South Asian monsoon precipitation in CMIP5: a link between intermodel spread and the representations of tropical convection

Samson Hagos¹ · L. Ruby Leung¹ · Moetasim Ashfaq² · Karthik Balaguru¹

Received: 11 September 2017 / Accepted: 15 March 2018 / Published online: 20 March 2018 © The Author(s) 2018

Abstract

CMIP5 models exhibit a mean dry bias and a large inter-model spread in simulating South Asian monsoon precipitation but the origins of the bias and spread are not well understood. Using moisture and energy budget analysis that exploits the weak temperature gradients in the tropics, we derived a non-linear relationship between the normalized precipitation and normalized precipitable water that is similar to the non-linear relationship between precipitation and precipitable water found in previous observational studies. About half of the 21 models analyzed fall in the steep gradient of the non-linear relationship where small differences in the normalized precipitable water in the equatorial Indian Ocean (EIO) manifest in large differences in normalized precipitation in the region. Models with larger normalized precipitable water in the EIO during spring contribute disproportionately to the large inter-model spread and multi-model mean dry bias in monsoon precipitation through perturbations of the large-scale winds. Thus the intermodel spread in precipitable water over EIO leads to the dry bias in the multi-model mean South Asian monsoon precipitation. The models with high normalized precipitable water over EIO also project larger response to warming and dominate the inter-model spread in the multi-model projections of monsoon rainfall. Conversely, models on the flat side of the relationship between normalized precipitation and precipitable water are in better agreement with each other and with observations. On average these models project a smaller increase in the projected monsoon precipitation than that from multi-model mean. This study identified the normalized precipitable water over EIO, which is determined by the relationship between the profiles of convergence and moisture and therefore is an essential outcome of the treatment of convection, as a key metric for understanding model biases and differentiating model skill in simulating South Asian monsoon precipitation.

1 Introduction

The South Asian monsoon is a prominent large-scale circulation feature. It influences a significant fraction of the world population that depends on the monsoon rainfall for food, energy production and many other economic activities. Understanding the physical processes that control the monsoon and its response to natural and anthropogenic forcings is of high societal and scientific value. Because of the complex processes involved, representing the monsoon and projecting its future changes has been a major challenge in climate modeling. Many global climate models in the

Samson Hagos samson.hagos@pnnl.gov

² Oak Ridge National Laboratory, Oak Ridge, TN, USA

Coupled Model Intercomparison Project (CMIP) display a dry bias in their simulation of the present day South Asian monsoon precipitation. While the multi-model mean of the latest set of models in CMIP5 shows some improvements in the spatial distribution of precipitation over that of CMIP3, the inter-model spread remains large and the delayed as well as weak monsoon rainfall persisted in many models (Sperber et al. 2013).

The 10-year mean (1996–2005) seasonal cycle of precipitation in the models and the projected summer monsoon rainfall under the RCP8.5 scenario are displayed in Fig. 1. Figure 1a shows the seasonal cycle of the 10-year mean all India precipitation (AIP, defined as the average precipitation over the land region from 5° N to 30° N and 70° E to 90° E) from the historical simulations of 21 CMIP5 models and the observed precipitation from TRMM-3B42 (Huffman et al. 2007), GPCP (Huffman et al. 2001) and rain gauge data from the India Meteorological Division (IMD). The list of CMIP5 models and the relative strength of their monsoon



¹ Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA 99352, USA



Fig. 1 a The seasonal cycle of present day (1996–2005) average All India Precipitation (mm/day). The red and blue lines correspond to the CMIP5 historical model simulations. The averaging area is the land portion of the red box in Fig. 2. The models are defined as strong/weak monsoon models depending on whether their JJAS mean precipitation exceeds the averaged of the three observations (black lines). **b** Future precipitation (2091–2100 JJAS mean) in RCP 8.5 scenario versus the present day precipitation (1996–2005). Each red/ blue dot corresponds to a CMIP5 simulation with weak/strong monsoon. The black line marks the multi-model mean increase of 20%

precipitation compared to the observations are shown in Table 1. Besides the considerably drier multi-model mean summer rainfall than observed, the large inter-model spread is remarkable. Atmospheric models with prescribed observed sea surface temperatures (SSTs) reproduce most of the bias and spread, which indicates that the modeling issues originate, to first order, from the atmosphere (Ashfaq et al. 2016; Bollasima and Ming 2012; Meehl et al. 2006), and more specifically relate to the treatment of convection (Turner and Slingo 2009; Bush et al. 2015). The summer dry bias is also associated with stronger easterly surface winds in spring that cool the adjacent Arabian Sea, which in turn could amplify the dry bias in AIP in summer (Levine et al. 2013; Levine and Turner 2012). The CMIP5 models produce a significant inter-model spread in their projection of the response of the AIP to warming. Figure 1b shows the

