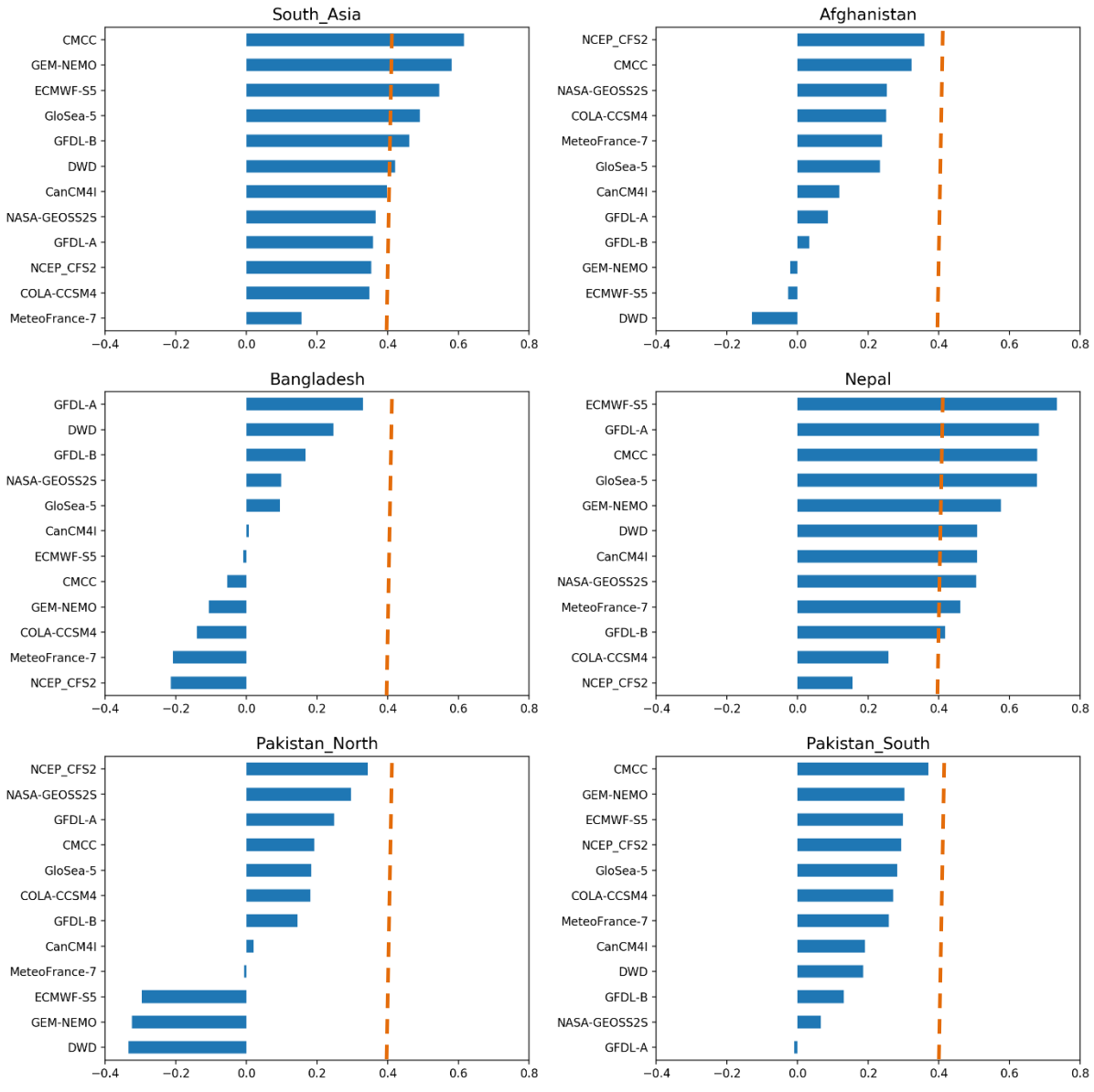

Figure 6. The 12 seasonal prediction systems ranked by Pearson's correlation between the spatially averaged precipitation in CHIRPS observations and each of the models for the South Asia region and country specific domains for the JJAS season from 1993 to 2016. Note, only correlations of above 0.4 (to the right of the orange dashed line) are significantly different from zero.

3.1.2 Results for the country-specific domains in JJAS

The ROC curves and reliability diagrams for the ARRCC focal countries, Afghanistan, Bangladesh, Nepal and Pakistan, indicate large variability between domains and models (Figures A4-A8). In general, the upper and lower terciles show greater skill and reliability than the middle tercile, corroborating the spatial ROC results in Figure 5 and A2. It's worth noting the small number of high probability forecasts in some of the models as shown by the frequency diagrams, and hence reliability results for these high probability forecasts should be regarded with caution. In the reliability diagrams, the forecasts are generally overconfident for all domains and models, shown by the lines with gradients less than the diagonal. This indicates that in subsets of years where the event has a high chance of occurring (i.e. it has a high frequency of occurrence) the forecast probability even higher, and in subsets where the event has a low probability of occurring (low observed frequency) the forecast probability is even lower. The following sections discuss the results for each of the ARRCC focal countries in turn.

Afghanistan

Unlike the majority of South Asia, JJAS is the dry season in Afghanistan. Unaffected by monsoon precipitation, prolonged droughts are common during this season (Qutbudin et al., 2019). Most models exhibit fairly low ROC skill (Figure A4) during JJAS, which may partly be explained by a lack of precipitation, and thus observational uncertainty in this season (Figure 2). CMCC and COLA-CCSM4 models exhibit the highest ROC skill, shown to be mainly in the northeast of Afghanistan by the spatial maps (Figure 4 and Figure 5). Furthermore, DWD-GCFS2 and ECMWF-S5 show positive skill for the below-normal tercile, supported by the spatial ROC maps which highlight areas of ROC scores exceeding 0.8 in parts of the north and west of the country by these models. Reliability is low for all models.

Bangladesh

Bangladesh receives more than 75% of its rainfall during the JJAS season, mainly driven by weak tropical depressions that originate in the Bay of Bengal (Shahid, 2010). Rainfall distribution varies considerably, with the lowest amounts in central-western parts and highest in eastern parts (Figure 2). Floods are a common occurrence in Bangladesh and result from excess rainfall falling both inside and outside the country (Ahasan et al., 2011).

Almost all models for the Bangladesh domain exhibit low ROC scores and reliability (Figure A5), and some no better than climatology (e.g. GEM-NEMO). According to the ROC and reliability plots, ECMWF-S5 has the greatest skill and reliability, with the ROC area under curve almost 0.6, followed by GloSea-5. However, the correlation of the spatially averaged precipitation (Figure 6) suggests GFDL-A and DWD are the best performing models, although

still not significant. The lack of skill in Bangladesh, also confirmed by the spatial skill maps and bar plots, highlights a need for improved understanding, regional observational datasets, and model development in this region. Potential reasons for the poor skill will be discussed in section 3.3 Drivers of South Asian precipitation variability.

Nepal

The summer monsoon is the predominant feature driving precipitation in Nepal, which receives approximately 80% of its precipitation in association during the JJAS season, with July and August being the wettest months. The central and eastern regions of Nepal receive more precipitation than the west (Figure 2), due to blocking and steering of flow and moisture by the steep Himalayan orography. Although there is large variation in spatial distribution from year to year (Shrestha, 2000).

Most of the models demonstrate high positive ROC skill over the Nepal domain, with ROC scores exceeding 0.6 for many models, in particular CanCM4I, ECMWF-S5 and GloSea-5 (Figure A6). Reliability is also good for some models, especially GloSea-5 for the below and above tercile categories. However, some models have much lower skill, including CFS2_NCEP and MeteoFrance-7. The spatial skill suggest the highest skill appears to be in western parts of Nepal (Figure 5).

Pakistan

Pakistan has a mostly arid climate, although it does remain humid in a small area in the north. The monsoon season occurs from July to September, when most rainfall falls in north and northeastern parts of the country, although the distribution of precipitation greatly varies from year to year (Chaudhry & Rasul, 2004). Extreme rainfall events in this northern area can cause severe flooding and loss of life during this season. For example, when an unusually strong monsoon depression travels all the way from the Bay of Bengal westwards to Pakistan.

The majority of models for Pakistan exhibit generally low skill (ROC scores less than 0.6) but poor reliability (Figures A7-A8). In the north domain, where the majority of Pakistan's precipitation falls in association with the southwest monsoon, ROC skill is generally low, but with some models showing good reliability, such as CMCC. In the south domain, some models possess positive ROC scores of above 0.6, namely CMCC, COLA-CCSM4 and ECMWF-S5, although reliability is largely poor. The spatial maps indicate that the areas of higher skill are mainly around the centre of the country.

3.2 Skill of models in predicting precipitation from October to December

In this section, we analyse various metrics measuring the skill and reliability of seasonal precipitation variability for the 12 seasonal prediction systems, considering the northeast monsoon season from October to December (OND). This is a much drier season for most of South Asia as the monsoon flows from the northeast, predominantly affecting the far south and southeast (Figure 7). Even though the amount of precipitation falling over Afghanistan and parts of Pakistan appears minimal in Figure 7, OND is an important season here as precipitation makes up 10-30% of their annual precipitation (Figure 8).

