Skilful seasonal predictions of Global Monsoon summer precipitation with DePreSys3

ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Skilful seasonal predictions of global monsoon summer precipitation with DePreSys3

To cite this article: Paul-Arthur Monerie *et al* 2021 *Environ. Res. Lett.* **16** 104035

IOP Publishing Environ. Res. Lett. 16 (2021) 104035 https://doi.org/10.1088/1748-9326/ac2a65

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
25 May 2021

REVISED
24 September 2021

ACCEPTED FOR PUBLICATION
27 September 2021

PUBLISHED
7 October 2021

LETTER

Skilful seasonal predictions of global monsoon summer precipitation with DePreSys3

Paul-Arthur Monerie[1], **Jon I Robson**[1], **Nick J Dunstone**[2] **and Andrew G Turner**[1]

[1] Department of Meteorology, National Centre for Atmospheric Science (NCAS), University of Reading, Reading, United Kingdom
[2] Met Office Hadley Centre, Exeter, United Kingdom

E-mail: p.monerie@reading.ac.uk

Keywords: DePreSys3, seasonal prediction, global monsoon, skill, summer monsoon precipitation, dynamic/thermodynamic

Supplementary material for this article is available online

Abstract

We assess skill of the Met Office's DePreSys3 prediction system at forecasting summer global monsoon precipitation at the seasonal time scale (2–5 month forecast period). DePreSys3 has significant skill at predicting summer monsoon precipitation ($r = 0.68$), but the skill varies by region and is higher in the northern ($r = 0.68$) rather than in the southern hemisphere ($r = 0.44$). To understand the sources of precipitation forecast skill, we decompose the precipitation into several dynamic and thermodynamic components and assess the skill in predicting each. While dynamical changes of the atmospheric circulation primarily contribute to global monsoon variability, skill at predicting shifts in the atmospheric circulation is relatively low. This lower skill partly relates to DePreSys3's limited ability to accurately simulate changes in atmospheric circulation patterns in response to sea surface temperature forcing. Skill at predicting the thermodynamic component of precipitation is generally higher than for the dynamic component, but thermodynamic anomalies only contribute a small proportion of the total precipitation variability. Finally, we show that the use of a large ensemble improves skill for predicting monsoon precipitation, but skill does not increase beyond 20 members.

1. Introduction

Global monsoon precipitation variability has substantial effects on about two thirds of the world's population (Wang and Ding 2006). Therefore, understanding the factors that drive monsoon variability, and its predictability, is societally important. However, simulating and predicting monsoon precipitation is still challenging, with many prediction systems exhibiting moderate-to-no skill on seasonal to multi-annual time scales (Bellucci *et al* 2013, Saha *et al* 2016).

Recently developed prediction systems have shown substantial skill at predicting tropical precipitation, for both seasonal to multi-annual time scales (Jia *et al* 2014, Dunstone *et al* 2020). Prediction models can predict aspects of Sahel precipitation from several months to several years ahead (Gaetani and Mohino 2013, Martin and Thorncroft 2014a, Rodrigues *et al* 2014, Siegmund *et al* 2015, Mohino *et al* 2016, Sheen *et al* 2017). On a seasonal time scale, skill has also been found over southern China (Lu *et al* 2017), India (Rajeevan *et al* 2012, Johnson *et al* 2017, Jain *et al* 2019, Chevuturi *et al* 2021), East and Austral Africa (Landman and Beraki 2012, Beraki *et al* 2014, Monerie *et al* 2018, Walker *et al* 2019), East Asia (Liu *et al* 2018), south America (Jones *et al* 2012), and Australia (King *et al* 2020). However, there are relatively few studies focusing on understanding the predictability of global monsoon precipitation (e.g. Saha *et al* 2016).

The predictability of tropical precipitation on seasonal time scales relies on the slowly varying lower boundary conditions (Charney and Shukla 1981), and thus is largely dependent on the ability of models to predict anomalous sea surface temperatures (SSTs) and their remote effects on monsoon precipitation. The El Niño Southern Oscillation (ENSO) is well known as a driver of tropical climate variability and is key to predicting monsoon precipitation (Shukla and Paolino 1983, Wang *et al* 2018, Sohn *et al* 2019, Dunstone *et al* 2020). In addition, anomalous North Atlantic and Indian Ocean SSTs also allow prediction of variations in monsoon precipitation (Mohino

et al 2016, Wang et al 2018). Therefore, our ability to predict monsoon precipitation could be dependent on SST conditions (e.g. ENSO phase) and on factors that can modulate SST—monsoon teleconnections (e.g. internal climate variability and external forcing) (Annamalai et al 2007, Chen et al 2010, O'Reilly et al 2017, Weisheimer et al 2017, Monerie et al 2018).

There has also been little work to understand the processes leading to skill in global monsoon predictions. For instance, precipitation variability is associated with both thermodynamic and dynamic mechanisms (Seager et al 2014). Thermodynamic changes are due to changes in surface temperature and specific humidity and, hence, may be highly predictable on a regional or even global scales. However, dynamic changes are associated with changes in the strength and pattern of the atmospheric circulation, which might lead to a reduced prediction skill. It is not clear which of these mechanisms is the dominant contributor to skill. We fill this gap by decomposing precipitation anomalies following Chadwick et al (2016) and quantifying skill at predicting each component.

Unpredictable noise acts to reduce our ability to predict monsoon precipitation. The large-ensemble approach aims to reduce unpredictable noise and hence elevate prediction skill by focusing on the predictable signal. However, the relative size of large ensemble required for predicting the global monsoon has not been assessed. DePreSys3 is a unique opportunity to analyse skill with a large ensemble, in which 40 members are available. In addition, DePreSys3 allows assessment of prediction skill over a long period (1959–2016) while previous prediction systems cover a shorter period (Scaife et al 2019).

This study aims at filling the aforementioned gaps, focusing on causes of skill at predicting global monsoon precipitation at a seasonal time scale. In addition, results help define the necessary ensemble size needed to predict global monsoon precipitation with a single climate model.

We address the following questions:

- Can we predict global monsoon precipitation, up to six months ahead?
- What are the sources of skill for monsoon precipitation in terms of dynamic and thermodynamic contributions?
- What ensemble size do we need to predict monsoon precipitation?

The paper is organised as follows: section 2 describes the simulations and the methodologies used. In section 3 we quantify skill at predicting monsoon precipitation. Sources of skill are shown in sections 4 and 5 provides an estimation of the ensemble sizes required to reduce unpredictable noise. Section 6 concludes.

2. Model and methods

2.1. The DePreSys3 prediction system

DePreSys3 is a Atmosphere-Ocean General Circulation Model developed at the Met Office (Dunstone et al 2016, 2018), and based on the Hadley Centre Global Environment Model version 3, global coupled configuration v2 (HadGEM3-GC2; Williams et al 2015). The ocean model is the Global Ocean version 5.0 (Megann et al 2014), based on the Nucleus for European Models of the Ocean Model (NEMO; Madec 2008). The ocean is run at a quarter degree resolution using the NEMO grid with 75 vertical levels (the ORCA025L75 grid; Bernard et al 2006). The sea-ice model is CICE version 4.1 (Hunke and Lipscomb 2004) from the United States Los Alamos National Laboratory. The atmosphere model is the Global Atmospheric version 6.0 (GA6) of the Met Office Unified Model, run at N216 resolution (∼60 km in the mid-latitudes) with 85 vertical levels. The land surface model is the Joint UK Land Environment Simulator version 6.0 (JULES; Best et al 2011). The different components are coupled using OASIS3 (Valcke 2013).

Two sets of hindcasts have been performed. The first set is performed by initialising simulations on 1 November each year between 1959 and 2016 (i.e. 58 start dates). In the second set, hindcasts are started from 1 May, each year between 1960 and 2016 (i.e. 57 start dates). Forty ensemble members are generated for both hindcast sets by using different seeds to a stochastic physics scheme (Dunstone et al 2016). Each hindcast is forced by the historical evolution of external forcings (greenhouse gases, aerosols, ozone, solar radiation and volcanoes). After 2005, external forcing is taken from the RCP4.5 scenario, as in the Climate Model Intercomparison Project (CMIP5) protocol (Taylor et al 2012). DePreSys3 is full-field initialized by relaxing a coupled integration of HadGEM3-GC2 towards gridded observations (see Dunstone et al 2016). Three-dimensional ocean temperature and salinity are relaxed toward the Met Office statistical ocean reanalysis (Smith and Murphy 2007, Smith et al 2015) and sea-ice concentration is relaxed towards HadISST (Hadley Centre Sea Ice and Sea Surface Temperature; Rayner et al 2003), both at one day relaxation time scale. The atmosphere model is initialised from ERA-40, before 1979, and from ERA-Interim afterwards with a six hourly relaxation time scale.