 Table 1
 The list of CMIP5 models used in the analysis and comparison of their JJAS mean All India Precipitation with observations

Model name	All India Mon- soon precipita- tion
BNU-ESM	Strong
INM	Strong
MIROC-ESM-CHEM	Strong
CNRM-CM5	Weak
MIROC5	Strong
CCSM4	Strong
BCC-CSM1	Strong
CamESM2	Weak
FGOALS-g2	Strong
GFDL-ESM2G	Strong
CMCC-CM	Strong
GFDL-CM3	Strong
NorESM1-M	Strong
IPSL-CM5A-LR	Weak
MPI-ESM-MR	Weak
MPI-ESM-LR	Weak
IPSL-CM5A-MR	Weak
ACCESS1-0	Weak
MRI-CGCM3	Weak
CSIRO-Mk-3-6-0	Weak
HadGEM2-ES	Weak

Models with their JJAS mean All India Precipitation in the historical simulations larger or smaller than the observations are referred to models with strong or weak monsoon, respectively

projected JJAS mean monsoon precipitation for 2091–2100 under RCP8.5 versus the present day (1996–2005). The projected summer monsoon rainfall correlates well with the present-day rainfall, showing a multi-model mean of 20% increase (the black regression line), with a standard deviation of 26%. Note that models with low present-day rainfall, also referred to as weak monsoon, appear to show a stronger response to warming than the multi-model mean (red dots above the regression line). Conversely, the wetter or strong monsoon models appear to exhibit a weaker response (blue dots below the regression line). This point will be revisited later.

In this study, the origin of the multi-model mean bias and spread in the historical simulations and their implications for uncertainty in the projected changes are examined using moisture budget analysis that takes advantage of the weak temperature gradients in the tropics. In particular, we examine the apparent bimodality of models: models with weak monsoon and stronger response to warming versus models with strong monsoon and weaker response to warming (Fig. 1b), which are denoted consistently using red and blue, respectively, in all the figures. Our analysis focuses on two 10-year windows, 1996–2005 for present day and 2091–2100 for future under the RCP8.5 scenario. Daily precipitation, moisture, and wind data from the 21 climate models that participated in CMIP5 (Table 1) and three global reanalyses including ERA-Interim (ERAI, Dee et al. 2011), the Modern-Era Retrospective Analysis for Research and Applications (MERRA, Rienecker et al. 2011) and NCEP Climate Forecast System Reanalysis (CSFR, Saha et al. 2010) are used.

2 The link to equatorial convection

To examine the regional processes that influence summer AIP, the correlations between the summer AIP with antecedent spring precipitation over the surrounding regions are first analyzed. Correlations of summer AIP with summer precipitation in the surrounding areas have been considered as well, but such correlations could simply be the signature of a shift in precipitation patterns, not an evidence of causality. Analyzing the correlations between spring precipitation with the summer AIP highlights possible causality rather than concurrent changes in spatial distribution. For each of the 21 models, the 10-year mean JJAS precipitation over the land grid points in the red box (earlier referred to as all India precipitation or AIP) is calculated. For each of the 21 models, a global map of mean pre-monsoon (March-April-May) precipitation is also calculated. For each grid point on the global map, the correlation between the set of 21 pre-monsoon precipitation values and the set of 21 summer AIP is calculated to determine the correlation across the 21 models.

Figure 2a shows the resulting map of correlations in shadings. The correlation is considered statistically significant if it is larger than the 95% confidence level from a Student t test. The region in the blue box (hereafter referred to as equatorial Indian Ocean or EIO) represents an area where the MAM precipitation from the models has a strong negative correlation with the mean summer monsoon precipitation in the red box. To understand the physical meaning of this correlation map and hence the dynamical connection between spring EIO rainfall and large-scale circulation and the South Asian summer monsoon, consider the difference in MAM 850 hPa winds between models with weak and strong summer monsoon shown as arrows in Fig. 2a. By weakening the tropospheric thermal contrast with the land, tropospheric heating from excess convection over the EIO in spring induces anomalous northeasterly and northerly winds over India that reduce the quasi-geostrophic monsoon southwesterly flow (Webster 1987; Sun et al. 2010; Yang and Lau 2006) that transport moisture to the region. The latter can delay the onset of the summer monsoon and reduce the monsoon rainfall. Hence models that produce more precipitation



Fig. 2 a Correlation between the 10 year mean JJAS all India precipitation (AIP) averaged over the land area within the red box and the 10 year MAM mean precipitation over every grid point across 21 CMIP5 models. The blue box marks the equatorial Indian Ocean where MAM precipitation is negatively correlated with the JJAS AIP. Both are calculated using 21 CMIP5 models. The arrows indicate the MAM mean 850 hPa wind vectors (m/s) in weak monsoon models minus that is strong monsoon models. **b** Same as **a** but correlation of JJAS AIP with MAM SST at each grid point across the models

over the EIO in spring produce less precipitation over India during summer.