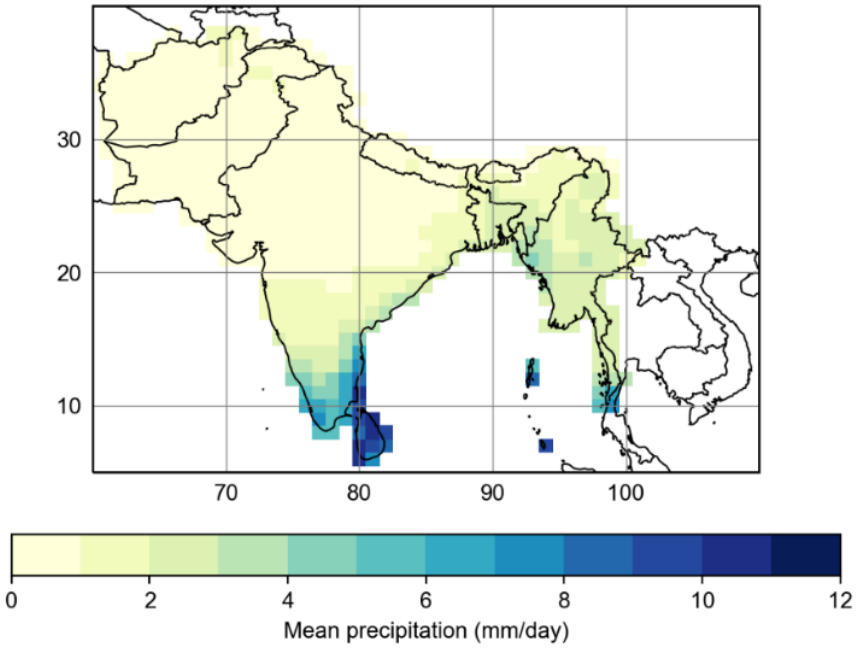

Figure 7. Mean precipitation across South Asia in mm/day for the OND season from 1993 to 2016 according to the CHIRPS observation dataset

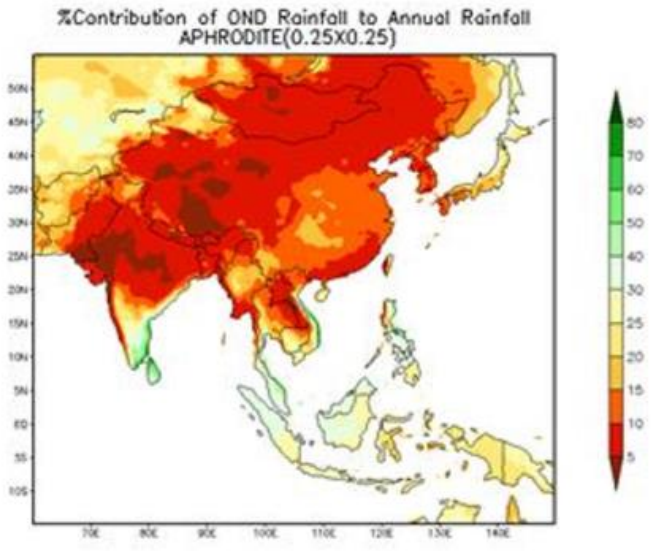

Figure 8. Percentage contribution of total annual precipitation during the OND season. Source: MMS and RIMES (2017). Data source: APHRODITE (0.25 x 0.25) data set (Yatagai et al., 2012)

3.2.1 Results for the South Asia Region in OND

As with the results for JJAS, there is considerable spatial variability across the region for the OND season, as shown in the spatial skill maps in Figure 9 and Figure 10. The areas of negative and weak correlation mostly correspond with the areas that receive very little precipitation during this season, such as in central and northern India, and Nepal. In the south and southeast of the region, where the highest rainfall occurs, the majority of models possess higher skill, with correlations exceeding 0.4. Almost all models show positive and significant correlation for parts of the northwest, particularly in Afghanistan and northern Pakistan, where ROC scores exceed 0.7 in places; although note that precipitation is still relatively low here and thus observational uncertainty should be taken into account. Models show contrasting results for the northeast of the region. For example, Bangladesh ranges from positive and significant correlation exceeding 0.4 in CFS2 and ECMWF-S5, to negative and significant correlation below -0.4 in CanCM4. Similarly, results in Bhutan, northeast India and Myanmar are mixed. The area in the far northwest of India which exhibited poor skill for the JJAS season has generally positive and, in a few models such as ECMWF-S5, significant skill for the OND period.

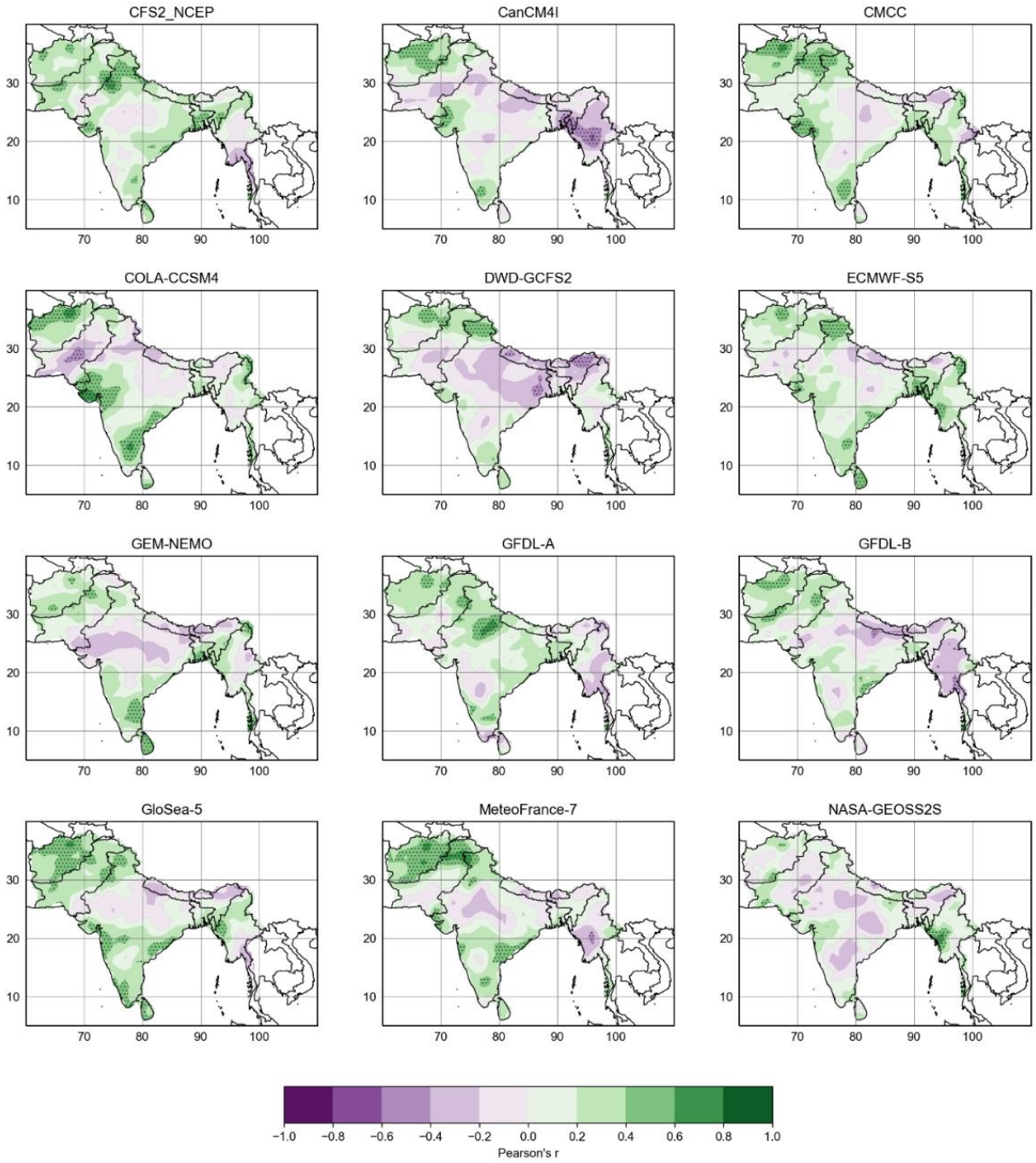

Figure 9. Pearson's correlation between precipitation in the CHIRPS observations and the 12 different seasonal prediction systems listed for the OND season from 1993 to 2016. Stippling marks statistical significance at the 95% confidence level.

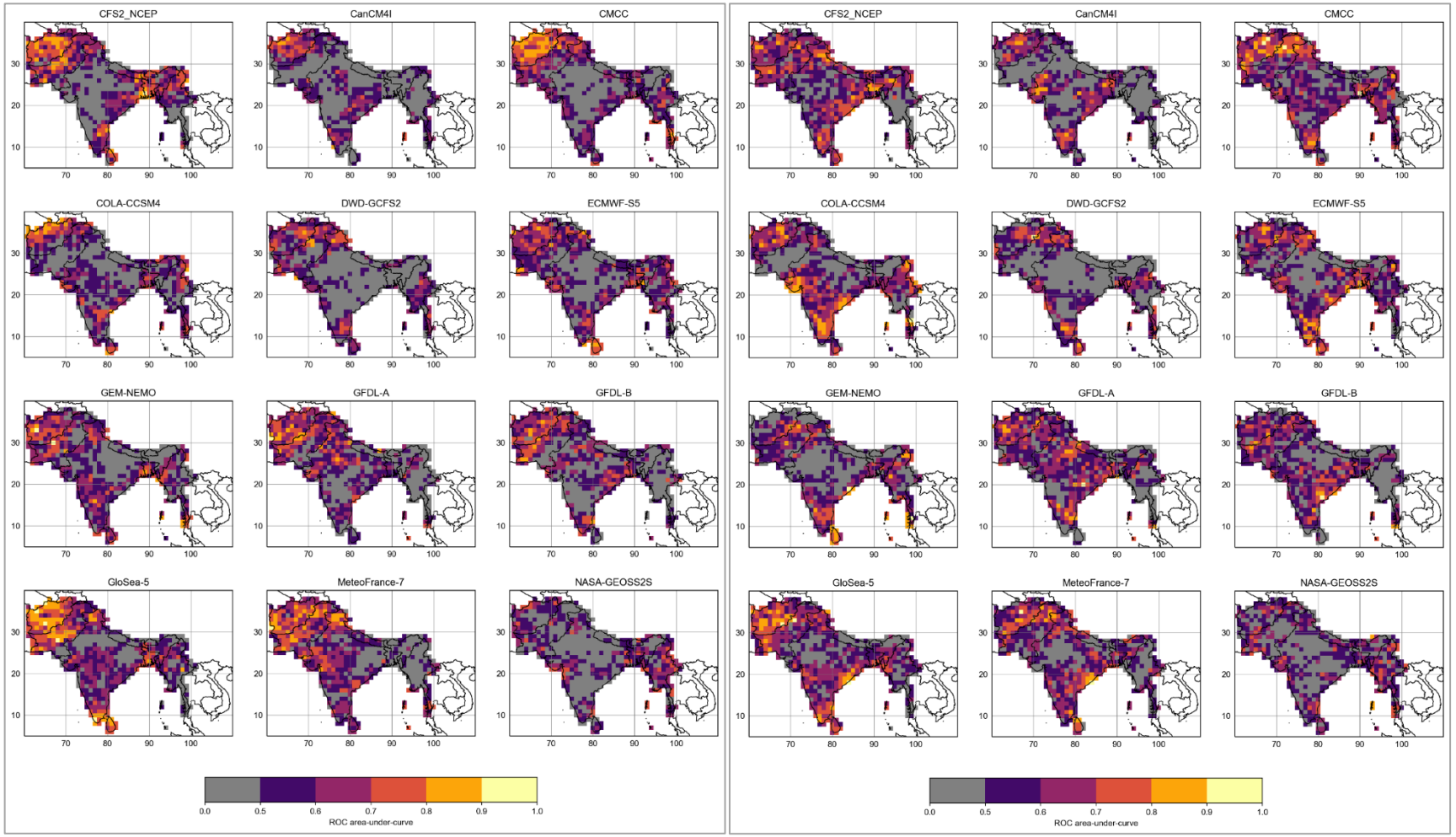

Figure 10: Maps showing the area under the ROC curve for the below-normal tercile (left) and above-normal tercile (right) for each model for OND precipitation. Values greater than 0.5 indicate a skilful forecast.