DePreSy3 is the Met Office decadal prediction system, however in this study we focus on the seasonal time scale (2–5 month forecast period). DePreSys3 is based upon the same physical climate model, and is at the same resolution, as the Met Office operational seasonal prediction system (GloSea5; MacLachlan et al 2015) but the longer DePreSys3 hindcasts period (1960–2016 vs 1992–2016 of GloSea5) allows for a

more robust evaluation of seasonal hindcast skill over the global monsoon regions.

2.2. Observations and reanalysis

Prediction skill is evaluated using observations and reanalysis. For precipitation we use the Global Precipitation Climatology Centre (GPCC) version v7 (Schneider *et al* 2014). GPCC is available over 1901–present on a 0.5° grid. We also use data from the Climate Research Unit (CRU; Harris *et al* 2014), version 4.03, which spans 1901-present. For a large range of atmospheric variables we used the data from the National Centers for Environmental Prediction (NCEP) reanalysis (R-1; Kanamitsu *et al* 2002). NCEP is given on a 2.5° resolution (144 × 72) with 17 vertical levels. NCEP spans 1948 to present.

2.3. Precipitation metrics

Observed and simulated precipitation are first interpolated onto a common 1° horizontal resolution grid when computing the monsoon precipitation indices. Precipitation is first interpolated to a common 2.5° resolution prior to assessing skill at each grid point. Monsoon domains are defined following Wang *et al* (2011), using GPCC. Specifically, grid points are selected where the annual observed precipitation range (i.e. the difference between May to September [MJJAS] and November to March [NDJFM]) exceeds 2.5 mm d^{-1}. Note, we only consider precipitation that falls within the tropical latitudes [30° S–30° N] and over land. Monsoon domains are shown in figure 1(a) and are named NAM (North America), NAF (North Africa), SAS (South Asia), EAS (East Asia), SAM (South America), SAF (South Africa), and AUS (Australia). Since we seek enough spread for a probabilistic forecast, we remove the first month of each simulation. We assess the ability of DePreSys3 at simulating precipitation over the northern hemisphere using hindcasts initialised in May, focusing on JJAS, i.e. on a 2–5 month forecast period. Over the southern hemisphere, the 2–5 month forecast period is defined using hindcasts initialised in November and focusing on DJFM.

We also assess global monsoon predictability, averaging together precipitation initialised in May for the northern hemisphere precipitation (focusing on JJAS) and initialised in November for the southern hemisphere precipitation (focusing on DJFM). This metric is hereafter called GM_nm.

2.4. Bias adjustment

Once initialised from reanalysis, models drift to their preferred (and imperfect) mean climatology. We remove the drift following the procedure described in the World Climate Research Program recommendation (ICPO 2011), as:

$$dr(\tau) = \frac{1}{nm}\sum_{j=1}^{n}\sum_{i=1}^{m} Y_j^i(\tau) - \frac{1}{n}\sum_{j=1}^{n} X_j(\tau)$$

where Y and X are given for a member i and a start date j for respectively DePreSys3 and the corresponding observations/reanalysis, spanning n start dates and m members. The drift, dr, is only lead-time (τ) dependent and is assumed to be start independent. Here, we assume that the ICPO method reliably removes drift for a large range of variables and over several regions. Note that the drift correction method does not impact our estimation of the model skill (anomaly correlation coefficient (ACC) values).

2.5. Evaluation of the model skill

We evaluate skill using the ACC between the ensemble-mean prediction from DePreSys3 and observations. The statistical significance of the ACC value is assessed by performing a Monte Carlo procedure through resampling (5000 permutations). For a given lead time, we randomly resampled time-series (of 57 and 58 years) using blocks of 5 year periods and filled until the size of the original time-series is reached, to preserve a multi-annual variability. Correlation between DePreSys3 and observed/reanalysed time-series are then computed for each permutation to form a distribution of ACC scores. ACC is then considered significant at $p \leqslant 0.05$ when values are greater than the 95th percentile of the permutation distribution (i.e. a one-sided test).

In this study the skill is always shown relative to the long-term trend, removing a linear trend for each-grid point and for each monsoon index. Note that removing the linear trend does not dramatically impact the results on prediction skill (not shown).

2.6. Decomposition

We decompose precipitation anomalies into terms documenting precipitation anomalies due to thermodynamic and dynamic changes. Held and Soden (2006) assumed that precipitation can be approximated by,

$$P = M * q$$

where, P is precipitation, M^* is a proxy for convective mass-flux from the boundary layer to the free troposphere (with $M^* = p/q$), and q is the near surface specific humidity.

A change in precipitation ($\Delta P = \Delta(M^* q)$) is computed for each month, regarding the 1959–2016 and 1960–2016 mean periods. Anomalies in precipitation are reformulated in terms of thermodynamic (ΔP_{therm}), dynamic (ΔP_{dyn}) and cross-nonlinear (ΔP_{cross}) components, following Chadwick *et al* (2013, 2016) as:

$$\Delta P = M * \Delta q + q\Delta M * + \Delta q \Delta M *$$

with,

$$\Delta P = \Delta P_{\text{therm}} + \Delta P_{\text{dyn}} + \Delta P_{\text{cross}}$$

where ΔP_{therm} is the anomaly in precipitation due to a change in specific humidity, with no change of the

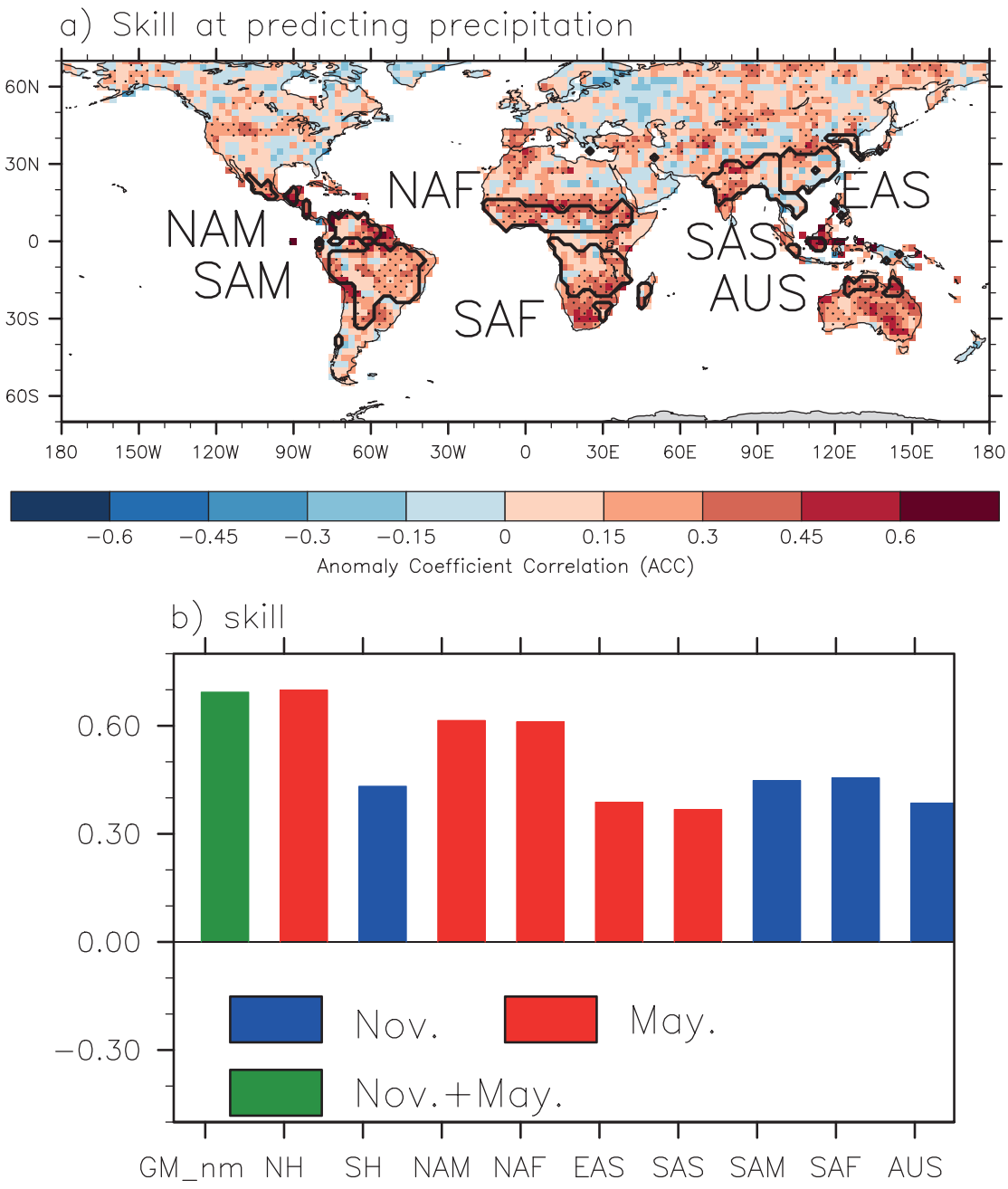

Figure 1. (a) Skill (anomaly coefficient correlation) at predicting summer precipitation in DePreSys3, at the 2–5 month forecast period. ACC values are shown for JJAS (DJFM) in the northern (southern) hemisphere, with the simulations initialised in May (November), over the period 1960–2016 (1959–2016). Stippling indicates that ACC is significantly different to zero according to a Monte-Carlo procedure with 5000 permutations and a 95% confidence level (see text for details). Black contours show the monsoon domains computed from GPCC. (b) Skill at predicting precipitation at the 2–5 month forecast period for different monsoon domains, the global monsoon, and the northern (NH) and southern (SH) hemispheres. Results for JJAS (DJFM) with the simulations initialised in May (November) are shown in red (blue). The combined GM definition (GM_nm) is shown in green. All bars are significant at the 95% level, method as in panel (a). ACC are computed with respect to GPCC and using ensemble-mean values for DePreSys3.

atmospheric circulation (constant M^*), ΔP_{dyn} is the anomaly in precipitation due to a change in the atmospheric dynamics, with no change in specific humidity value (q), and ΔP_{cross} is the anomaly in precipitation due to changes in both dynamics and specific humidity.