Similar analysis is also performed to investigate the correlations between pre-monsoon (MAM) SSTs with summer AIP and the result is shown in Fig. 2b. Summer AIP shows strong positive correlation with spring SST mainly over Arabian Sea. This positive correlation is consistent with the anomalous easterly winds (Fig. 2a) in spring that keep the Arabian Sea cool in weak monsoon models. The cold Arabian Sea SST bias has been shown by Levine et al. (2013) and Levine and Turner (2012) using numerical experiments to contribute to a weaker summer monsoon in climate models. Hence we hypothesize from Fig. 2 that excess convection over the EIO during spring reduces the tropospheric temperature gradient between land and ocean and induces anomalous northeasterly winds that counter the southwesterly transport of moisture into India while cooling the Arabian Sea. The reduced moisture transport and cool Arabian Sea SST can both delay the onset of the South Asian monsoon and reduce the summer monsoon rainfall. The absence of statistically significant correlations of summer AIP with spring SSTs over the EIO (Fig. 2b) implies that the excess convection and inter-model spread in spring EIO precipitation are primarily of atmospheric origin. Therefore, we focus our analysis on convection over the EIO that may hold the key to understanding the origin of the monsoon biases and spread in models.

2.1 Relationship between precipitation and moisture convergence

In the last subsection it was shown that the magnitude of the monsoon precipitation is strongly linked to the spring season precipitation or convection over the equatorial Indian Ocean. Hence we can reframe the question of the source of bias and uncertainty in monsoon precipitation as an issue of inter-model spread in spring season precipitation over the EIO. Specifically what contributes to the EIO precipitation differences among the models? A simple moisture budget analysis would breakdown the contributions of evaporation, advection and moisture convergence to the EIO precipitation but would provide little insight on the non-linear feedbacks of diabatic heating processes on those terms and the possible role of the representation of convection in the models in the model biases. To address this challenge we present a novel approach that considers the moisture budget along with the energy budget under the weak temperature gradient approximation in the tropics. Because of the weak Coriolis force and the weak horizontal temperature gradients, diabatic heating in the tropics is primarily balanced by vertical advection of potential temperature. This implies a strong coupling between the large-scale vertical velocity and precipitation, which has been the basis for constructing tropical circulation models of simple-to-moderate complexity (Sobel et al. 2001; Sugiyama et al. 2009). The strong coupling has also been exploited to investigate the interactions between various forms of diabatic processes and moisture transport over the tropical oceans and monsoon environments (Zhang and Hagos 2009; Hagos and Zhang 2010; Zhang et al. 2008).

One of the implications of the weak temperature gradient approximation is that the profile of wind convergence has a bi-modal structure. In areas of weak precipitation, the updrafts, if any, are capped by strong subsidence of dry air that results in a shallow circulation, while in areas of stronger precipitation, the updrafts peak near the mid troposphere. To demonstrate this, the following analysis is performed using model outputs for spring over the EIO:

- 1. For each model and reanalysis, 10 years of daily vertically integrated moisture convergence, precipitation and profiles of humidity and divergence are calculated for each $2^{\circ} \times 2^{\circ}$ grid point over the blue box in the EIO region shown in Fig. 2.
- 2. Then the daily vertically integrated moisture convergence values are assigned to one of 30 equally sized bins ranging from -20 to 40 mm/day.
- 3. For each bin, the mean values of the vertically integrated moisture convergence, precipitation and the profiles of humidity and divergence are calculated.

Figure 3 shows the results of this analysis. The red dashed contours represent the levels of zero divergence and indicate the depth of the convergence layer while the blue contours indicate the anomalous specific humidity, which is the specific humidity in every bin minus the specific humidity in the bin where the vertically integrated moisture convergence is zero, which is marked by the vertical dashed line near 5 mm/day precipitation. To the right of the vertical dashed line is the deep convection regime with wind convergence in the lower troposphere and wind divergence in the upper troposphere, positive anomaly of specific humidity, and stronger precipitation. Conversely, to the left of the vertical dashed line is the subsidence regime with wind divergence near the surface, negative anomaly of specific humidity, and weak precipitation. Thus Fig. 3 captures the relationship of wind divergence and anomalous specific humidity with precipitation from a large number of daily events over the equatorial Indian Ocean.