Similar to the JJAS season, there is considerable disparity between models, with no single model that performs well everywhere, but each exhibiting positive and negative correlations in different locations. By spatially averaging precipitation over the region, Figure 11 (top-left) shows the models with the highest average correlation over the region are MeteoFrance-7, COLA-CCSM4 and GFDL-A, whereas DWD and CanCM4I perform least well. Note how the high correlations for MeteoFrance-7 appear to be mostly driven by the very high correlations in Afghanistan and northern Pakistan, as elsewhere the skill pattern is similar to other models, highlighting the importance of assessing the average values alongside spatial maps (i.e. Figure 9). As with the JJAS season, these results demonstrate how models have varying strengths and weaknesses for the different country-specific domains.

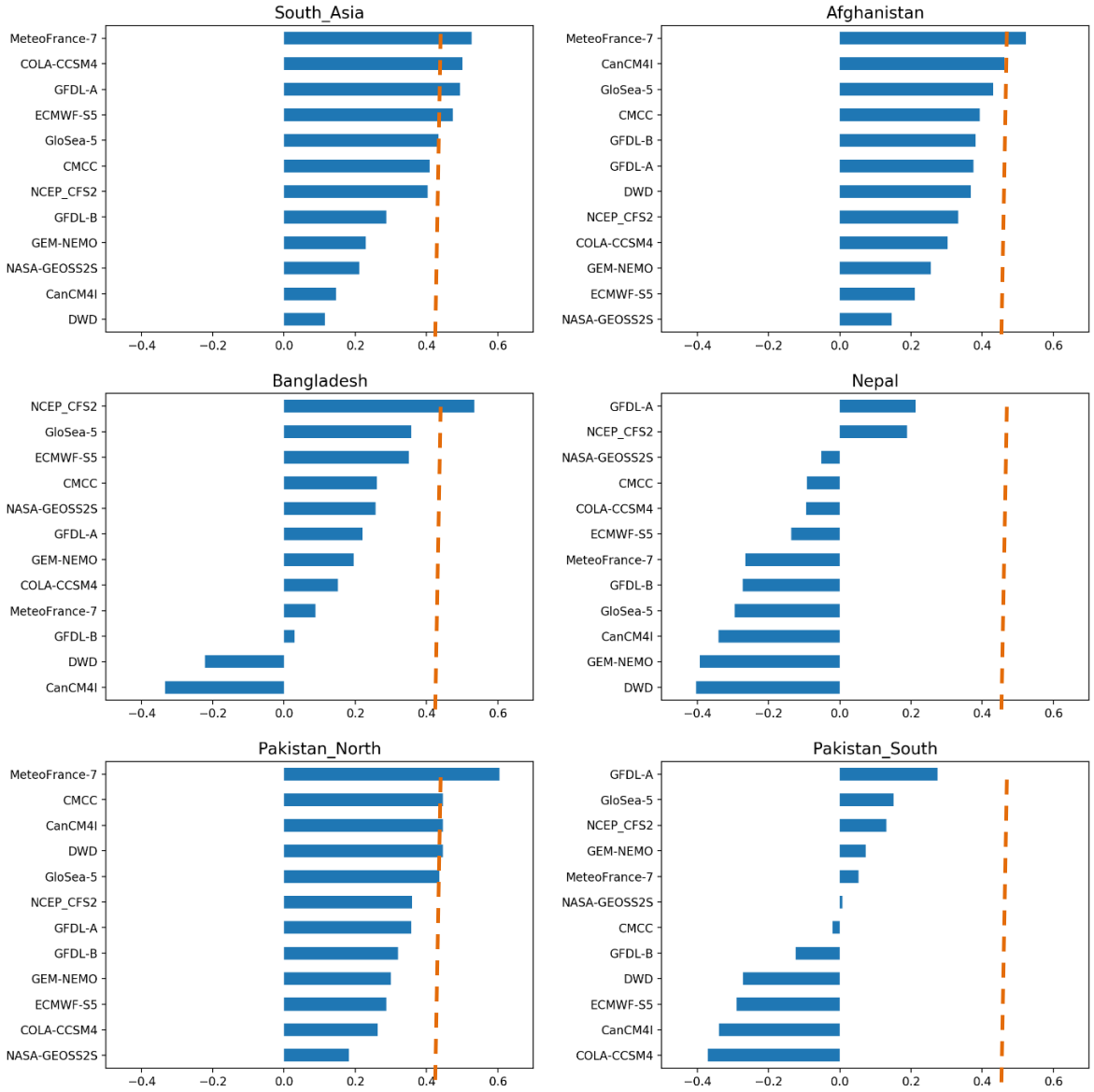

Figure 11. The 12 seasonal prediction systems ranked by Pearson's correlation between the spatially averaged precipitation in CHIRPS observations and the models for the South Asia region and country specific domains and the OND season from 1993 to 2016. Note, only correlations of above 0.4 (to the right of the orange dashed line) are significantly different from zero.

3.2.2 Results for the country-specific domains in OND

As seen in all other analyses, the ROC curves and reliability diagrams for the ARRCC focal countries (Afghanistan, Bangladesh, Nepal and Pakistan) indicate large variability between domains and models (Figures A11-A15). As expected, the upper and lower terciles generally show greater skill and reliability than the middle tercile. Again, only the section of the reliability diagram with a sufficient number of forecasts will be considered. As with the JJAS results, the forecasts generally appear overconfident, indicated by the lines with gradients less than the diagonal in the reliability diagrams. The following sections discuss more on the results for each country, referring to the ROC and reliability diagrams in Appendix 9, as well as the spatial skill maps showing Pearson's correlation (Figure 9) and the ROC skill (Figure 10).

Afghanistan

Western disturbances drive spells of precipitation across Afghanistan during the OND season, although most of these occur from December to March.

Almost all models show positive ROC skill and moderate reliability over Afghanistan (Figure A11), with the CMCC, GloSea-5 and MeteoFrance-7 models performing particularly well, with some ROC scores around 0.7. The NASA and COLA models perform less well according to these metrics. The bar plots show that about half of the models have average correlation exceeding 0.4, with MeteoFrance-7 performing particularly well (Figure 11). The spatial ROC maps illustrate how most of the models exhibit ROC scores of above 0.6 across large swathes of the country, in particular the north and west.

Bangladesh

From October to November, Bangladesh receives around 15% of its annual rainfall as the monsoon rains retreat (Rafiuddin et al., 2009).

As with the JJAS season, ROC and reliability plots for Bangladesh indicate a lack of skill in the region for the OND season (Figure A12). CFS shows the most skill and reliability, and the spatial skill maps suggest significant positive correlations of above 0.4 in the south and southeast of the country. The bar plots also show that CFS has significant correlation exceeding 0.5 over this domain.

Nepal

The northeast monsoon is marked by occasional short spells of precipitation in October, otherwise Nepal stays predominantly dry during the OND period. Although, occasional western disturbances can bring spells of precipitation to western Nepal from December (Sigdel & Ikeda, 2013).

As displayed in the correlation maps, the skill for Nepal over the OND season is much lower than for JJAS (Figure A13), although precipitation amounts are also much lower during this season. ROC skill scores are generally low (around 0.5), and reliability is also limited. GFDL-A and CFS2 (the upper tercile) exhibit positive skill, but this is weak and non-significant, which is supported by the bar plots in Figure 11. The spatial maps also signify low but positive skill uniformly distributed across Nepal.

Pakistan

Winter precipitation in Pakistan is mainly associated with western disturbances. Most of the precipitation falls from December to March, hence the mean precipitation in the OND season (Figure 7) is still fairly low. The majority of precipitation (including a large amount of snowfall) affects the north of Pakistan (Salma et al., 2012).

For the OND season, the models possess higher skill for the north domain compared to the south (Figure A14 and A15); noting that the opposite is the case for the JJAS season. The spatial maps corroborate these findings, with the highest skill in the north of the region for most models, corresponding with slightly higher precipitation totals. The models with the highest skill and reliability are CMCC, CFS2 and MeteoFrance-7 in the north and CMCC, GloSea-5 and MeteoFrance-7 (lower tercile) in the south. The bar plots also clearly demonstrate the higher skill in the north of the country than the south, with MeteoFrance-7 being particularly skilful in the north domain with a spatially-averaged correlation of 0.6 (Figure 11).

3.3 Drivers of South Asian precipitation variability

The WMO recommend that when selecting models for the SASCOF seasonal forecast, the selection process should also be based on their ability to simulate climate drivers that have an influence on the area of interest and relevant season (WMO, 2020). Therefore, in this section we will explore the link between precipitation and the two main climate drivers in the region: ENSO and IOD, for the seasons and regions of interest. As noted in section 1.2 Predictability of seasonal precipitation in South Asia, there are other important drivers of South Asian precipitation, but only links with ENSO and IOD are covered in this section.

3.3.1 The observed ENSO-precipitation relationship

The observed ENSO-precipitation relationship is investigated by calculating the Pearson's correlation coefficient between observed precipitation from 1993 to 2016 for the JJAS and OND seasons and the observed ONI index (see details in Table 3).

For the JJAS season, Figure 12 indicates the correlation between the ONI index and observed precipitation is negative for much of South Asia, particularly in Nepal and central and southern India. The largely negative correlation confirms the increased likelihood of anomalously dry (wet) conditions during the warm (cool) phase of ENSO (corresponding with positive (negative) values on the ONI index). The spatial pattern of correlation has similarities to the spatial skill plots for precipitation (Figure 4 and Figure 5). This is not surprising; a stronger ENSO teleconnection increases potential predictability, and models that can capture the teleconnection will tend to have better skill in the teleconnected area. In contrast, areas where the correlation is much weaker, for example in Bangladesh and northwest of the region, correspond with areas of lower skill in the spatial precipitation skill plots.