Further decomposition of ΔP_{dyn} allows to document changes that are due to the strength of the tropical mean circulation (ΔP_{weak}) and to a shift in the pattern of the circulation (ΔP_{shift}), as

$$\Delta P_{weak} = q\Delta M^*_{weak}$$

$$\Delta P_{shift} = q\Delta M^*_{shift}$$

with,

$$\Delta M^*_{weak} = -\alpha M*$$

where

$\alpha = -$(tropical mean ΔM^*/tropical mean M^*)

α is scaled by the strength of the mean tropical circulation. Finally,

$$\Delta M^*_{\text{shift}} = \Delta M* - \Delta M^*_{\text{weak}}.$$

The decomposition is performed at the monthly time step (as in Chadwick *et al* 2016, Rowell and Chadwick 2018) prior to computing the seasonal means and the area-weighted averages.

2.7. Quantification of the variability

We use a covariance analysis to quantify the part of the precipitation variance that is due to each term. Following Kent *et al* (2015), the precipitation variance can be written as the sum of the covariance matrix for all components:

$$\sigma_P^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \text{cov}\left(\Delta P_i, \Delta P_j\right).$$

Here, $n = 3$ when decomposing precipitation using ΔP_{therm}, ΔP_{dyn} and ΔP_{cross} and $n = 4$ when the dynamical term is decomposed further (i.e. using ΔP_{therm}, ΔP_{shift}, ΔP_{weak} and ΔP_{cross}). Subscripts i and j are different precipitation terms. Hereafter, $\text{cov}\left(\Delta P_i, \Delta P_j\right)$ denotes the covariance between two terms (ΔP_i and ΔP_j) and between itself (when $i = j$). $\text{cov}\left(\Delta P_i*\right)$ denotes the sum of the covariances between ΔP_i and all terms (including ΔP_i itself). It is worth nothing that the thermodynamic component is negatively correlated with ΔP_{weak} due to the fact that both ΔP_{therm} and ΔP_{weak} are associated with changes in tropical SST (Ma *et al* 2011, Kent *et al* 2015). Therefore, $\text{cov}\left(\Delta P_{\text{weak}}, \Delta P_{\text{therm}}\right)$ is negative. As a result, the variance explained by a term can exceed 100%. For instance, in DePreSys3, $\text{cov}\left(\Delta P_{\text{dyn}}*\right)$ is of 110% of the total precipitation variance for several monsoon domains.

3. Skill at predicting tropical precipitation

Figure 1 shows the precipitation skill at a 2–5 month forecast period, for each grid point and when averaged over each monsoon domain. Displayed time series (figure 2) show the ability of DePreSys3 to predict precipitation anomaly magnitude.

DePreSys3 exhibits substantial skill at predicting summer precipitation over the tropics (figure 1(a)). Significant ACC values are found over North and South America, the Sahel, southern Africa, northern India and Australia (figures 1(a) and (b), 2(b)–(j)). We have also evaluated ACC using CRU TS observation and find similar results (figure S1 (available online at stacks.iop.org/ERL/16/104035/mmedia)). In terms of monsoon domains, skill is strongest over the NAM and NAF (ACC = 0.61) domains, while skill is moderate, but still significant, over the EAS (ACC = 0.39) and SAS (ACC = 0.37) domains. Skill is also significant over the southern hemisphere with significant ACC values for the SAM (ACC = 0.45), SAF (ACC = 0.45) and AUS (ACC = 0.39) monsoon domains.

When assessed over the relatively long 1950–2016 period, the skill at predicting summer monsoon precipitation is due to DePreSys3's ability to simulate both interannual and multi-annual fluctuations in summer monsoon precipitation. Skill at predicting interannual and multi-annual variations in precipitation is shown for most of the monsoon domains (table 1; figures S2 and S3).

A substantial amount of skill in predicting NAF summer precipitation is due to the ability of DePreSys3 to simulate the multi-annual monsoon precipitation variability (ACC = 0.71; table 1; figure S2(e)), with the drying trend to the 1980s and the limited precipitation recovery of the 1990s (figures 2(e) and S2(e)). This seesaw in precipitation has been previously associated with Atlantic multidecadal variability (AMV) (Martin and Thorncroft 2014b). Therefore, we attribute a part of the skill in predicting NAF precipitation to be due to the high ACC values of the North Atlantic SSTs (figure S4), as shown in Mohino *et al* (2016). The skill at predicting NAM interannual variability is high (ACC = 0.71; table 1; figures 2(d) and S3) and is suggested to be due to the ability of DePreSys3 to simulate tropical Pacific SSTs (figure S4), which have strong effects on NAM precipitation (figures S5 and S6). The skill at predicting AUS precipitation mostly arises from the ability of DePreSys3 to simulate interannual precipitation variability (ACC = 0.61; table 1; figures 2(j) and S3) while skill is rather low for the multi-annual variability (ACC = 0.03; figure S2).

Skill at predicting global monsoon precipitation (GM_nm) is high (ACC = 0.68) and the model is also able to capture the large magnitude of anomalies in GM precipitation. We expect common sources of variability across the northern and the southern Hemisphere because of internal variability (e.g. AMV, ENSO, Interdecadal Pacific Oscillation) (Wang *et al* 2017, Monerie *et al* 2019) and external forcing (e.g. Ackerley *et al* 2011). This provides motivation for showing the skill for NH and SH monsoon precipitation. The skill is higher over the northern Hemisphere (in JJAS, ACC = 0.68) than over the southern Hemisphere (in DJFM, r = 0.44) (figures 1(b), 2(b) and (c)). This difference in skill mainly arises from DePreSys3's ability at simulating the multi-annual variability of the NH monsoon precipitation (ACC = 0.55; table 1: figure S2) compared to the SH (ACC = 0.24). In comparison, the interannual variability is well simulated by DePreSys3 for both NH (ACC = 0.71) and SH (ACC = 0.60) precipitation (table 1: figure S3).