In both the models and the reanalyses, the transition between the subsidence and deep convection regimes occurs near 5 mm/day of daily mean precipitation. Using this bimodality, for each grid point a given day is defined as a deep convection day (denoted by the subscript d) with wind divergence of $\nabla \cdot (v)_d = \nabla \cdot (v)$ if the daily precipitation is greater than 5 mm/day or a subsidence day (denoted by the script s) with wind divergence $\nabla \cdot (v)_s = \nabla \cdot (v)$ if the daily precipitation is less than or equal to 5 mm/day. The 5 mm/ day threshold corresponds roughly to the transition of the vertically integrated moisture convergence from negative to positive values indicated by the dashed blue lines in Fig. 3, which also capture the transition in wind divergence from subsidence to updraft.

An important outcome of the weak temperature gradients and the bi-modality is that the contributions of the moisture convergences in the two regimes are strongly related to the precipitation. Consider the monthly mean precipitation and the monthly mean moisture convergence at a specific grid point, the latter can be partitioned into contributions by deep convection days and contributions by subsidence days as discussed above. Figure 4 shows the relationship between the



Fig. 3 Vertical profiles of wind divergence (color shaded) and specific humidity anomaly (g/kg, blue contours) vs precipitation (mm/ day) from reanalysis data and model simulations over the equatorial Indian Ocean in spring (MAM). The vertical blue dashed lines indicate the precipitation value where moisture convergence

 $-\frac{1}{g}\int_{pt}^{ps}q(\nabla \cdot v)dp$ is zero. The strong (weak) monsoon models have blue (red) caption. The strong (weak) monsoon models are labeled using blue (red) captions above the panels. The red dashed contours indicate levels where wind divergence is zero



Fig. 4 The relationship between monthly mean deep moisture convergence $-\frac{1}{g} \int_{pt}^{ps} q(\nabla \cdot v)_d dp$ and monthly mean total precipitation during spring (MAM) in the reanalysis or model simulations. The dots corre-

spond to deep convection days on grid points in the blue box in Fig. 2. The slope of the linear regression lines represent α_d in Eq. (6)



Fig. 5 The relationship between monthly mean shallow moisture convergence $-\frac{1}{g} \int_{pt}^{ps} q(\nabla \cdot v)_s dp$ and monthly mean total precipitation in spring (MAM). The dots correspond to subsidence days on grid

points in the blue box in Fig. 2 in the reanalysis data or model simulations. The slopes of the linear regression lines represent α_s of the model or reanalysis in Eq. 8

moisture convergences contributed by deep convection and the total monthly mean precipitation from both subsidence and deep convection days. Each dot in the scatter plot represents a grid point in the EIO box in Fig. 2. Similarly Fig. 5 shows the relationship of the moisture convergence contributed by subsidence days to the total monthly precipitation. In both cases the moisture convergence terms are linearly related to the total monthly mean precipitation. These linear relationships will be revisited in the next subsection in the context of moisture budget analysis.

2.2 The moisture budget equation under weak temperature gradient

In the last subsection, it was shown that the monthly mean moisture convergence contributed by days with deep convection is linearly related to the monthly total precipitation and this is also the case for the moisture convergence contributed by days with subsidence, although the regression is weaker for the subsidence regime. This fact will be exploited to simplify the moisture budget equation into a form that allows us to gain insight into the source of bias and uncertainty. We start out with the moisture budget equation:

$$P = E - \frac{1}{g} \int_{pt}^{ps} q \nabla \cdot \mathbf{v} dp - \frac{1}{g} \int_{pt}^{ps} \mathbf{v} \cdot \nabla q dp \tag{1}$$

where P and E are the monthly mean precipitation and evaporation respectively, q is specific humidity, v is the horizontal velocity vector and p_s and p_t are the surface and top level pressure. As discussed in the last section, the moisture convergence (the second term on the RHS of Eq. 1) can be partitioned into two parts as follows;

$$P = E - \frac{1}{g} \int_{pt}^{ps} q(\nabla \cdot \mathbf{v})_s dp - \frac{1}{g} \int_{pt}^{ps} q(\nabla \cdot \mathbf{v})_d dp - \frac{1}{g} \int_{pt}^{ps} \mathbf{v} \cdot \nabla q dp$$
(2)

where the two convergence terms for subsidence (s) and deep convection (d) are monthly means of daily instances with precipitation < 5 and > 5 mm/day, respectively as discussed above. Using the continuity equation for the convergence term for deep convection,

$$-\frac{1}{g}\int_{pt}^{ps}q(\nabla\cdot\mathbf{v})_{d}dp = \frac{1}{g}\int_{pt}^{ps}q\left(\frac{\partial\omega}{\partial p}\right)_{d}dp$$
(3)

where ω is the pressure vertical velocity, and likewise for the subsidence regime.