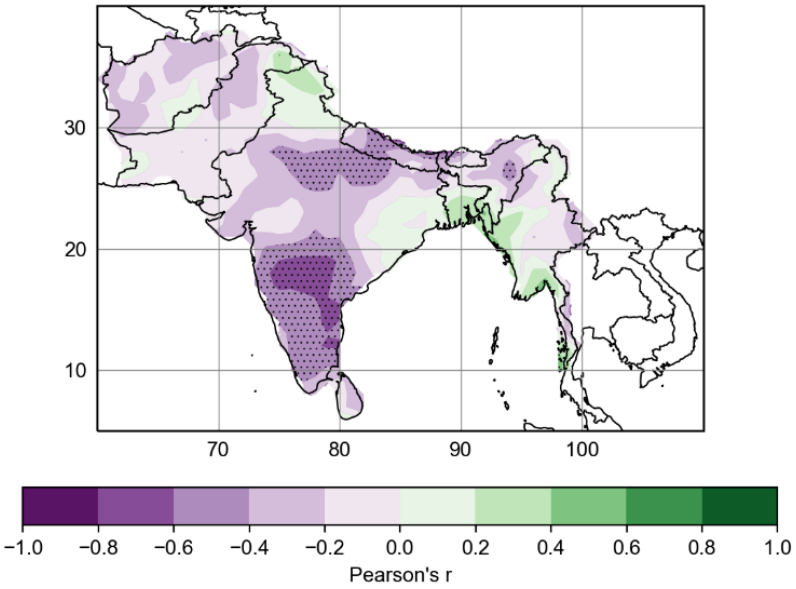

Figure 12. Map of Pearson's correlation between the ONI index and precipitation observations, both for the JJAS period (no time lag) from 1993 to 2016.

For the OND season, the relationship between the ONI index and observed precipitation is mixed (Figure 13), but as seen in JJAS, the areas which indicate a stronger ENSO teleconnection are generally the areas where models have better skill. A negative and, in places, significant correlation is present over the east of the region including Bangladesh, indicating a warm (cool) phase of ENSO is associated with less (more) precipitation. The opposite is true for Sri Lanka and the far northwest of the region, in particular Afghanistan and northern Pakistan, where much of the correlation is significant and positive. Elsewhere, a weaker ENSO teleconnection indicated by lower correlation values corresponds with areas of lower skill, including across large swathes of India and Nepal. The stronger (weaker) correlation values also appear to correspond with areas of higher (lower) mean precipitation (Figure 7).

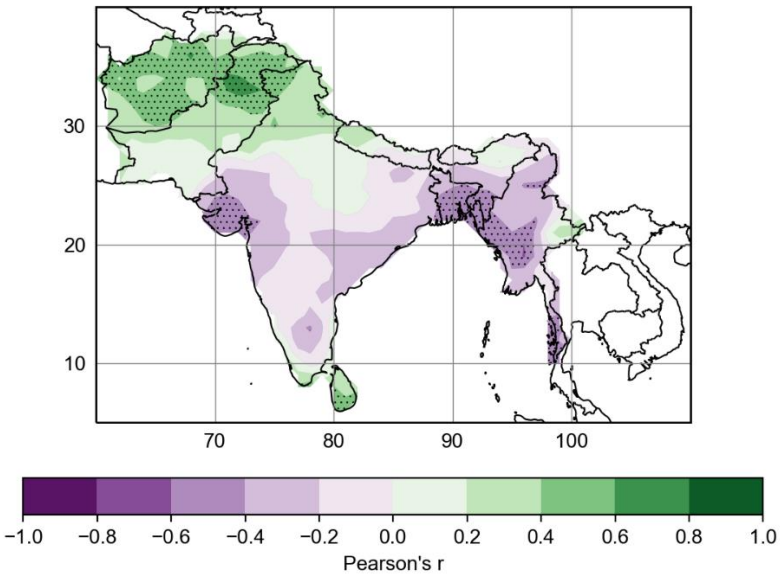

Figure 13. Map of Pearson's correlation between the ONI index and precipitation observations, both for the OND period (no time lag) from 1993 to 2016.

3.3.2 Model performance and their relationship with ENSO

Results from section 3.3.1 The observed ENSO-precipitation relationship suggest that the ONI index and precipitation correlation have a similar spatial pattern to the skill maps for the various models. In this section, we will explore the following question: do the models that have a stronger ENSO-precipitation relationship also have higher skill? To explore this question, model skill is plotted against the ENSO-precipitation relationship by comparing the correlation between the spatially averaged model and observed precipitation (x-axis) and correlation between the ONI index and spatially averaged model precipitation (y-axis). The Pearson's correlation coefficient (r) and p-value (p) between the two correlations have also been calculated. Uncertainty in these correlation statistics is large because of the small sample size therefore care should be taken not to infer anything too definitive from these r and p values. Furthermore, it is important to remember that the models will pick up on a number of other teleconnections other than ENSO and thus the relationship between model skill and either of these two drivers should not be overstated when analysing these plots.

When JJAS precipitation is averaged over the entire South Asia region, according to the results in Figure 14 (top-left panel), there is a negative correlation between the ENSO teleconnection and model skill. The negative slope signifies that when a model has a more negative correlation between ONI and precipitation, it tends to have higher skill. Interestingly, the observation line suggests that in the "real world" (albeit limited by the precision of the CHIRPS observations dataset), the ENSO-precipitation relationship is much weaker than

many of the skilful models suggest. The spatial plots of correlation between precipitation and ONI index for each of the models (Figure A16) also appear to exaggerate the negative correlation across a large swathe of South Asia, when compared to the correlation with observed precipitation (Figure 12). Therefore, even though the models appear oversensitive to the influence of ENSO, this oversensitivity also appears to improve their skill in forecasting precipitation. This is likely because the if the variability is driven by ENSO, then models that capture the ENSO response are likely to be more skilful than those that do not, even if they overestimate it. The correlation diagnostic may also amplify this effect. Note that, ENSO occurs in a minority of years, and thus much of the correlation score may come from just a few years in the series.

For the country-specific domains, the ENSO teleconnection appears to have less influence on model skill, with the exception being the Nepal domain (see Figure 14; middle-right panel), which has a negative correlation. The observation line also signifies a high correlation between observed precipitation and the ONI index, which is much closer to that of the most skilful model. Thus, these results, as well as those from the correlation between the observations in Figure 12, suggest that precipitation variability in Nepal is influenced by ENSO. The other domains for the ARRCC focal countries have much lower skill and a weaker link with ENSO. For the Afghanistan and Pakistan North domains, the slope of the line of best fit is positive (opposite to all other domains), albeit with weak correlation between the points. The observation line is near zero for both of these domains, suggesting ENSO has a weak influence in this area, and from our previous analyses shows that both precipitation amounts and skill are low here, especially in Afghanistan. However, interestingly the top performing model is CFS2 in both of these domains, and is the only model which has a highly positive ENSO-precipitation correlation. For Bangladesh and the southern Pakistan domains there is no significant relationship between model skill and the ENSO-precipitation relationship, which is likely due to the fact that the ENSO teleconnection has less influence here and thus there is less predictability for models to exploit.

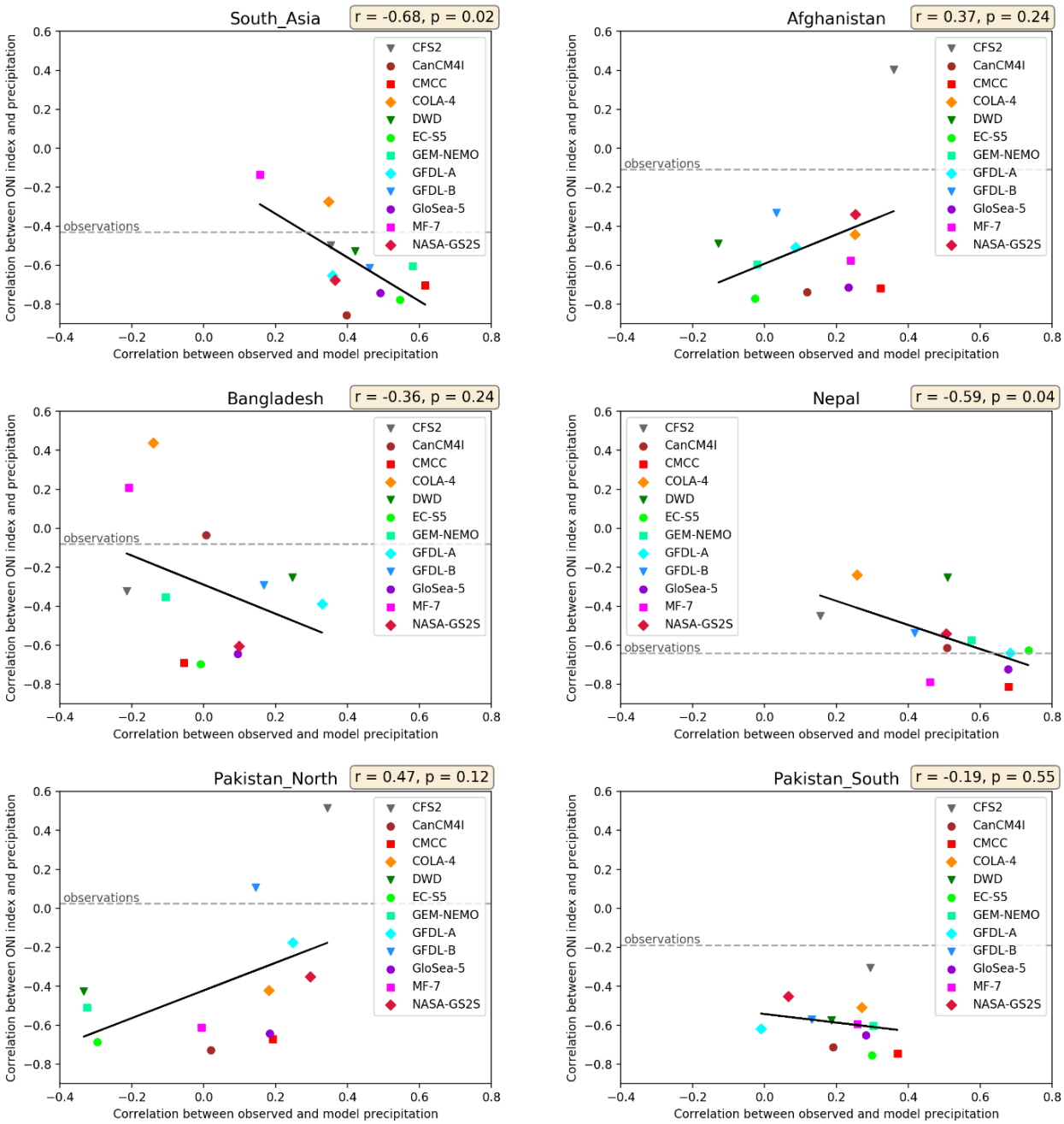

Figure 14. Scatterplot of correlation between the observed ONI index and model precipitation (y-axis) against correlation between observed and model precipitation (x-axis) for the JJAS season from 1993 to 2016. The correlation coefficient (r) and p-value (p) are stated in the box at the top-right of each plot; note that p<0.05 represents significance at 95% confidence level. The dashed grey line marked "observations" represents the correlation between the ONI index and observed precipitation. Precipitation is spatially averaged over South Asia and each of the country-specific domains in the plot titles. The black line represents the line of best fit between the 12 points.