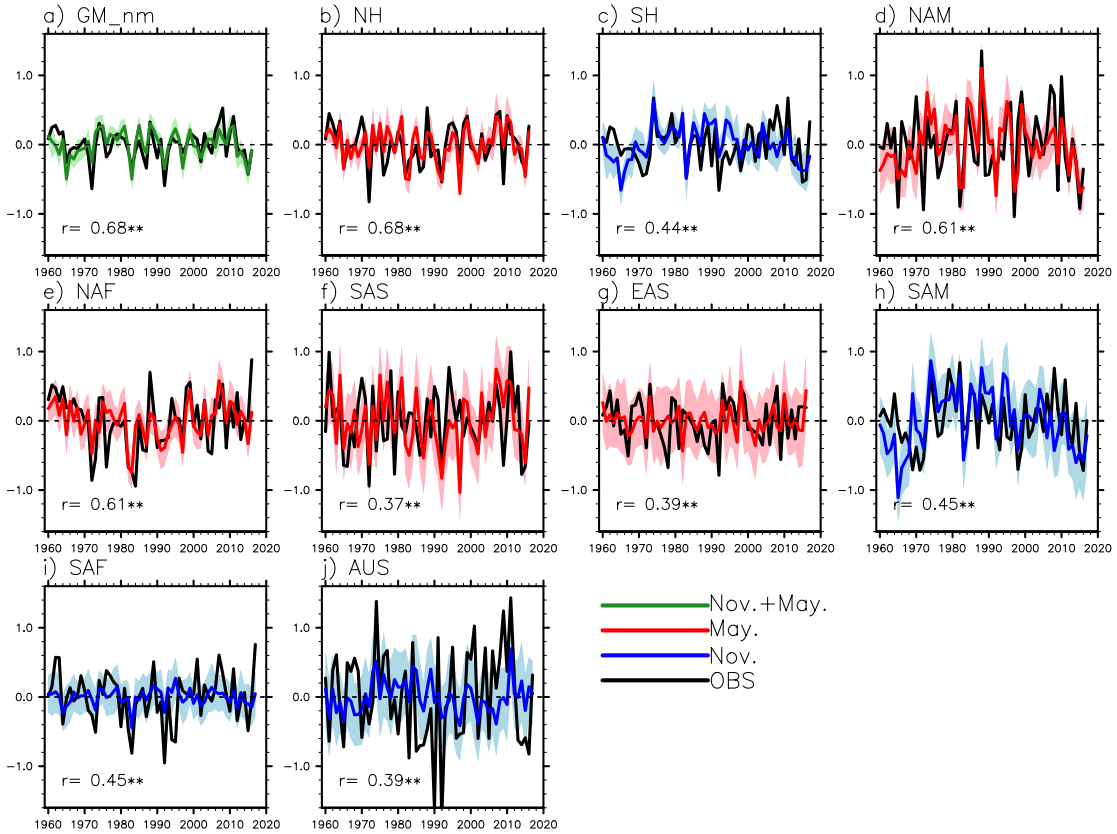

Figure 2. Linearly detrended summer monsoon precipitation time series (mm d^{-1}), for (a) the global monsoon, (b) the northern hemisphere, (c) the southern hemisphere and ((d)–(j)) the different monsoon domains. Time series are plotted for the JJAS (DJFM) season with the simulations initialised in May (November), with a red (blue) line. The GM monsoon domain is shown in green (initialised in May and November for northern and southern hemispheres, respectively). Observations (GPCC) are shown in black. Shading indicates the DePreSys3 ensemble spread defined by the ensemble standard deviation. The correlation between GPCC and the DePreSys3 ensemble-mean is added in each panel, two starts are added when the ACC is significantly different to zero according to a Monte-Carlo procedure with 5000 permutations and a 95% confidence level.

Table 1. Skill (Pearson's correlation between GPCC and DePreSys3) at predicting precipitation time series, for the total (interannual + multi-annual), and interannual and multi-annual variability separately in global and regional monsoon precipitation. The multi-annual evolution is extracted by performing a four year running mean. The interannual precipitation evolution is defined as the deviation of precipitation from the multi-annual component. (One) Two starts are added when the ACC is significantly different to zero according to a Monte-Carlo procedure with 5000 permutations and a (90) 95% confidence level.

Monsoon indices	GM_nm	NH	SH	NAM	NAF	SAS	EAS	SAM	SAF	AUS
Total variability	**0.68****	**0.68****	**0.44****	**0.61****	**0.61****	**0.37****	**0.39****	**0.45****	**0.45****	**0.39****
Interannual variability	**0.72****	**0.71****	**0.60****	**0.71****	**0.51****	**0.39****	**0.34****	**0.44****	**0.54****	**0.61****
Multi-annual variability	**0.51****	**0.55****	0.24	**0.37****	**0.71****	**0.27****	**0.50****	**0.43****	**0.30***	0.03

We expect skill at simulating monsoon precipitation variability to be associated with the ability of DePreSys3 to simulate ocean modes of variability. ENSO is an important source of skill at a seasonal lead time (Robertson *et al* 2015). We find that DePreSys3 has a high skill over the tropical Pacific Ocean at the 2–5 month forecast period (figure S4), which is consistent with the high skill in predicting interannual summer monsoon precipitation variability (table 1 and figure S3). In addition, low-frequency modes of variability (e.g. AMV and Interdecadal Pacific Oscillation (IPO)) and their respective impacts over land could also contribute to skill at predicting summer monsoon precipitation (table 1). High skill is found for indices of both AMV (ACC = 0.7; defined as the surface air temperature averaged over ocean points [0°–60° N; 80° W–0° E]), and IPO (ACC = 0.8; defined as the difference in surface air temperature, over ocean points between [25°–45° N; 170°–90° W] and [10° S–10° N; 150°–210° E], following Huang *et al* (2020)).

4. Unravelling sources of skill

4.1. Explaining summer monsoon precipitation variance

Section 3 did not indicate mechanisms responsible for the skill. Therefore, we further analyse the sources of skill by decomposing precipitation into terms representing the dynamic and thermodynamic

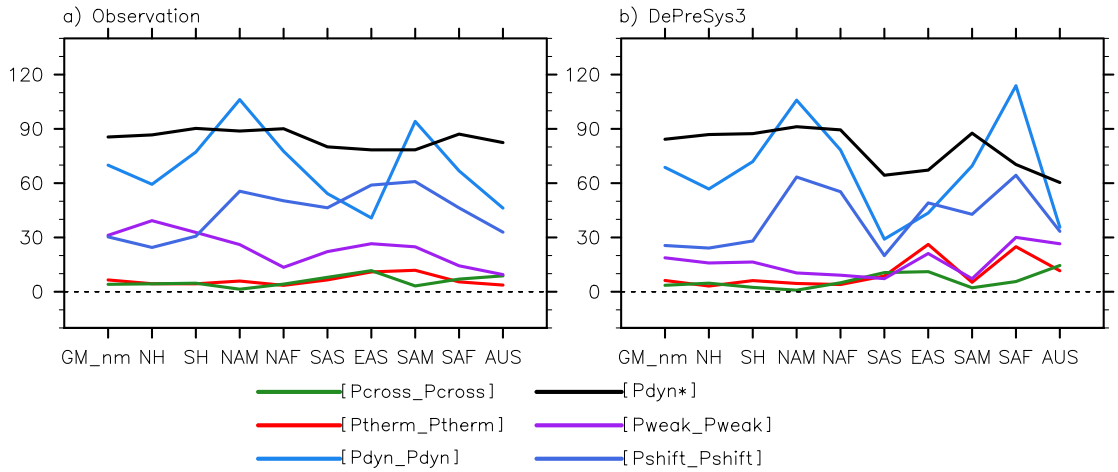

Figure 3. Covariance between several selected precipitation terms, in percent of the precipitation variance (%), for ΔP_{dyn} (blue), ΔP_{therm} (red) and ΔP_{cross} (green) and for ΔP_{weak} (purple) and ΔP_{shift} (dark blue). The sum of the covariances between each term and ΔP_{dyn} (cov $(\Delta P_{dyn}*)$) is shown with the black line. Results are given for each monsoon domain and for (a) observations/reanalysis and (b) DePreSys3.

contributions to the summer monsoon precipitation variability and by analysing the ability of DePreSys3 to capture the evolution of these precipitation terms. Tropical precipitation can be decomposed into different components, including dynamic (ΔP_{dyn}), thermodynamic (ΔP_{therm}) and cross non-linear (ΔP_{cross}) terms (see section 2.6). Thus, we use this decomposition to explore the reasons for skill. However, we first need to assess the relative importance of each term in explaining the total monsoon precipitation variability before assessing the skill of each precipitation component separately.

We find that precipitation variability is mostly associated with changes in the atmospheric dynamics (r ranging from 0.7 to 0.9 over the monsoon domains between precipitation and ΔP_{dyn}; not shown). However, the terms are not independent, and we document the importance of each term with a co-variance analysis (see section 2.7). The result is expressed as a percentage of the precipitation variance explained in order to stress the respective importance of each term to summer monsoon precipitation variability (figure 3). For clarity, we only show the covariances that explain most of the summer monsoon precipitation variance.

Figure 3(a) shows that, in reanalysis, ΔP_{dyn} is the dominant driver of the precipitation variance. ΔP_{dyn} explains most of the monsoon precipitation variance (cov($\Delta P_{dyn}*$) explains ~90% of precipitation variance for all monsoon domains). ΔP_{shift} is the main contributor to ΔP_{dyn}, and is the main source of the summer monsoon precipitation variance (figure 3(a)). The dominance of ΔP_{dyn} (and of ΔP_{shift}) shows that monsoon precipitation variability is mostly dominated by changes in atmospheric circulation. However, the relative importance of terms is monsoon domain dependent. The contribution of ΔP_{dyn} to summer monsoon precipitation variance is highest over NAM and SAM monsoon domains and is lowest over the AUS and EAS monsoon domains (figure 3(a); ~40% in reanalysis). In contrast, the ΔP_{therm} and ΔP_{cross} terms only account for a moderate part (less than 10%) of the summer monsoon precipitation variance (figure 3(a)).

DePreSys3 summer monsoon precipitation variance is also dominated by the dynamic components, while thermodynamic and cross components only moderately contribute to the precipitation variance (figure 3(b)). Like in observations, the covariance between ΔP_{dyn} and all terms (i.e. cov($\Delta P_{dyn}*$)) contributes more strongly to the precipitation variance over the NAM and SAF monsoon domains than over the SAS, EAS and AUS monsoon domains. Over the Australian and southern African monsoon domains, the contributions of ΔP_{weak} and ΔP_{therm} are relatively large.