From the weak temperature gradient form of the energy equation (Holton 1992),

$$-S_p \omega_d = \frac{J_d}{C_p} \tag{4}$$

where S_p is the static stability and J_d is the total diabatic heating. In the tropics, diabatic heating is dominated by latent heating associated with precipitation, so J_d can be approximated by the product of precipitation and a normalized diabatic heating profile \hat{J}_d (Schumacher et al. 2007).

$$J_d = P\hat{j}_d \tag{5}$$

Substituting (4) and (5) into (3) yields

$$-\frac{1}{g}\int_{pt}^{ps}q(\nabla\cdot\nu)_d dp \simeq \alpha_d P + \beta_d \tag{6}$$

where

$$\alpha_d = \frac{1}{C_p g} \int_{pt}^{ps} q \frac{\partial}{\partial p} \left(\frac{\hat{j}(p)}{S_p}\right) dp \tag{7}$$

The validity of Eq. 6 was already demonstrated in the last section by the strong linear relationships between the moisture convergence contributed by deep convection and total precipitation shown in Fig. 4. Hence α_d and β_d in Eq. (6) can be estimated from the slope and intercept of the linear regression fit line for each model (Fig. 4). From Eq. (7), α_d is related to the precipitable water content in the column and the profiles of diabatic heating, moisture and potential temperature (stability).

Similarly, for the subsidence regime,

$$-\frac{1}{g}\int_{pt}^{ps}q(\nabla\cdot v)_s dp\simeq \alpha_s P+\beta_s \tag{8}$$

where α_s and β_s are the slope and intercept of the linear fit line shown in Fig. 5. In deriving Eq. (8), we make the same assumption that the diabatic heating is dominated by latent heating, as for deep convection. The linear regression fit shown in Fig. 5 has weaker correlation compared to Fig. 4 and β_s is not small. These suggest that other sources of diabatic heating such as radiative cooling also play an important role in determining the large-scale circulation in the subsidence regime.

Substituting (6) and (8) into (2), moving $\alpha_d P$ and $\alpha_s P$ to the left hand side, and defining a *normalized precipitable water* as

$$pw_N = 1 - abs(1 - (\alpha_d + \alpha_s)) \tag{9}$$

by dividing both sides of (3) by the right hand side of the equation, we obtain

$$pr_N = \frac{1}{(1 - pw_N)} \tag{10}$$

where

$$pr_{N} = \frac{P}{\left(E - \frac{1}{g}\int_{ps}^{pt} \mathbf{v} \cdot \nabla(q)dp + \beta_{d} + \beta_{s}\right)}$$
(11)

is defined as the *normalized precipitation*. As noted above, the normalized precipitable water is a non-dimensional quantity that represents the effectiveness of moisture supply in generating precipitation, which depends on the thermodynamic profiles (Eq. 7). The normalized precipitation refers to the ratio of precipitation to external supply of moisture that is not directly linked to the convergence in the precipitating column. Equation (10) is analogous to the conventional relationship between precipitation and precipitable water (Bretherton et al. 2004). It shows that, for a given column, the fraction of moisture supply by evaporation and advection that is converted to precipitation is related to the moisture and divergence profiles. Therefore the relationship is intrinsic to the treatment of convection in the respective model.

The above analysis shows that the well-known non-linear relationship between precipitation and precipitable water follows directly from the conservations of moisture and energy under the weak temperature gradient approximation. Specifically, the relationship between the profiles of divergence and moisture determine how precipitation is related to precipitable water (Eq. 7) in each model. As shown in Fig. 3, models with deeper moisture profile and/or shallower divergence

profile can produce larger moisture convergence and precipitation over the EIO in spring, which induce anomalous northeasterly winds (Fig. 2a) and cooler Arabian Sea SSTs (Fig. 2b) and reduce summer monsoon rainfall. In the next section, we discuss the implications of the non-linear relationship between precipitation and precipitable water for the model bias and inter-model spread in precipitation over EIO during spring and AIP over summer.