The same plots for the OND period show contrasting results from those for JJAS, with no clear correlation present for the South Asia region (Figure 15; top-left panel). However, the spatial map of correlation between observed precipitation and the ONI index (Figure 13) suggests there is a significant relationship in places, with areas of negative and positive correlation. When spatially averaging the entire area, where neither positive or negative correlation dominates (unlike in JJAS where most of the area is dominated by a negative correlation) is likely to be the reason that the correlation between ONI and precipitation appears deceptively low (Figure 14; top-left panel).

The country-specific domains reveal there is a potential relationship between ENSO and precipitation for Afghanistan and Pakistan North domains. Most of the models appear to pick up on this relationship and have good positive skill, with the best performing models generally having a slightly stronger ENSO-precipitation relationship than the others. The results for the Pakistan South domain imply a positive relationship between model performance and the ENSO teleconnection. However, note that some models suggest a negative ENSO-precipitation correlation whilst the models with most skill suggest a positive correlation; closer to the observations line, which is weakly positive. For the Bangladesh domain, the ENSO teleconnection looks to have more of an influence in the OND season, compared to JJAS when no obvious ENSO relationship is apparent. The negative correlation implies that the models that have a stronger ENSO teleconnection have higher skill; somewhat surprising given the model results for the ROC and reliability plots (aside from the CFS2 model) in Figures A12-i and A12-ii are fairly poor. Negligible ENSO influence is detected by the results for the Nepal domain during OND, although this season experiences very little precipitation.

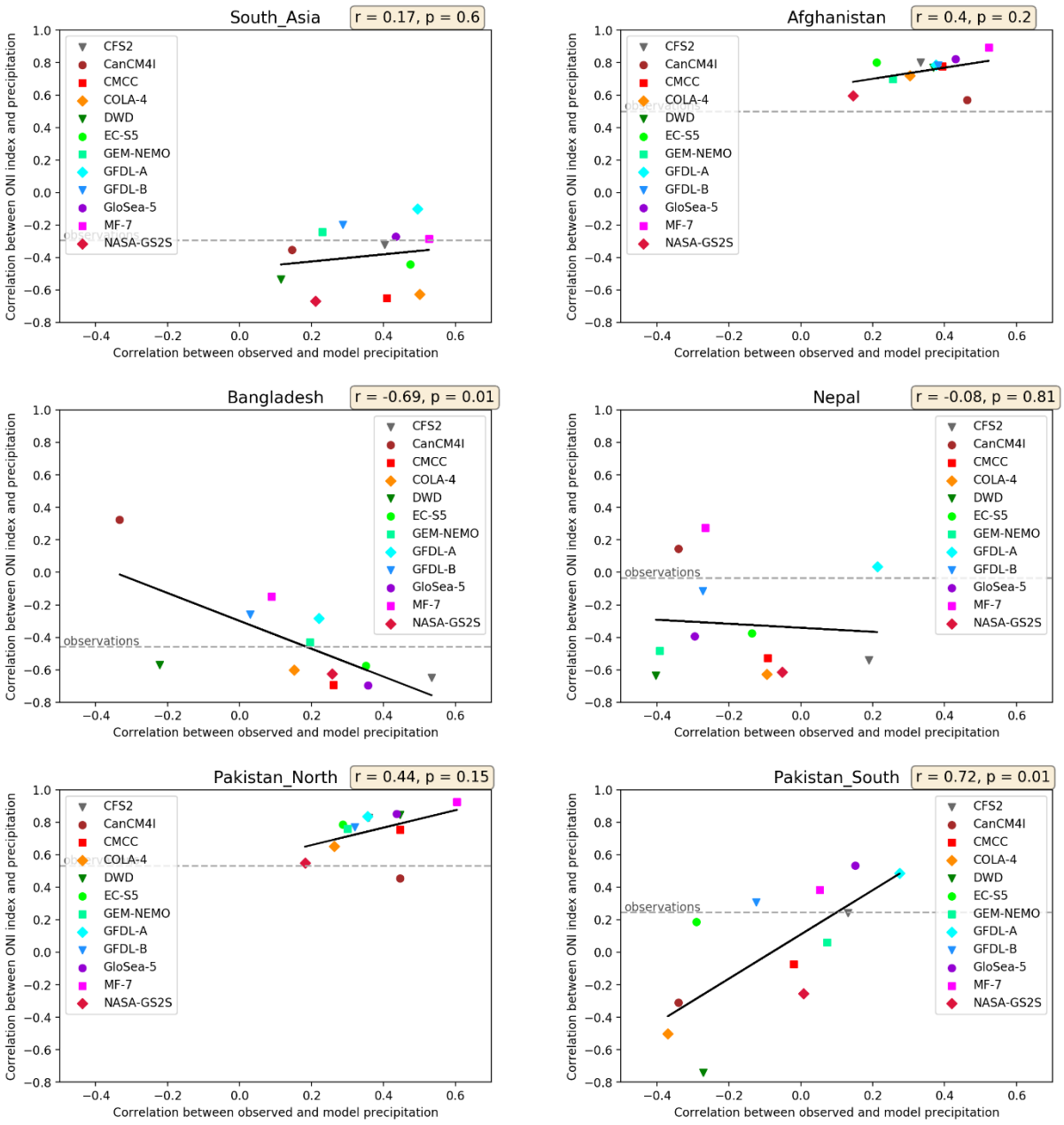

Figure 15. As caption for Figure 14, but for the OND season.

3.3.3 The IOD-precipitation relationship in observations and models

The relationship between the observed IOD-precipitation relationship is investigated by calculating the Pearson's correlation coefficient between observed precipitation from 1993 to 2016 for the JJAS and OND seasons and the observed IODMI index (see details Table 3) and shown in Figure 16.

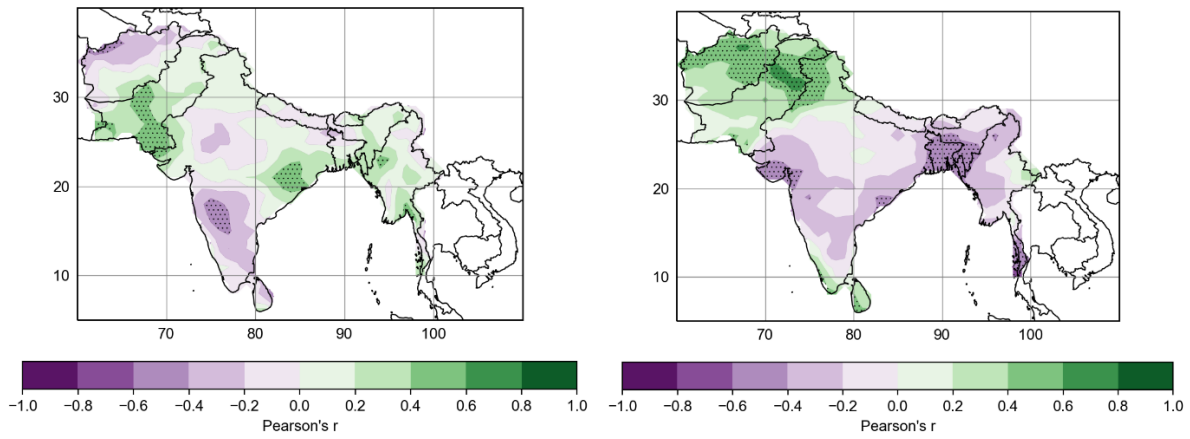

Figure 16. Map of Pearson's correlation between the IODMI index and CHIRPS observations, both for the JJAS period (left) and OND period (right) (no time lag) from 1993 to 2016.

There is a much weaker relationship between observed South Asian precipitation variability and the IOD compared with ENSO for the JJAS season, indicated by the low correlations of less than 0.4 or above -0.4 for much of the area (Figure 16; left). Studies suggest that, in general, positive (negative) IOD events correlate with increased (decreased) southwest monsoon precipitation totals over the monsoon trough region (Ashok et al., 2001, Behera et al., 1999), which on average runs from the north of the Bay of Bengal to west Rajasthan and adjoining Pakistan. This coincides with an area of positive and significant correlation in part of east India. For OND, there appears to be a more significant relationship; although interestingly, the pattern is remarkably similar to the ENSO plot in Figure 13 during this same season. The similar pattern suggests that a positive (negative) IOD event has a similar effect on South Asian precipitation to a El Niño (La Niña) event and/or that ENSO and IOD can co-occur (Cherchi & Navarra, 2013).

Scatterplots comparing the IOD and model precipitation relationship with model skill can be found for the JJAS and OND season in Figures A17-i and A17-ii respectively. For the JJAS season, as seen in the plots showing the observed precipitation-IOD relationship, the model precipitation-IOD relationship is generally weak and does not appear to be linked to model skill. Whereas for the OND season, the results look remarkably similar to those for the ENSO relationship during IOD in Figure 15, and therefore will not be discussed.

4. Summary and Recommendations for Further Work

4.1 Discussion and summary of results for South Asia

In this study, the ability of 12 dynamical prediction systems to capture precipitation variability in South Asia has been assessed for both the southwest (JJAS) and northeast (OND) monsoon seasons, with further analysis into how spatial variability in model skill varies with spatial variability in the strength of teleconnection with ENSO and IOD.