4.2. Skilful prediction of the thermodynamic and dynamic terms

We now assess the predictability of the different precipitation components. Figure 4 shows the skill at predicting the different components of precipitation with DePreSys3, and figure 5 shows the skill for the terms averaged over the monsoon domains, both at a 2–5 month forecast period.

The skill at predicting the thermodynamic term (ΔP_{therm}) is high over the tropics, and particularly over the American monsoon domains, South Africa, Indonesia and East Asia (figure 4(a)). We also find that the skill at predicting ΔP_{therm} is statistically significant for all monsoon domains, and for GM_nm (figure 5(a)). The significant skill at predicting ΔP_{therm} is consistent with the ability of DePreSys3 to predict surface air temperature and specific humidity (figures S4 and S7). The precipitation variability is strongly associated with changes in atmospheric

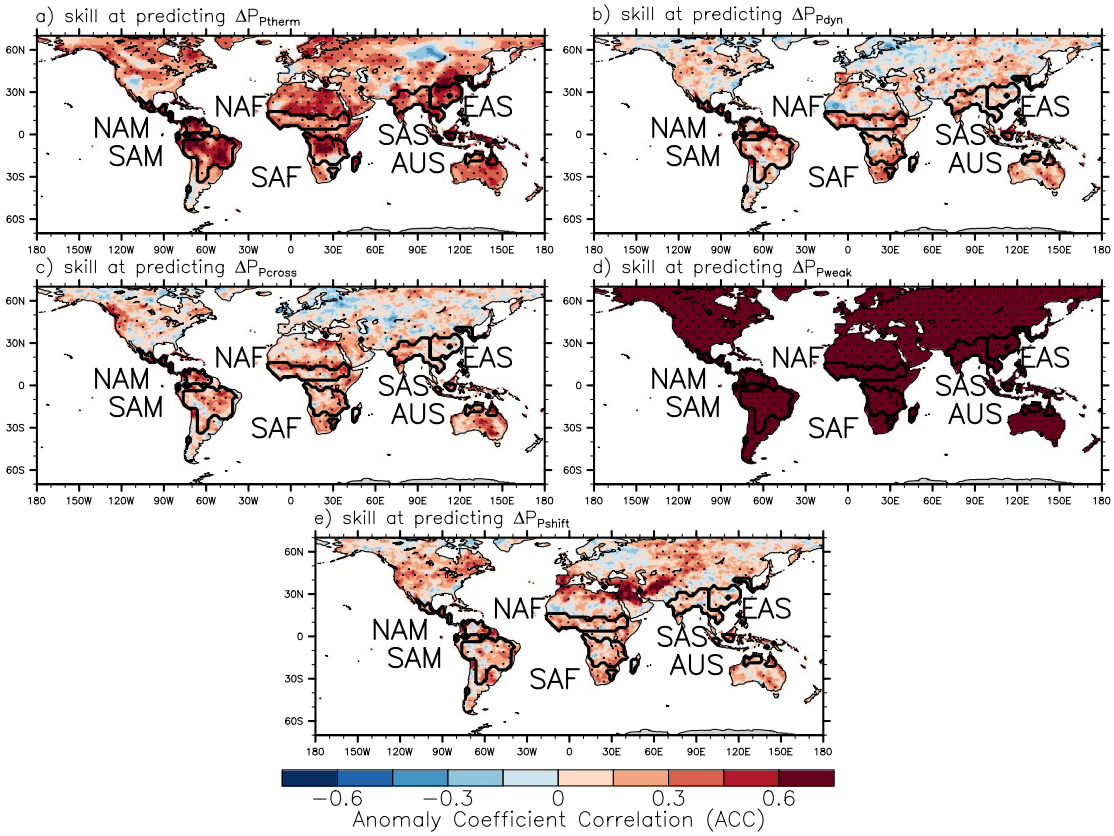

Figure 4. Skill (ACC) at predicting (a) ΔP_{therm}, (b) ΔP_{dyn}, (c) ΔP_{cross}, (d) ΔP_{weak} and (e) ΔP_{shift}. ACC values are shown for JJAS (DJFM) in the northern (southern) hemisphere, with the simulations initialised in May (November), over the period 1960–2016 (1959–2016). Stippling indicates that ACC is significantly different to zero according to a Monte-Carlo procedure with 5000 permutations and a 95% confidence level. ACC are computed with respect to observations/reanalysis.

circulation (figures 3(a) and (b)). However, skill in ΔP_{dyn} is rather low over the tropics (figure 4(b)), although it is significant when averaged over most of the monsoon domains (figure 5(b)). The lowest skill is evident at the grid-box scale for ΔP_{dyn}, but recovers to near the level of ΔP_{therm} for the area-average indices, and skill in ΔP_{dyn} is event larger than skill in ΔP_{therm} for NAF. We hypothesize that the low skill at the grid-box scale in ΔP_{dyn} may be due to noise in the verifying data.

The skill at predicting ΔP_{shift} and ΔP_{cross} is relatively low when assessed for each grid point (figures 4(c) and (e)). Nevertheless, there is significant skill in ΔP_{cross} when averaged over the monsoon domains (figure 5(c)). However, the skill at predicting precipitation associated with shifts of the circulation (ΔP_{shift}) is the lowest, which highlights deficiencies in predicting atmospheric circulation variability (figure 5(e)). The skill at predicting ΔP_{weak} is high and is the same for each grid point for a given hemisphere (figures 4(d) and 5(d)), because it is a tropical-mean quantity.

It is unclear whether the low skill is due to DePreSys3's inability to simulate large-scale or regional changes in atmospheric circulation. Wang *et al* (2018) proposed a NH monsoon circulation index, which is significantly positively correlated with global monsoon precipitation. This index is defined by computing changes in zonal wind shear, between the 850 westerlies and the 200 hPa easterlies, and averaged over a large area (between 0°–20° N and 120° W–90° E). In DePreSys3, the skill is high for the wind shear averaged over the tropics, in both JJAS (ACC = 0.7) and DJFM (ACC = 0.8) (significant at the 95% confidence level) (figure S9). Therefore, we conclude that DePreSys3 can predict the important large-scale atmospheric dynamics associated with the northern hemisphere summer monsoon. Hence, the relatively low ACC values of ΔP_{shift} are likely to be due to errors in simulating regional-scale atmospheric circulation.

The skill at predicting summer monsoon precipitation is not solely due to ability of DePreSys3 to predict ΔP_{dyn}. This suggests that, even if accounting for a relatively small proportion of the explained precipitation variance, ΔP_{therm} and ΔP_{cross} could be helpful for predicting precipitation. For instance, DePreSys3 has skill at predicting EAS and AUS precipitation (figure 1(b)), that cannot be attributed to a prediction of ΔP_{dyn} (figure 5(b)) or ΔP_{shift} (figure 5(e)). However, we show in figures 3(a) and (b) that the contribution of ΔP_{dyn} and ΔP_{shift} is anomalously low for the EAS and AUS precipitation variability, compared to the other monsoon domains, and the