2.3 Bias and inter-model spread over equatorial Indian ocean

In Sects. 2.1 and 2.2, we used the bi-modality of divergence profile and the linear relationship between moisture convergence and precipitation to reduce the moisture conservation equation to a relationship between the long-term mean precipitation normalized by the sum of evaporation and advection (pr_N) and the precipitable water normalized by the wind convergence (pw_N) , as given by Eq. (10). Figure 6a shows the long-term mean normalized precipitable water and normalized precipitation over the EIO during spring across the CMIP5 model simulations. It shows where each





Fig. 6 a The non-linear relationship between normalized precipitable water (PW) and normalized precipitation from the present day (1996–2005, filled circles) and future (2091–2100, under RCP 8.5, open circles) climate simulations and three global reanalysis datasets. The strong (weak) monsoon models are indicated by the blue (red) circles. The dashed black curve illustrates the non-linear relationship in Eq. (10) derived from the energy and moisture budget with the weak temperature gradient approximation. **b** The relationship between daily

mean precipitation and precipitable water from the CMIP5 models, each indicated by a gray line in the cluster on the left for the present day and the cluster on the right for the future. The locations of the precipitation weighted mean precipitable water (Eq. 12) and the corresponding daily precipitation are marked by the blue (red) circles for the strong (weak) monsoon models and in other colors for the global reanalyses for the present day and blue (red) diamonds for the future

model represented by a dot falls on the theoretical non-linear relationship between the normalized precipitable water and normalized precipitation (Eq. 10) represented by the dashed curve. Many models have normalized precipitable water close to one, so small differences in the normalized precipitable water manifest in a large spread in the normalized precipitation. Because of the non-linear relationship, these models produce larger precipitation compared to models lying on the relatively flat part of the curve. Physically, the non-linear relationship implies that models that effectively utilize moisture from local convergence produce more precipitation than models that rely on moisture supply from evaporation and advection. Not surprisingly models that lie on the steep part of the curve are more likely to have more precipitation over the equatorial Indian Ocean during spring and weaker summer monsoon (more red than blue circles). Most models on the flat part also overestimate the normalized precipitable water compared to the ERA-interim and MERRA reanalyses, but their normalized precipitation bias is very small. Despite the large differences in normalized precipitable water among the three global reanalyses, their differences in normalized precipitation are small because they lie on the flat part of the non-linear curve.

To further understand the physical processes controlling pw_N we revisit Fig. 3 as the normalized precipitable water is related to the depth of the moist layer relative to the depth of the convergence layer. In Fig. 3, the models are sorted by increasing normalized precipitable water. If the moist layer (represented by the blue contours) is shallow and the convergence layer is comparatively deep (represented by blue shadings), the convergence is importing relatively dry air and hence is less efficient in supporting precipitation. That is the case in the reanalyses and in the strong monsoon models. For example, GFDL-ESM2G and CCSM4 depicted in the upper panels have low values of pw_N and have shallow moist layer relative to the convergence layer. In the weak monsoon models, the moist layer is relatively deep, so moisture is effectively utilized by the convergence to produce stronger precipitation in the equatorial Indian Ocean (e.g., CSIRO-Mk-3-6-0). In other words, if the moist layer is deep compared to the convergence layer, moisture supply from the updraft dominates the balance with precipitation so $pw_N \sim 1$. Hence pw_N is intimately related to the model representation of convection, which influences the profiles of moisture and convergence and moisture-precipitation feedback.

The non-linear relationship between the normalized 10-year mean precipitation and precipitable water across the models also manifests in the actual (non-normalized) daily precipitation and precipitable water in the EIO for each model. Figure 6b shows the actual rather than normalized precipitable water and precipitation obtained by constructing the histogram of 30 bins from the daily precipitable water values (gray lines) for each model. The dots represent the

mean value of precipitable water for each model and are defined as

$$PW_m = \frac{\sum pw \cdot P}{\sum P} \tag{12}$$

where PW and P in this case represent daily values including all grid points in the blue box (Fig. 2a) and all the 10 years. The precipitable water at every grid point is weighed according to its contribution to the mean precipitation because at any given grid point only a small fraction of days have enough precipitable water to produce precipitation. The figure shows the non-linear relationships in gray lines for each model and reanalysis based on daily precipitation and precipitable water. Again, these non-linear relationships are reminiscence of the well-known relationship between precipitation and precipitable water observed in the tropics (Bretherton et al. 2004) and has been proposed to arise from self-organized criticality due to continuous phase transitions in tropical cloud populations (Peters and Neelin 2006; Peters et al. 2009). Our derivation and analysis in Sect. 2.2 show that the non-linear relationship follows from moisture and energy conservation under the weak temperature gradient approximation. For each model the markers indicate the long-term average precipitable water and the corresponding precipitation. Similar to the normalized values, the intermodel spread of mean precipitation increases with the mean precipitable water. This is particularly apparent from the triangular distribution of the markers. The non-linear relationship between precipitation and precipitable water favors a higher multi-model mean precipitation because the same increase in precipitable water has a stronger effect at high precipitable water (steep part of the curve) than a comparable increase at lower precipitable water (the flat part of the curve). Furthermore the inter-model spread increases with global warming; that is, as the model spread in precipitable water increases with warming, the spread in precipitation also increases following the shape of the non-linear curves.