Most models have skill in predicting precipitation variability in South Asia, although there is considerable spatial variability. For instance, in the JJAS season, most models exhibit moderate to good skill for large swathes of central and northern India and Nepal, with correlations of 0.4 to 0.8 and ROC scores of 0.6 to 1.0, which may be explained by a strong ENSO teleconnection in these areas. Whereas in places where the connection with ENSO is less clear, such as Bangladesh and northwest India, most models exhibit much lower skill (correlations below 0.4). Moreover, many of the models exhibit relatively lower skill (correlations below 0.4) over parts of the monsoon trough region. The location of the trough line varies throughout the JJAS season, but on average, it follows a line from the north of the Bay of Bengal to west Rajasthan and adjoining Pakistan. This area is strongly influenced by intraseasonal variations caused by the north-south swing of the trough about its normal location and low-pressure systems moving along the trough axis. Models on the seasonal timescale are not always able to simulate the intraseasonal variability associated with the monsoon trough, hence the lower skill here. In contrast, most models for the OND season demonstrate moderate positive skill (correlations of 0.4 to 0.8) for Afghanistan and northern Pakistan, as well as the far southeast. As with the JJAS season, this corresponds with areas demonstrating stronger ENSO and IOD teleconnections. Skill in the northeast of the region is rather mixed, even though there appears to be a strong correlation with ENSO in places. Areas of lower skill (correlations less than 0.4) mainly correspond with areas of minimal precipitation during this season, such as Nepal and central India. Reliability is particularly poor for the majority of domains assessed over both seasons. However, reliability can be improved through calibration, and this is regularly done in SASCOF through the use of CPT.

There is considerable variation between models depending on the domain chosen and for different seasons. For example, the South Asia region bar plots of models ranked by Pearson's correlation (Figure 11) suggest MeteoFrance-7 has the lowest correlations for JJAS and highest for OND, signifying the stark differences in model skill from one period to another. In regard to developing the regional forecast for SASCOF, this clearly supports the WMO's case for using an MME (as explained in section 1.3 National and Regional Seasonal Forecasts for South Asia), and this analysis does not suggest that any of these 12 models should be

completely dismissed. However, at the country-level there are clearly models that exhibit substantially more skill over others. Careful consideration should therefore be made when selecting models for the seasonal forecast at the national level and the models that possess negative skill for the region and season of interest should be disregarded.

Most of the models show skill in capturing precipitation where an ENSO teleconnection is identified by the observations (Figure 12 and Figure 13). Furthermore, the models that simulate the strongest ENSO-precipitation relationship typically exhibit more skill; a finding that has been reported by other studies (e.g. Jain et al. (2019)). The IOD teleconnection is less pronounced than that for ENSO in the JJAS season. Interestingly in the OND season, the influence of the IOD on precipitation shows a very similar pattern to the influence of ENSO. The majority of IOD events tend to be associated with El Niño events and many studies have discussed the independence of the IOD from ENSO (e.g. Allan et al., 2001; Ashok et al., 2001). Thus, perhaps the similar correlation patterns are not hugely surprising during OND when both SST anomalies are particularly prevalent.

4.2 Discussion and summary of country-specific results

The strength of the ENSO and IOD teleconnections partly explain the differing results for each of the ARRCC focal countries. Next, we will summarise our findings and potential drivers for each of the countries in turn.

Afghanistan

In Afghanistan, models exhibit fairly poor skill for the JJAS season; an inconsequential result as this is a commonly dry season. Whereas for OND, spells of precipitation occur with the passage of western disturbances. Models perform better for this season, with both ENSO and IOD having a moderate to high positive correlation (0.4 to 0.8) with precipitation. Moreover, Afghanistan receives the majority of their precipitation from December to March, and therefore similar analysis focussing on subsequent months would be beneficial.

Bangladesh

Models perform particularly poorly for Bangladesh during JJAS, and unfortunately this is a season of substantial rainfall amounts and variability. For all models, correlations fail to exceed 0.4 and ROC scores and reliability are also low. Neither ENSO or IOD have a strong correlation with precipitation over Bangladesh in this season, which likely explains the poor model skill. Another potential reason could be the large intraseasonal rainfall variability, driven by the boreal summer intraseasonal oscillation (Fujinami et al., 2011), which is the dominant mode of rainfall variability during JJAS. Seasonal prediction systems are limited in their ability

to capture such short-range oscillations a season in advance (Lee et al., 2015). Although the OND season is a drier season than for JJAS as the monsoon withdraws, our findings indicate that some models exhibit moderate skill, with correlations exceeding 0.4 in the southeast of Bangladesh. The higher skill could be explained by an apparent stronger ENSO and IOD association with precipitation during OND than JJAS.

Nepal

The majority of models perform well over Nepal for the JJAS period, particularly in the west. Given the highly varied topography and climate of the region, the good skill is somewhat surprising, but ENSO appears to be a strong driver of Nepal's precipitation. The IOD appears to have less of an influence according to the relationship between the observed IODMI index and precipitation (Figure 16). Studies (e.g. Bohlinger & Sorteberg (2018) and Sharma et al. (2020)) agree there exists a strong negative correlation between ENSO and Nepal rainfall, which can be linked to the strengthening of the monsoonal trough over Nepal. Sharma et al. (2020) also find a strong significant negative relationship with the Indian Ocean from the preceding spring to concurrent summer. The models demonstrate much poorer skill during OND, although this is a predominantly dry season for Nepal and both ENSO and IOD are only very weakly correlated with OND precipitation.

Pakistan

In general, the seasonal prediction systems for Pakistan exhibit low skill for the JJAS season. However, the spatial plots suggest some areas of better skill in places, particularly around central areas, with correlations of 0.4 to 0.6 in places. Correlation with ENSO appears weak during this period, although an area of significant and positive correlation exists between the IOD and precipitation over the south of Pakistan (Figure 16). The scatterplots do not show any relationship between this teleconnection and model skill, thus suggesting that the models are not picking up on this potential teleconnection. A recent study by Syed et al. (2019) assessed statistical predictors of summer monsoon precipitation. They found potential significant teleconnections, including SST anomalies in the tropical Atlantic, the equatorial southeast Indian Ocean in northern Pakistan and the North Asia sea level pressure tendency in southern Pakistan. A statistical model based on these predictors was shown to exhibit good predictive skill for the primary monsoon region of Pakistan. The results for the OND season reveal more optimistic results, particularly in the north domain, with many of the models possessing correlations of 0.4 to 0.6 in the north. ENSO and IOD appear to be influential teleconnections here during OND.

4.3 Future recommendations for further study

Skill assessments, such as this one, take a snapshot in the performance of different seasonal prediction systems, providing information to inform the forecast produced in the SASCOF. Expanding this analysis in the following ways would also be beneficial:

1. Additional variables: The models' ability to capture other climate variables, such as temperature, would also be of interest for many sector users, for example those in the agriculture and health sectors, especially since there has been interest in including temperature in the SASCOF regional seasonal forecast. Another variable of great interest is monsoon onset timing. The ability of models to predict the exact date of the onset is still limited to 2-3 weeks in advance (Pradhan et al., 2017), although some models have been shown to make skilful onset predictions based on tercile categories, for example, early, late or normal (Chevuturi et al., 2019, 2021). Onset forecasts have also been produced by hybrid statistical/dynamical models using the Climate Predictability Tool (CPT) as part of the Greater Horn of Africa RCOF (GHACOF).
2. Additional models: Other seasonal prediction systems not used here should be included in future skill assessments, particularly those commonly included in the SASCOF forecast production such as the Japan Meteorological Agency seasonal system.
3. Additional seasons: Some parts of South Asia, namely Afghanistan and Pakistan, receive more precipitation in association with western disturbances from December to March. Therefore, this analysis could be extended to subsequent seasons for these areas.
4. Skill of multi-model ensemble combinations: Additional analysis on the skill of a multi-model ensemble at both regional and national levels, with experiments into different model combinations and how these compare with the skill scores of single skilful models, could be useful in informing the model selection process in the production of the national and regional seasonal forecasts.
5. Skill of the SASCOF regional forecast: A useful exercise would be to compare the skill of the dynamical models with the regional forecasts developed in SASCOF, as performed in West Africa by Pirret et al (2019) and in East Africa by Walker et al. (2019). Since the SASCOF forecast began being produced in 2010, there is an insufficient number of forecasts available for a robust skill assessment. Another barrier is that the SASCOF forecast consists of the dominant tercile category at each grid point and does not include information on the other two tercile categories. Reliable future verification of the skill of the SASCOF forecast would require the forecast probability of all three tercile categories. Including all tercile categories in future

SASCOF forecasts could also provide additional and potentially useful information for various users of the forecast.

6. Additional observation datasets with inclusion of Maldives: Examining observation uncertainty through the analysis of additional observations datasets would be valuable for assessing the trustworthiness of the skill results, particularly for regions with low precipitation amounts. Moreover, CHIRPS does not cover the Maldives, and thus, unfortunately, this study has not included this region. Therefore, future skill assessments should look to utilise station or precipitation datasets with ocean coverage which include the Maldives.
7. Country-specific guidance: The findings from this report could be tailored to produce country-specific guidance notes on the assessed skill of seasonal prediction systems to guide seasonal forecast production at the national-level.

4.4 Final conclusions

South Asia is particularly vulnerable to variability in seasonal precipitation amounts. Skilful seasonal forecasts can allow additional time to support long-term strategic planning, inform policymakers, and ultimately mitigate against the devastating impacts of floods and drought and capitalise on potential benefits from likely climatic conditions. However, throughout this study, it has been clear that further work is required to improve the skill of dynamical seasonal prediction systems in capturing South Asian precipitation variability. Gaps in model skill exist for differing locations and seasons, in particular over Bangladesh during their peak rainfall season (JJAS). The ability for models to accurately simulate the large-scale climate drivers and their complex relationship with South Asian precipitation will be a continued challenge for many years to come.

Fortunately, there are methods to enhance the skill of dynamical models as well as additional information that can support the production of regional and national seasonal forecasts, which can be particularly valuable in locations with poor model skill. First, statistical models, such as those developed in Pakistan by Syed et al. (2019), can provide additional information to support dynamical model output, especially where there are known predictors which have been shown to have a clear relationship with precipitation. Although, as with dynamical models, statistical methods have their limitations and solely relying on a statistical model should be done so with caution. Second, hybrid methods can be used to make use of the best characteristics of dynamical and statistical methods and can often enhance the skill and reliability of dynamical model output, for example through the use of CPT, as referenced in section 1.2 Predictability of seasonal precipitation in South Asia.