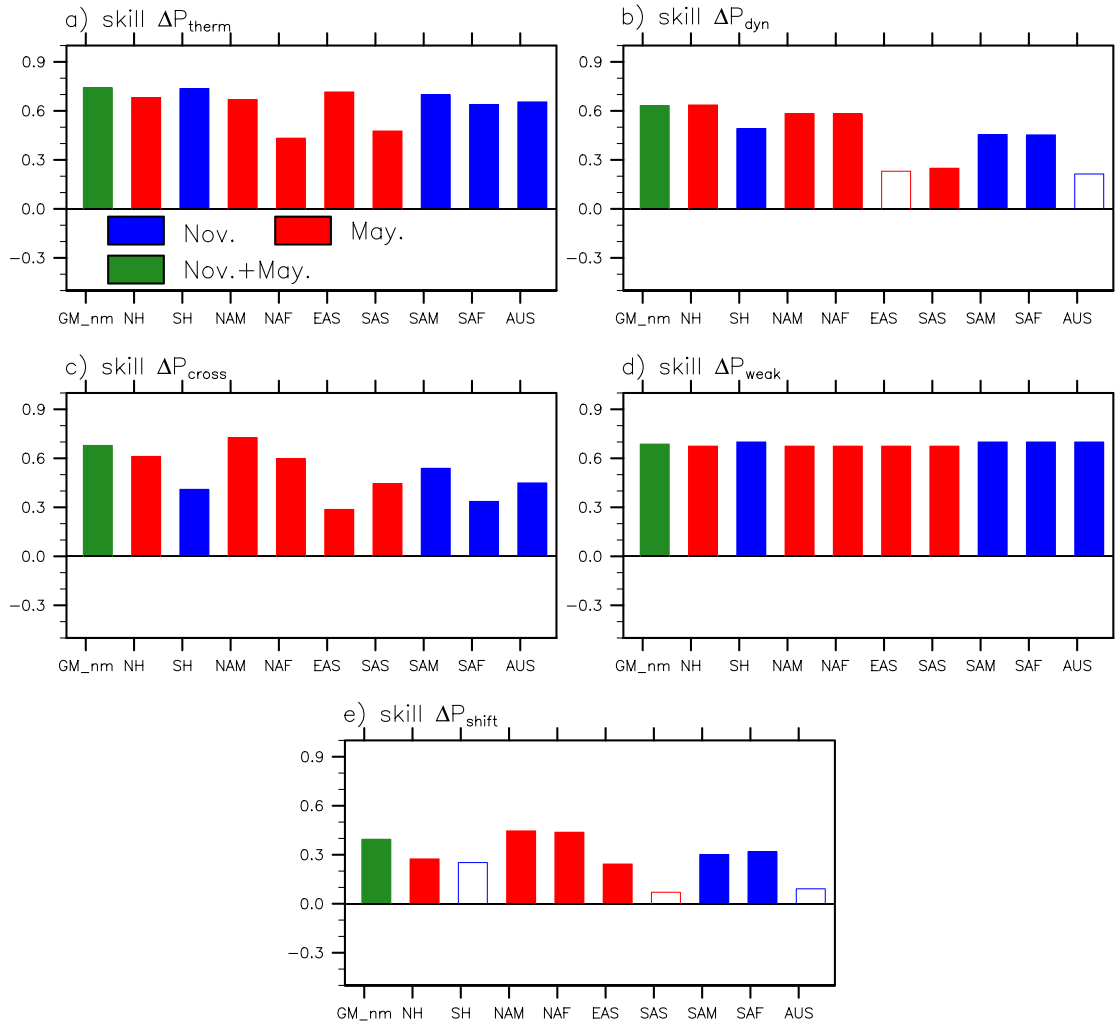

Figure 5. Skill at predicting the weighted-area averaged (a) ΔP_{therm}, (b) ΔP_{dyn}, (c) ΔP_{cross}, (d) ΔP_{weak} and (e) ΔP_{shift} terms over each monsoon domain, the global monsoon and the northern (NH) and southern (SH) hemispheres, for 2–5 month forecast periods. Results for JJAS (DJFM) with the simulations initialised in May (November) are shown in red (blue). The global monsoon is shown in green. Filled bars indicate that the ACC is significantly different to zero according to a Monte-Carlo procedure with 5000 permutations and a 95% confidence level. ACC are computed with respect to observations/reanalysis.

relative importance of the other pairs of covariances could be of importance, through for instance feedbacks between precipitation drivers.

The relatively low skill at predicting ΔP_{shift} shows that simulating changes in atmospheric circulation is challenging. The low skill at predicting ΔP_{shift} could be due to the inability of DePreSys3 to simulate remote effects of SSTs on precipitation. For instance, we assess the ability of DePreSys3 to simulate effects of SSTs on shifts in circulations on SAS and AUS monsoon precipitation (figure S8) and we note strong differences between observation and DePreSys3. Therefore, we suggest the low skill in SAS and AUS monsoon precipitation to be partly due to errors in simulating teleconnections between the tropical Pacific SSTs and the monsoons.

5. Ensemble size

We assume that a substantial proportion of summer monsoon precipitation variability might be unpredictable. Therefore, low skill at predicting summer monsoon precipitation and ΔP_{shift} might be due to unpredictable variability rather than model errors. However, it is well known that increasing the ensemble size will reduce stochastic and unpredictable noise and, hence, increase skill, as in Scaife and Smith (2018) for predicting the North Atlantic Oscillation. Over the tropics the minimum number of members needed to extract substantial skill at predicting summer monsoon precipitation on seasonal time scales has not been assessed so far, in a unique explicit assessment for all monsoons and using a large ensemble. We fill this gap using the large DePreSys3 ensemble size. To test this, we have resampled the hindcast dataset to create new synthetic timeseries, randomly selecting m ensemble members for each start date. The ensemble-mean of each ensemble of m-members is performed before to compute the ACC relative to the observed timeseries. m vary from 1 to 40 and 50 000 permutations are used. We have defined the result as significant when at least 95%

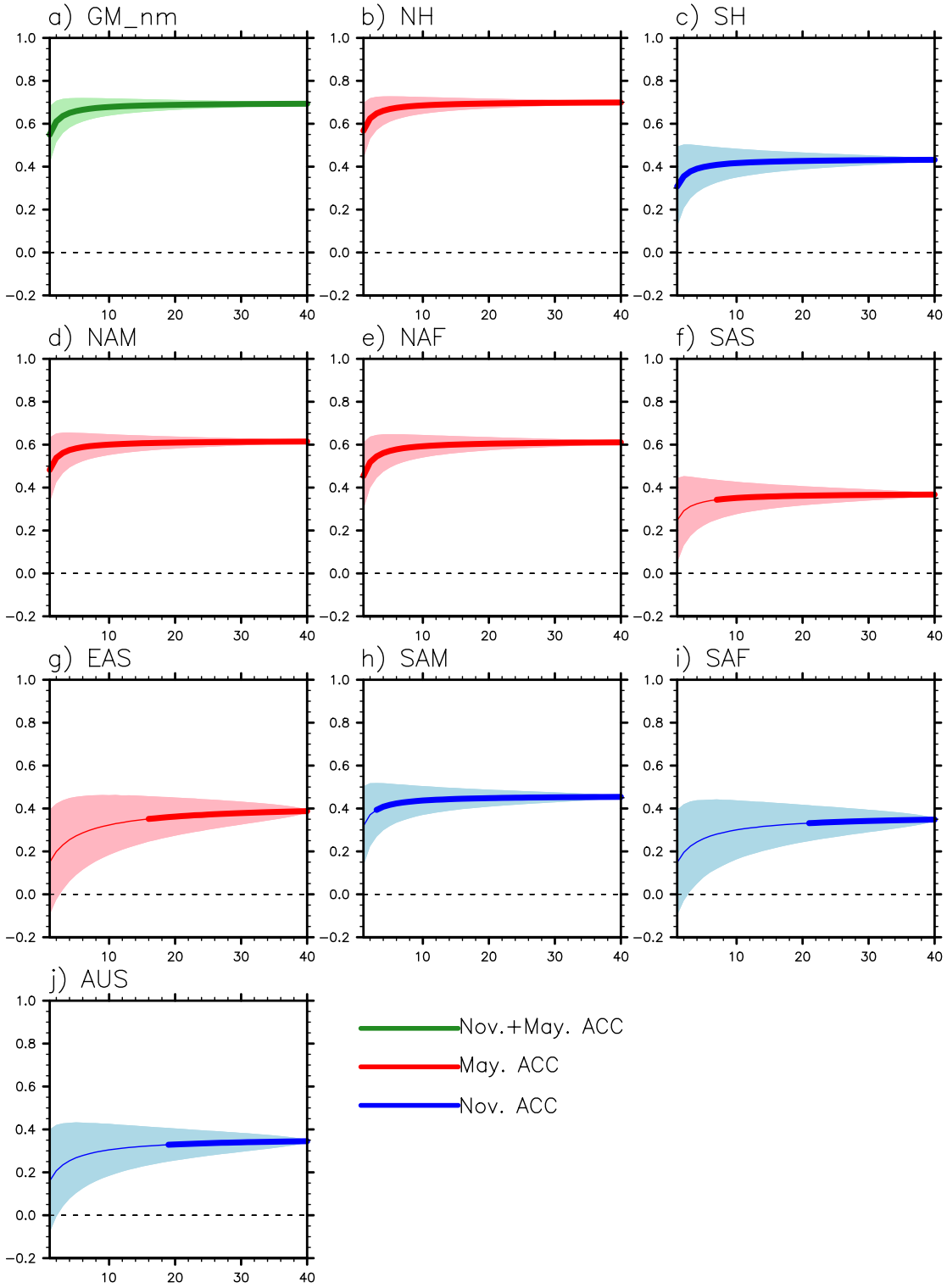

Figure 6. Skill at predicting summer monsoon precipitation, depending on the number of ensemble members (solid lines) (ranging from 1 to 40 members). The new ensemble members are computed by randomly re-sampling the data set and compared to observations (solid lines). A total of 50 000 permutations have been performed to re sample the data. Bold lines indicate that at least 95% of the 50 000 ensemble members show significant correlations (as defined here with a Student's *t*-test at the 95% confidence level). We show the median of the 50 000 correlations. ACC are computed with respect to GPCC.

of the 50 000 ensemble-means produce significant skill (as defined here with a Student's *t*-test at the 95% confidence level) at predicting summer monsoon precipitation.