3 Implications for monsoon bias, inter-model spread and projections

The implications of the bi-modality of model behaviors in the EIO in terms of where each model falls in the nonlinear relationship between precipitation and precipitable water for the monsoon bias and spread are examined further in this section. Here we divide the models into two groups based on the conditions over the EIO: those with pw_N less than the median of the 21 models (low pw_N hereafter) and the rest (high pw_N). A comparison of the seasonal cycle of monsoon precipitation from the two



Fig. 7 a The seasonal cycle of present day (1996–2005) average All India Precipitation (mm/day). The red and blue lines correspond to the mean of two groups of 10 CMIP5 models with low and high normalized precipitable water (*PWn*) (see text). The shaded areas indicate +/- one standard deviation for each group of models. **b** The model projected future precipitation under RCP 8.5 for the two groups of models versus the present day precipitation. The filled blue (red) circles correspond to models that have both strong (weak) monsoon and low (high) normalized precipitable water, so the normalized precipitable water in the equatorial Indian Ocean during spring (MAM) is a skillful predictor of the summer monsoon strength in these models. The open blue (red) circles correspond to models that have strong (weak) monsoon but high (low) normalized precipitable water. The red (blue) lines mark the multi-model mean increases of 28 and 15%

groups with observations is shown in Fig. 7. The precipitation from models with low pw_N is in more general agreement with the observed. It is apparent that the dry monsoon precipitation bias in the multi-model mean is driven almost exclusively by models of high pw_N , which exhibit excessive precipitation over the EIO and significant delay in monsoon onset, with the dynamical mechanism explained in Fig. 2. Using pw_N as a predictor of the summer monsoon strength, 8 out of 10 models with low pw_N are correctly predicted to have a strong monsoon and 8 out of 11 models with high pw_N are correctly predicted to have a weak monsoon so the normalized precipitable water over the EIO in spring is a skillful predictor for the summer monsoon rainfall. Another important difference between the two groups of models pertains to their projection of future summer mean monsoon precipitation (Fig. 7b). On average models with low pw_N project a 15% increase between 1996 and 2005 and 2091-2011 under the RCP8.5 scenario with little spread (blue line). The projected changes from models with high pw_N have a significant spread with a mean of 28% increase. Following the above discussion, the multi-model mean of 20% increase from all models projected by the end of the century is likely an overestimate that is heavily weighed by models with present day pw_N and pr_N in the EIO that are much larger than those from the reanalyses and therefore exhibit dry biases in monsoon precipitation (Fig. 7a).

4 Conclusion

As one of the most prominent circulation and hydrological features in the earth system, accurate simulation of the South Asian monsoon and building confidence in its projected response to anthropogenic forcing is of great societal value and a major challenge. Many of the global climate models that participated in CMIP5 display a dry bias in their simulation of the present day South Asian monsoon precipitation. Correlation analysis shows that the mean dry bias and inter-model spread in summer monsoon precipitation are respectively linked to the mean excess precipitation and its inter-model spread over the equatorial Indian Ocean earlier in spring (Fig. 2), which favors stronger easterlies and cooler Arabia Sea SSTs that delay the summer monsoon onset and weaken the monsoon precipitation. On the other hand, the absence of correlation between summer monsoon precipitation and SST in the EIO suggests that the monsoon bias is essentially of atmospheric origin. This study therefore uses moisture budget analysis under weak temperature gradients to identify the origin of the bias and spread over the equatorial Indian Ocean and examine the implications for that of the monsoon and its projected changes under the RCP 8.5 scenario.