Furthermore, there is a need for further and continued assessments to monitor the changes in performance of seasonal prediction systems in the region. With continued research, more widespread and reliable observations, and model developments from improved resolutions, parameterisations, and data assimilation techniques, forecast skill will continue to improve into the future. Whilst improving our understanding of model skill is vital, future work should also focus on producing forecasts which are user-relevant in order to increase their uptake and accessibility. By working closely with sector users to improve understanding on how forecasts are interpreted and applied in practice, seasonal forecasts can be co-developed to have the greatest societal impact for the South Asia region.

5. References

Ahasan, M., Chowdhary, M. A., & Quadir, D. (2011). Variability and Trends of Summer Monsoon Rainfall over Bangladesh. *Journal of Hydrology and Meteorology*, *7*(1), 1–17. https://doi.org/10.3126/jhm.v7i1.5612

Allan, R. J., Chambers, D., Drosdowsky, W., Hendon, H., Latif, M., Nicholls, N., et al. (2001). Is there an Indian Ocean dipole and is it independent of the El Niño-Southern Oscillation? *CLIVAR Exchanges*, (August 2015), 18–22. Retrieved from http://eprints.soton.ac.uk/19295/01/ex21.pdf

Aryal, J. P., Sapkota, T. B., Khurana, R., Khatri-Chhetri, A., Rahut, D. B., & Jat, M. L. (2020). Climate change and agriculture in South Asia: adaptation options in smallholder production systems. *Environment, Development and Sustainability*, *22*(6), 5045–5075. https://doi.org/10.1007/s10668-019-00414-4

Ashok, K., Guan, Z., & Yamagata, T. (2001). Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO. *Geophysical Research Letters*, *28*(23), 4499–4502. https://doi.org/10.1029/2001GL013294

Behera, S. K., Krishnan, R., & Yamagata, T. (1999). Unusual ocean-atmosphere conditions in the tropical Indian Ocean during 1994. *Geophysical Research Letters*, *26*(19), 3001–3004. https://doi.org/10.1029/1999GL010434

Bjerknes, J. (1969). Atmospheric Teleconnections from the Equatorial Pacific. *Monthly Weather Review*, *97*(3), 163–172. https://doi.org/10.1175/1520-0493(1969)097<0163:ATFTEP>2.3.CO;2

Bohlinger, P., & Sorteberg, A. (2018). A comprehensive view on trends in extreme precipitation in Nepal and their spatial distribution. *International Journal of Climatology*, *38*(4), 1833–1845. https://doi.org/10.1002/joc.5299

Buizza, R., & Leutbecher, M. (2015). The forecast skill horizon. *Quarterly Journal of the Royal Meteorological Society*, *141*(693), 3366–3382. https://doi.org/10.1002/qj.2619

Cash, B. A., Manganello, J. V., & Kinter, J. L. (2019). Evaluation of NMME temperature and precipitation bias and forecast skill for South Asia. *Climate Dynamics*, *53*(12), 7363–7380. https://doi.org/10.1007/s00382-017-3841-4

Charney, J. G., & Shukla, J. (1981). Predictability of monsoons. *Monsoon Dynamics*, 99–109. https://doi.org/10.1017/cbo9780511897580.009

Chaudhry, Q.-U.-Z., & Rasul, G. (2004). Agro-Climatic Classification of Pakistan. *Quarterly Science Vision*, *9*(2), 3–4.

Cherchi, A., & Navarra, A. (2013). Influence of ENSO and of the Indian Ocean Dipole on the Indian summer monsoon variability. *Climate Dynamics*, *41*(1), 81–103. https://doi.org/10.1007/s00382-012-1602-y

Chevuturi, A., Turner, A. G., Woolnough, S. J., Martin, G. M., & MacLachlan, C. (2019). Indian summer monsoon onset forecast skill in the UK Met Office initialized coupled seasonal forecasting system (GloSea5-GC2). *Climate Dynamics*, *52*(11), 6599–6617. https://doi.org/10.1007/s00382-018-4536-1

Chevuturi, A., Turner, A. G., Johnson, S., Weisheimer, A., Shonk, J. K. P., Stockdale, T. N., & Senan, R. (2021). Forecast skill of the Indian monsoon and its onset in the ECMWF seasonal forecasting system 5 (SEAS5). *Climate Dynamics*, *1*, 3. https://doi.org/10.1007/s00382-020-05624-5

Daron, J., Allen, M., Bailey, M., Ciampi, L., Cornforth, R., Costella, C., et al. (2020). Integrating seasonal climate forecasts into adaptive social protection in the Sahel. https://doi.org/10.1080/17565529.2020.1825920

van den Dool, H. (2006). *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press. https://doi.org/10.1093/oso/9780199202782.001.0001

Fröhlich, K., Dobrynin, M., Isensee, K., Gessner, C., Paxian, A., Pohlmann, H., et al. (2021). The German Climate Forecast System: GCFS. *Journal of Advances in Modeling Earth Systems*, (May), 1–28. https://doi.org/10.1029/2020ms002101

Fujinami, H., Hatsuzuka, D., Yasunari, T., Hayashi, T., Terao, T., Murata, F., et al. (2011). Characteristic intraseasonal oscillation of rainfall and its effect on interannual variability over Bangladesh during boreal summer. *International Journal of Climatology*, *31*(8), 1192–1204. https://doi.org/10.1002/joc.2146

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 1–21. https://doi.org/10.1038/sdata.2015.66

Goddard, L., Mason, S. J., Zebiak, S. E., Ropelewski, C. F., Basher, R., & Cane, M. A. (2001). Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology*, *21*(9), 1111–1152. https://doi.org/10.1002/joc.636

Golding, N., Hewitt, C., Zhang, P., Liu, M., Zhang, J., & Bett, P. (2019). Co-development of a seasonal rainfall forecast service: Supporting flood risk management for the Yangtze River basin. *Climate Risk Management*, *23*(January), 43–49. https://doi.org/10.1016/j.crm.2019.01.002

Goswami, B. N., & Xavier, P. K. (2003). Potential predictability and extended range prediction of Indian summer monsoon breaks. *Geophysical Research Letters*, *30*(18). https://doi.org/10.1029/2003GL017810

Gowariker, V., Thapliyal, V., Sarkar, R., Mandal, G., & Sikka, D. (1989). Parameteric and power regression models : new approach to long range forecasting of monsoon rainfall in India. *Mausam*, *40*, 115–122.

Graham, R. J., Gordon, M., McLean, P. J., Ineson, S., Huddleston, M. R., Davey, M. K., et al. (2005). A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus A: Dynamic Meteorology and Oceanography*, *57*(3), 320–339. https://doi.org/10.3402/tellusa.v57i3.14666

Hahn, D. G., Shukla, J., Hahn, D. G., & Shukla, J. (1976). An Apparent Relationship between Eurasian Snow Cover and Indian Monsoon Rainfall. *Journal of the Atmospheric Sciences*, *33*(12), 2461–2462. https://doi.org/10.1175/1520-0469(1976)033<2461:AARBES>2.0.CO;2

Hamill, T. M. (1997). Reliability diagrams for multicategory probabilistic forecasts. *Weather and Forecasting*, *12*(4), 736–741. https://doi.org/10.1175/1520-0434(1997)012<0736:RDFMPF>2.0.CO;2

Hartmann, H. C., Pagano, T. C., Sorooshian, S., & Bales, R. (2002). Confidence Builders: Evaluating Seasonal Climate Forecasts from User Perspectives. *American Meterological Society*, 683–698. https://doi.org/10.1002/sfr.30692

IPCC. (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge: Cambridge University Press*.

Jain, S., Scaife, A. A., & Mitra, A. K. (2019). Skill of Indian summer monsoon rainfall prediction in multiple seasonal prediction systems. *Climate Dynamics*, *52*(9–10), 5291–5301. https://doi.org/10.1007/s00382-018-4449-z

Johnson, S. J., Turner, A., Woolnough, S., Martin, G., & MacLachlan, C. (2017). An assessment of Indian monsoon seasonal forecasts and mechanisms underlying monsoon interannual variability in the Met Office GloSea5-GC2 system. *Climate Dynamics*, *48*(5), 1447–1465. https://doi.org/10.1007/s00382-016-3151-2

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019

Ju, J., & Slingo, J. (1995). The Asian summer monsoon and ENSO. *Quarterly Journal of the Royal Meteorological Society*, *121*(525), 1133–1168. https://doi.org/10.1002/qj.49712152509

Kim, H.-M., Webster, P. J., Curry, J. A., & Toma, V. E. (2012). Asian summer monsoon prediction in ECMWF System 4 and NCEP CFSv2 retrospective seasonal forecasts. *Climate Dynamics*, *39*(12), 2975–2991. https://doi.org/10.1007/s00382-012-1470-5

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., et al. (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*(4), 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1

Köhn-Reich, L., & Bürger, G. (2019). Dynamical prediction of Indian monsoon: Past and present skill. *International Journal of Climatology*, *39*(8), 3574–3581. https://doi.org/10.1002/joc.6039

Kripalani, R. H., & Kumar, P. (2004). Northeast monsoon rainfall variability over south peninsular India vis-à-vis the Indian Ocean dipole mode. *International Journal of Climatology*, *24*(10), 1267–1282. https://doi.org/10.1002/joc.1071

Kucharski, F., & Abid, M. A. (2017). Interannual Variability of the Indian Monsoon and Its Link to ENSO. *Oxford Research Encyclopedia of Climate Science*, (November), 1–24. https://doi.org/10.1093/acrefore/9780190228620.013.615

Kumar, A., Chen, M., & Wang, W. (2013). Understanding prediction skill of seasonal mean precipitation over the tropics. *Journal of Climate*, *26*(15), 5674–5681. https://doi.org/10.1175/JCLI-D-12-00731.1

Lang, X., & Wang, H. (2010). Improving extraseasonal summer rainfall prediction by merging information from GCMs and observations. *Weather and Forecasting*, *25*(4), 1263–1274. https://doi.org/10.1175/2010WAF2222342.1

Lee, S. S., Wang, B., Waliser, D. E., Neena, J. M., & Lee, J. Y. (2015). Predictability and prediction skill of the boreal summer intraseasonal oscillation in the Intraseasonal Variability Hindcast Experiment. *Climate Dynamics*, *45*(7–8), 2123–2135. https://doi.org/10.1007/s00382-014-2461-5

Maclachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., et al. (2015). Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1072–1084. https://doi.org/10.1002/qj.2396

Marzban, C. (2004). *A Comment on the ROC Curve and the Area Under it as Performance Measures*. Retrieved from http://www.nhn.ou.edu/

Mason, I. (1982). A model for assessment of weather forecasts. *Aust. Meteorol. Mag. 30*, 291–303.