Figure 6 indicates that, as expected, skill increases with *m*, a larger ensemble allowing for better skill. For predicting global monsoon and northern and southern Hemisphere summer monsoon precipitation we only need a limited ensemble size and skill converges with ensembles of ∼6–10 members (figures 6(a)–(d)). A limited number of members are also needed for most of the monsoon domains, as

seen for the NAM, NAF and SAM monsoon domains (figures 6(e), (f) and (i)). However, an ensemble of 20 members or more, could be needed for predicting EAS, SAS and AUS summer monsoon precipitation (figures 6(h), (j) and (k)). This raises the importance of using relatively large ensembles over some monsoon domains.

6. Discussion and conclusions

We find that DePreSys3 (Dunstone *et al* 2016) has significant skill at predicting summer monsoon precipitation on a 2–5 month forecast period, using predictions initialised annually over 1959–2016. However, the skill depends on the specific monsoon domains being considered. The highest skill is found for the NAM and NAF monsoon domains (ACC = 0.61), while the lowest skill is found for the SAS monsoon domain (ACC = 0.37). Skill at predicting interannual monsoon variability is high (ACC = 0.72) for global monsoon precipitation and is associated with the high skill of the tropical Pacific SSTs (figure S4). However, there is significant skill on multi-annual time scales too, but the skill is monsoon domain dependent; low over the Australian monsoon domain (ACC = 0.03) and high over the NAF monsoon domain (ACC = 0.71). Although skill at predicting monsoon precipitation was shown to be strongly model dependent (Rodrigues *et al* 2014), DePreSys3 has comparable skill to other prediction systems for global monsoon precipitation (Saha *et al* 2016) and the individual monsoon domains. However, results of this study are not directly comparable to other prediction systems because of differences in ensemble size and length of the hindcast period covered.

We have assessed whether predictability in monsoon rainfall arises from thermodynamically (ΔP_{therm}) and dynamically (ΔP_{dyn}) driven components by using the decomposition method of Chadwick *et al* (2016). Significant skill is obtained for both ΔP_{therm} and ΔP_{dyn}, although ΔP_{dyn} generally contributes less skill than ΔP_{therm}. Overall, we show that the interannual variability of monsoon precipitation is primarily due to shifts of the atmospheric circulation (ΔP_{shift}), which is not well captured in predictions. Hence, it is critical to improve predictions of ΔP_{shift} to improve the skill of monsoon precipitation prediction. However, we acknowledge that prediction skill could depend on the observations/reanalysis that are used for verification, and that the prediction skill of ΔP_{shift} is expected to be more uncertain that the prediction skill of ΔP_{therm}.

We find that deficiencies in the predictions of monsoon precipitation are largely explained by DePreSys3's inability to simulate ΔP_{shift}. More specifically, we show that skill is low at capturing ΔP_{shift} over the Australian and South Asian monsoon domains, because of errors in the simulated teleconnections between Pacific and Indian SSTs and land monsoon precipitation. However, ΔP_{shift} only explains a moderate part of the precipitation variance of AUS and SAS precipitation, suggesting that the improvement of the skill in these regions could be limited even if predictions of ΔP_{shift} were improved. Therefore, additional efforts should be devoted to understanding feedbacks between the different precipitation terms and their biases in DePreSys3, especially over South Asia, Australia and Indonesia.

Another way to increase skill at predicting precipitation is to reduce the unpredictable noise by increasing the ensemble size of the predictions. For most of the monsoon domains, and for global monsoon precipitation, we show that an ensemble of 5–10 members is necessary to extract significant skill for prediction. However, at least 20 members are necessary to get useful predictions of precipitation over southern Africa and over the Australian monsoon domain. Increasing the number of members beyond this offers no significant increase in skill for monsoon forecasts on this time scale.

In Summary, our results suggest that improving our ability to simulate shifts of the circulation would lead to an important improvement of the predictive skill of summer monsoon precipitation. This improvement could result from detailed assessment of the role of systematic biases in the mean-state simulation of SST and monsoon circulation on ENSO—monsoon teleconnections (e.g. Turner *et al* 2005) and prediction skill (e.g. Lee *et al* 2010). Nevertheless, we expect model error and prediction skill to be model-dependent and, hence, analysing skill of several prediction systems will be important to further define the sources of skill. In addition, a multi-model combination could further improve skill because of differences in structural models biases (e.g. Dunstone *et al* 2020).

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The contribution of P A M to this work was funded by the EMERGENCE project under the Natural Environment Research Council (NERC Grant NE/S004890/1). J R was supported by the Natural Environment Research Council (NERC) via the National Centre for Atmospheric Science (NCAS), and J R was additionally funded by the NERC ACSIS program. Part of AGT's contribution to this work was conducted through the Weather and Climate Science for Service Partnership (WCSSP) India, a collaborative initiative between the Met Office, supported by the UK Governments Newton Fund, and the Indian Ministry of Earth Sciences (MoES). N D was supported by the UK-China Research & Innovation Partnership Fund through the Met Office Climate Science

for Service Partnership (CSSP) China as part of the Newton Fund; P A M and J R were also part-funded via the partnership via the DOVE project. We thank Doug Smith, Leon Hermanson and the Met Office for providing the DePreSys3 output. We also thank the two anonymous reviewers for their comments and suggestions.

ORCID iDs

Paul-Arthur Monerie https://orcid.org/0000-0002-5304-9559
Jon I Robson https://orcid.org/0000-0002-3467-018X
Nick J Dunstone https://orcid.org/0000-0001-6859-6814
Andrew G Turner https://orcid.org/0000-0002-0642-6876

References

Ackerley D, Booth B B B, Knight S H E, Highwood E J, Frame D J, Allen M R and Rowell D P 2011 Sensitivity of twentieth-century Sahel rainfall to sulfate aerosol and CO_2 forcing *J. Clim.* **24** 4999–5014

Annamalai H, Hamilton K and Sperber K R 2007 The South Asian summer monsoon and its relationship with ENSO in the IPCC AR4 simulations *J. Clim.* **20** 1071–92

Bellucci A, Gualdi S, Masina S, Storto A, Scoccimarro E, Cagnazzo C, Fogli P, Manzini E and Navarra A 2013 Decadal climate predictions with a coupled OAGCM initialized with oceanic reanalyses *Clim. Dyn.* **40** 1483–97

Beraki A F, DeWitt D G, Landman W A and Olivier C 2014 Dynamical Seasonal Climate Prediction Using an Ocean–Atmosphere Coupled Climate Model Developed in Partnership between South Africa and the IRI *J. Clim.* **27** 1719–41

Bernard B *et al* 2006 Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution *Ocean Dyn.* **56** 543–67

Best M J *et al* 2011 The Joint UK Land Environment Simulator (JULES), model description—part 1: energy and water fluxes *Geosci. Model Dev.* **4** 677–99

Chadwick R, Boutle I and Martin G 2013 Spatial patterns of precipitation change in CMIP5: why the rich do not get richer in the tropics *J. Clim.* **26** 3803–22

Chadwick R, Good P and Willett K 2016 A simple moisture advection model of specific humidity change over land in response to SST warming *J. Clim.* **29** 7613–32

Charney J G and Shukla J 1981 *Predictability of Monsoons. Monsoon Dynamics* ed J Lighthill and R P Pearce (Cambridge: Cambridge University Press) (https://doi.org/10.1017/CBO9780511897580)

Chen W, Dong B and Lu R 2010 Impact of the Atlantic Ocean on the multidecadal fluctuation of El Niño–Southern Oscillation–South Asian monsoon relationship in a coupled general circulation model *J. Geophys. Res. Atmos.* **115** D17109

Chevuturi A, Turner A G, Johnson S, Weisheimer A, Shonk J K P, Stockdale T N and Senan R 2021 Forecast skill of the Indian monsoon and its onset in the ECMWF seasonal forecasting system 5 (SEAS5) *Clim. Dyn.* **56** 2941–57

Dunstone N *et al* 2018 Skilful seasonal predictions of summer European rainfall *Geophys. Res. Lett.* **45** 3246–54

Dunstone N *et al* 2020 Skilful interannual climate prediction from two large initialised model ensembles *Environ. Res. Lett.* **15** 094083

Dunstone N, Smith D, Scaife A, Hermanson L, Eade R, Robinson N, Andrews M and Knight J 2016 Skilful predictions of the winter North Atlantic Oscillation one year ahead *Nat. Geosci.* **9** 809–14

Gaetani M and Mohino E 2013 Decadal prediction of the Sahelian precipitation in CMIP5 simulations *J. Clim.* **26** 7708–19

Harris I, Jones P D, Osborn T J and Lister D H 2014 Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset *Int. J. Climatol.* **34** 623–42

Held I M and Soden B J 2006 Robust responses of the hydrological cycle to global warming *J. Clim.* **19** 5686–99

Huang X *et al* 2020 The recent decline and recovery of Indian summer monsoon rainfall: relative roles of external forcing and internal variability *J. Clim.* **33** 5035–60

Hunke E C and Lipscomb W H 2004 The Los Alamos sea ice model, documentation and software Version 3.1 (Los Alamos, NM)

ICPO 2011 Data and bias correction for decadal climate predictions. International CLIVAR Project Office Publication Series 150:5. (available online at http://www.wcrp-climate.org/decadal/references/DCPP_Bias_Correction.pdf)

Jain S, Scaife A A and Mitra A K 2019 Skill of Indian summer monsoon rainfall prediction in multiple seasonal prediction systems *Clim. Dyn.* **52** 5291–301

Jia L *et al* 2014 Improved seasonal prediction of temperature and precipitation over land in a high-resolution GFDL climate model *J. Clim.* **28** 2044–62

Johnson S J, Turner A, Woolnough S, Martin G and MacLachlan C 2017 An assessment of Indian monsoon seasonal forecasts and mechanisms underlying monsoon interannual variability in the Met Office GloSea5-GC2 system *Clim. Dyn.* **48** 1447–65

Jones C, Carvalho L M and Liebmann B 2012 Forecast Skill of the South American Monsoon System *J. Clim.* **25** 1883–89

Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo J J, Fiorino M and Potter G L 2002 NCEP–DOE AMIP-II reanalysis (R-2) *Bull. Am. Meteorol. Soc.* **83** 1631–43

Kent C, Chadwick R and Rowell D P 2015 Understanding uncertainties in future projections of seasonal tropical precipitation *J. Clim.* **28** 4390–413

King A D, Hudson D, Lim E-P, Marshall A G, Hendon H H, Lane T P and Alves O 2020 Sub-seasonal to seasonal prediction of rainfall extremes in Australia *Q. J. R. Meteorol. Soc.* **146** 2228–49

Landman W A and Beraki A 2012 Multi-model forecast skill for mid-summer rainfall over southern Africa *Int. J. Climatol.* **32** 303–14

Lee J-Y *et al* 2010 How are seasonal prediction skills related to models' performance on mean state and annual cycle? *Clim. Dyn.* **35** 267–83

Liu Y, Ren H-L, Scaife A A and Li C 2018 Evaluation and statistical downscaling of East Asian summer monsoon forecasting in BCC and MOHC seasonal prediction systems *Q. J. R. Meteorol. Soc.* **144** 2798–811

Lu B, Scaife A A, Dunstone N, Smith D, Ren H-L, Liu Y and Eade R 2017 Skillful seasonal predictions of winter precipitation over southern China *Environ. Res. Lett.* **12** 74021

Ma J, Xie S-P and Kosaka Y 2011 Mechanisms for tropical tropospheric circulation change in response to global warming *J. Clim.* **25** 2979–94

MacLachlan C *et al* 2015 Global seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system *Q. J. R. Meteorol. Soc.* **141** 1072–84

Madec G 2008 NEMO ocean engine

Martin E R and Thorncroft C D 2014b The impact of the AMO on the West African monsoon annual cycle *Q. J. R. Meteorol. Soc.* **140** 31–46

Martin E R and Thorncroft C 2014a Sahel rainfall in multimodel CMIP5 decadal hindcasts *Geophys. Res. Lett.* **41** 2169–75

Megann A, Storkey D, Aksenov Y, Alderson S, Calvert D, Graham T, Hyder P, Siddorn J and Sinha B 2014 GO5.0: the joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications *Geosci. Model Dev.* **7** 1069–92

Mohino E, Keenlyside N and Pohlmann H 2016 Decadal prediction of Sahel rainfall: where does the skill (or lack thereof) come from? *Clim. Dyn.* **47** 3593–612

Monerie P-A *et al* 2018 Predicting the seasonal evolution of southern African summer precipitation in the DePreSys3 prediction system *Clim. Dyn.* **52** 6491–510

Monerie P-A, Robson J, Dong B, Hodson D L R and Klingaman N P 2019 Effect of the Atlantic multidecadal variability on the global monsoon *Geophys. Res. Lett.* **46** 1765–75

O'Reilly C H, Heatley J, MacLeod D, Weisheimer A, Palmer T N, Schaller N and Woollings T 2017 Variability in seasonal forecast skill of Northern Hemisphere winters over the twentieth century *Geophys. Res. Lett.* **44** 5729–38

Rajeevan M, Unnikrishnan C K and Preethi B 2012 Evaluation of the ENSEMBLES multi-model seasonal forecasts of Indian summer monsoon variability *Clim. Dyn.* **38** 2257–74

Rayner N A *et al* 2003 Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century *J. Geophys. Res.* **108** 4407

Robertson A W, Kumar A, Peña M and Vitart F 2015 Improving and promoting subseasonal to seasonal prediction *Bull. Am. Meteorol. Soc.* **96** ES49–ES53

Rodrigues L R L, García-Serrano J and Doblas-Reyes F 2014 Seasonal forecast quality of the West African monsoon rainfall regimes by multiple forecast systems *J. Geophys. Res. Atmos.* **119** 7908–30

Rowell D P and Chadwick R 2018 Causes of the uncertainty in projections of tropical terrestrial rainfall change: East Africa *J. Clim.* **31** 5977–95

Saha S K, Sujith K, Pokhrel S, Chaudhari H S and Hazra A 2016 Predictability of global monsoon rainfall in NCEP CFSv2 *Clim. Dyn.* **47** 1693–715

Scaife A A *et al* 2019 Tropical rainfall predictions from multiple seasonal forecast systems *Int. J. Climatol.* **39** 974–88

Scaife A A and Smith D 2018 A signal-to-noise paradox in climate science *npj Clim. Atmos. Sci.* **1** 28

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M and Rudolf B 2014 GPCC's new land surface precipitation climatology based on quality-controlled *in situ* data and its role in quantifying the global water cycle *Theor. Appl. Climatol.* **115** 15–40

Seager R, Liu H, Henderson N, Simpson I, Kelley C, Shaw T, Kushnir Y and Ting M 2014 Causes of increasing aridification of the mediterranean region in response to rising greenhouse gases *J. Clim.* **27** 4655–76

Sheen K L *et al* 2017 Skilful prediction of Sahel summer rainfall on inter-annual and multi-year timescales **8** 14966

Shukla J and Paolino D A 1983 The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India *Mon. Weather Rev.* **111** 1830–7

Siegmund J, Bliefernicht J, Laux P and Kunstmann H 2015 Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high-resolution dynamical downscaling *J. Geophys. Res. Atmos.* **120** 7316–39

Smith D M, Allan R P, Coward A C, Eade R, Hyder P, Liu C, Loeb N G, Palmer M D, Roberts C D and Scaife A A 2015 Earth's energy imbalance since 1960 in observations and CMIP5 models *Geophys. Res. Lett.* **42** 1205–13

Smith D M and Murphy J M 2007 An objective ocean temperature and salinity analysis using covariances from a global climate model *J. Geophys. Res.* **112** C02022

Sohn S-J, Tam C-Y and Kug J-S 2019 How does ENSO diversity limit the skill of tropical Pacific precipitation forecasts in dynamical seasonal predictions? *Clim. Dyn.* **53** 5815–31

Taylor K E, Stouffer R J and Meehl G A 2012 An overview of CMIP5 and the experiment design *Bull. Am. Meteorol. Soc.* **93** 485–98

Turner A G, Inness P M and Slingo J M 2005 The role of the basic state in the ENSO–monsoon relationship and implications for predictability *Q. J. R. Meteorol. Soc.* **131** 781–804

Valcke S 2013 The OASIS3 coupler: a European climate modelling community software *Geosci. Model Dev.* **6** 373–88

Walker D P, Birch C E, Marsham J H, Scaife A A, Graham R J and Segele Z T 2019 Skill of dynamical and GHACOF consensus seasonal forecasts of East African rainfall *Clim. Dyn.* **53** 4911–35

Wang B and Ding Q 2006 Changes in global monsoon precipitation over the past 56 years *Geophys. Res. Lett.* **33** L06711

Wang B, Kim H-J, Kikuchi K and Kitoh A 2011 Diagnostic metrics for evaluation of annual and diurnal cycles *Clim. Dyn.* **37** 941–55

Wang B, Li J, Cane M A, Liu J, Webster P J, Xiang B, Kim H-M, Cao J and Ha K-J 2018 Toward predicting changes in the land monsoon rainfall a decade in advance *J. Clim.* **31** 2699–714

Wang P X *et al* 2017 The global monsoon across time scales: mechanisms and outstanding issues *Earth-Sci. Rev.* **174** 84–121

Weisheimer A, Schaller N, O'Reilly C, MacLeod D A and Palmer T 2017 Atmospheric seasonal forecasts of the twentieth century: multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution *Q. J. R. Meteorol. Soc.* **143** 917–26

Williams K D *et al* 2015 The met office global coupled model 2.0 (GC2) configuration *Geosci. Model Dev.* **8** 1509–24