MMS, & RIMES. (2017). 11th South Asian Climate Outlook Forum, (September), 25–27.

Mohanty, U. C., Nageswararao, M. M., Sinha, P., Nair, A., Singh, A., Rai, R. K., et al. (2019). Evaluation of performance of seasonal precipitation prediction at regional scale over India. *Theoretical and Applied Climatology*, *135*(3–4), 1123–1142. https://doi.org/10.1007/s00704-018-2421-9

Ogallo, L., Bessemoulin, P., Ceron, J.-P., Mason, S., & Connor, S. J. (2008). Adapting to climate variability and change: the Climate Outlook Forum process. *Bulletin of the World Meteorological Organization*, *57*(2), 93–102. Retrieved from http://www.uonbi.ac.ke/openscholar/logallo/files/ogallo_en.pdf

Pai, D. S., Suryachandra Rao, A., Senroy, S., Pradhan, M., Pillai, P. A., & Rajeevan, M. (2017). Performance of the operational and experimental long-range forecasts for the 2015 southwest monsoon rainfall. *Current Science*, *112*(1), 68–75. https://doi.org/10.18520/cs/v112/i01/68-75

Pant, G. B., & Parthasarathy, S. B. (1981). Some aspects of an association between the southern oscillation and indian summer monsoon. *Archives for Meteorology, Geophysics, and*

Bioclimatology Series B, *29*(3), 245–252. https://doi.org/10.1007/BF02263246

Pillai, P. A., Rao, S. A., Ramu, D. A., Pradhan, M., & George, G. (2018). Seasonal prediction skill of Indian summer monsoon rainfall in NMME models and monsoon mission CFSv2. *International Journal of Climatology*, *38*, e847–e861. https://doi.org/10.1002/joc.5413

Pradhan, M., Rao, A. S., Srivastava, A., Dakate, A., Salunke, K., & Shameera, K. S. (2017). Prediction of Indian Summer-Monsoon Onset Variability: A Season in Advance. *Scientific Reports*, *7*(1), 1–14. https://doi.org/10.1038/s41598-017-12594-y

Qutbudin, I., Shiru, M. S., Sharafati, A., Ahmed, K., Al-Ansari, N., Yaseen, Z. M., et al. (2019). Seasonal drought pattern changes due to climate variability: Case study in Afghanistan. *Water (Switzerland)*, *11*(5). https://doi.org/10.3390/w11051096

Rafiuddin, M., Uyeda, H., & Islam, M. N. (2009). Characteristics of monsoon precipitation systems in and around Bangladesh. *International Journal of Climatology*, *30*(7), n/a-n/a. https://doi.org/10.1002/joc.1949

Rajeevan, M., Pai, D. S., Dikshit, S. K., & Kelkar, R. R. (2004). IMD's new operational models for long-range forecast of southwest monsoon rainfall over India and their verification for 2003. *Current Science*, *86*(3), 422–431.

Rajeevan, M., Pai, A. D. S., Kumar, A. R. A., & Lal, A. B. (2007). New statistical models for long-range forecasting of southwest monsoon rainfall over India. https://doi.org/10.1007/s00382-006-0197-6

Rajeevan, M., Unnikrishnan, C. K., & Preethi, B. (2012). Evaluation of the ENSEMBLES multi-model seasonal forecasts of Indian summer monsoon variability. *Climate Dynamics*, *38*(11–12), 2257–2274. https://doi.org/10.1007/s00382-011-1061-x

Ramanathan, V., Chung, C., Kim, D., Bettge, T., Buja, L., Kiehl, J. T., et al. (2005). Atmospheric brown clouds: impacts on South Asian climate and hydrological cycle. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(15), 5326–33. https://doi.org/10.1073/pnas.0500656102

Ramu, D. A., Rao, S. A., Pillai, P. A., Pradhan, M., George, G., Rao, D. N., et al. (2017). Prediction of seasonal summer monsoon rainfall over homogenous regions of India using dynamical prediction system. *Journal of Hydrology*, *546*, 103–112. https://doi.org/10.1016/j.jhydrol.2017.01.010

Rasmusson, E. M., & Carpenter, T. H. (1983). The Relationship Between Eastern Equatorial Pacific Sea Surface Temperatures and Rainfall over India and Sri Lanka. *Monthly Weather Review*, *111*(3), 517–528. https://doi.org/10.1175/1520-0493(1983)111<0517:TRBEEP>2.0.CO;2

Ray, D. K., Gerber, J. S., Macdonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, *6*(1), 1–9. https://doi.org/10.1038/ncomms6989

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (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

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*(6751), 360–363. https://doi.org/10.1038/43854

Salma, S., Rehman, S., & Shah, M. a. (2012). Rainfall Trends in Different Climate Zones of Pakistan. *Pakistan Journal of Meteorology*, *9*(17), 37–47. Retrieved from http://www.pmd.gov.pk/rnd/rnd_files/vol8_issue17/4.pdf

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N. S., Plummer, D., et al. (2013). The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes. *Atmosphere-Ocean*, *51*(1), 104–125. https://doi.org/10.1080/07055900.2012.755610

Sanna, A., Borrelli, A., Athanasiadis, P. J., Materia, S., Storto, A., Navarra, A., et al. (2017). RP0285 – CMCC-SPS3: The CMCC Seasonal Prediction System 3, (RP0285), 85. Retrieved from https://www.cmcc.it/publications/rp0285-cmcc-sps3-the-cmcc-seasonal-prediction-system-3

Scaife, A. A., Ferranti, L., Alves, O., Athanasiadis, P., Baehr, J., Dequé, M., et al. (2019). Tropical rainfall predictions from multiple seasonal forecast systems. *International Journal of Climatology*, *39*(2), 974–988. https://doi.org/10.1002/joc.5855

Shahid, S. (2010). Recent trends in the climate of Bangladesh. *Climate Research*, *42*(3), 185–193. https://doi.org/10.3354/cr00889

Sharma, S., Hamal, K., Khadka, N., & Joshi, B. B. (2020). Dominant pattern of year-to-year variability of summer precipitation in Nepal during 1987–2015. *Theoretical and Applied Climatology*, *142*(3–4), 1071–1084. https://doi.org/10.1007/s00704-020-03359-1

Shrestha, M. L. (2000). Interannual variation of summer monsoon rainfall over Nepal and its relation to Southern Oscillation Index. *Meteorology and Atmospheric Physics*, *75*(1–2), 21–28. https://doi.org/10.1007/s007030070012

Sigdel, M., & Ikeda, M. (2013). Seasonal Contrast in Precipitation Mechanisms over Nepal Deduced from Relationship with the Large-Scale Climate Patterns. *Nepal Journal of Science and Technology*, *13*(1), 115–123. https://doi.org/10.3126/njst.v13i1.7450

SPECS. (2016). Methodologies for calibration and combination of global and downscaled s2d predictions. *SPECS Project Report*, (308378), 1–52. Retrieved from https://library.wmo.int/doc_num.php?explnum_id=10314

Srivastava, A., Singhal, A., & Jha, P. K. (2020). Climate Change—Implication on Water Resources in South Asian Countries (pp. 217–240). Springer, Singapore. https://doi.org/10.1007/978-981-15-4668-6_12

Stacey, J., Richardson, K., Krijnen, J., & Janes, T. (2019). Seasonal Forecasting in South Asia: A Review of the Current Status. Retrieved from https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/business/international/scipsa_review_seasonal_forecasting_south_asia_final.pdf

Syed, F. S., Farah, I., & Awan, J. A. (2019). Revisiting the predictors for the long-range forecasting of southwest monsoon rainfall over South Asia with a focus on Pakistan. *Weather*, *74*(S1), S46–S56. https://doi.org/10.1002/wea.3528

Vernieres, G., Rienecker, M. M., Kovach, R., & Keppenne, C. L. (2012). The GEOS-iODAS: Description and Evaluation. Retrieved February 4, 2021, from https://ntrs.nasa.gov/citations/20140011278

Walker, D. P., Birch, C. E., Marsham, J. H., Scaife, A. A., Graham, R. J., & Segele, Z. T. (2019). Skill of dynamical and GHACOF consensus seasonal forecasts of East African rainfall. *Climate Dynamics*, *53*(7–8), 4911–4935. https://doi.org/10.1007/s00382-019-04835-9

Walker, G. T. (1924). Correlation in Seasonal Variations of Weather, IX. A Further Study of World Weather. *Memoirs of the India Meteorological Department*, *24*(9), 275–333.

Wang, Z., Yang, S., Lau, N. C., & Duan, A. (2018). Teleconnection between summer NAO and East China rainfall variations: A bridge effect of the Tibetan Plateau. *Journal of Climate*, *31*(16), 6433–6444. https://doi.org/10.1175/JCLI-D-17-0413.1

Webster, P. J., Magaña, V. O., Palmer, T. N., Shukla, J., Tomas, R. A., Yanai, M., & Yasunari, T. (1998). Monsoons: Processes, predictability, and the prospects for prediction. *Journal of Geophysical Research: Oceans*, *103*(C7), 14451–14510. https://doi.org/10.1029/97JC02719@10.1002/(ISSN)2169-9291.TOGA1

Wilks, D. S. (2011). Statistical Methods in the Atmospheric Sciences - Third edition. Retrieved December 21, 2020, from https://books.google.co.uk/books?hl=en&lr=&id=IJuCVtQ0ySIC&oi=fnd&pg=PP2&ots=apFrsSzN

KY&sig=cTqu84dmYIqYt2M1nj3Xl8izU8E&redir_esc=y#v=onepage&q&f=false

WMO, W. M. O. (2019). *Manual on the Global Data-processing and Forecasting System: Annex IV to the WMO Technical Regulations* (2019 editi). Geneva: WMO.

WMO, W. M. O. (2020). *Guidance on Operational Practices for Objective Seasonal Forecasting 2020*. Retrieved from https://library.wmo.int/doc_num.php?explnum_id=10314

Yadav, R. K., Kumar, K. R., & Rajeevan, M. (2009). Increasing influence of ENSO and decreasing influence of AO / NAO in the recent decades over northwest India winter precipitation, *114*(June), 1–12. https://doi.org/10.1029/2008JD011318

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., Kitoh, A., et al. (2012). APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bulletin of the American Meteorological Society*, *93*(9), 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1