We show that under the weak temperature gradient approximation moisture convergence over subsidence and deep convection areas are linearly related to precipitation so the moisture budget equation can be reduced to a non-linear relationship between the normalized precipitation and normalized precipitable water (Eq. 10). The former represents precipitation divided by evaporation and advection and the latter represents the effectiveness of convergence profile at importing moisture. The steep gradient in the non-linear curve relating the normalized precipitation and precipitable water (Fig. 6a) implies that small differences in the normalized precipitable water manifest in large differences in the normalized precipitation. Such models (i.e., models that lie on the steep part of the curve) produce higher precipitation over the equatorial Indian Ocean, and contribute disproportionately to the large inter-model spread and multimodel mean dry bias in monsoon precipitation (Fig. 6b). On the other hand, models that show weaker sensitivity of normalized precipitation to normalized precipitable water (i.e., models that lie on the flat part of the curve in closer agreement with the reanalyses) simulate seasonal cycle of the monsoon precipitation that is in general in better agreement with observations (Fig. 7a). It should be noted that the non-linear relationship between precipitable water and precipitation is a diagnostic relationship that follows from the moisture and energy budgets so it does not imply causality between the two quantities. Rather the relationship depends on the relationship between the profiles of convergence and moisture, which is an essential outcome of the cumulus parameterization and parameter choices in models (for example entrainment, Bush et al. 2015). Non-linear feedbacks between the import of moisture by convergence and the diabatic heating the determines the profile of convergence likely play a significant role in the bi-modality of model behaviors shown in Fig. 6.

The relationship between the normalized precipitable water and normalized precipitation also affects modelprojected response to warming. Models on the steep part of the non-linear relationship show broader range of sensitivity to warming than those on the flat part (Figs. 6b, 7b). On average models with low normalized precipitable water (i.e., those with good agreement with observation and reanalysis) project a 15% increase between 1996 and 2005 and 2091-2011 under the RCP8.5 scenario with a spread of 10%, while models with higher precipitable water project a mean increase of about 28% increase with a spread of 31%. Therefore, both the multi-model mean projected 20% increase by the end of the century and the 26% spread among all the models are likely overestimates that are heavily skewed by the non-linearity and the larger number of models with higher precipitable water, as discussed above. Bollasina and Ming (2012) also noted the important role of model biases in the southwestern EIO on South Asian monsoon rainfall biases. Based on experiments from a single model, they attributed the relationship between EIO and monsoon rainfall biases to the model excess response to the local meridional SST gradient. Ashfaq et al. (2016) noted that diabatic processes over land also influence monsoon rainfall biases. Levine et al. (2013) showed that cold SST bias in the Arabian Sea contributes to weak summer monsoon in models. Here our analysis of correlations (Fig. 2) points to the convection over the EIO in spring as the source of anomalous northeasterly winds that delay the onset of the summer monsoon and cool the Arabian Sea that may further weaken the summer monsoon, as noted by Levine et al. (2013). We further identify the normalized precipitable water over the EIO as a critical parameter distinguishing model skill in simulating South Asian monsoon precipitation

and highlight the far-reaching impact of model biases and uncertainties in the treatment of tropical convection over the oceans on regional precipitation over land.

The non-linear nature of the processes that determine the normalized precipitable water and normalized precipitation implies that they not only amplify small differences in precipitable water to large spread in precipitation, but the non-linear relationship between precipitation and precipitable water can introduce a mean bias in the monsoon precipitation and asymmetric inter-model spread. In a warmer climate, uncertainty in climate sensitivity increases the inter-model spread in precipitable water, which further amplifies the inter-model spread in precipitation through the non-linear relationship between the two. The normalized precipitable water, defined here as the covariance between moisture convergence and precipitation, is a key metric for evaluating the fidelity of climate models, with high predictive power for where a CMIP5 model falls in the non-linear relationship between precipitation and precipitable water in the EIO and whether the model produces a strong or weak monsoon (Fig. 7a). As this parameter directly estimates the sensitivity of convection to small perturbations, it can be used to provide an important constraint on convective parameterizations.

Acknowledgements This research is supported by the U.S. Department of Energy Office of Science Biological and Environmental Research as part of the Regional and Global Climate Modeling Programs. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RLO1830.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ashfaq M, Rastogi D, Mei R, Touma D, Leung RL (2016) Sources of errors in the simulation of south Asian summer monsoon in CMIP5 GCMs. Clim Dyn. https://doi.org/10.1007/s0038 2-016-3337-7
- Bollasina M, Ming Y (2012) The general circulation model precipitation bias over the southwestern equatorial Indian Ocean and its implications for simulating the South Asian monsoon. Clim Dyn 40:3–4. https://doi.org/10.1007/s00382-012-1347-7

- Bretherton CS, Peters ME, Back LE (2004) Relationships between water vapor path and precipitation over the tropical oceans. J Climate 17:1517–1528
- Dee DP et al (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597
- Hagos S, Zhang C (2010) Diabatic heating, divergent circulation and moisture transport in the African monsoon system. Q J R Meteorol Soc 136(S1):411–425
- Holton JR, (1992) An introduction to dynamic meteorology, vol 48 of International geophysics series, ISSN 0074–6142
- Huffman GJ,RF, Adler MM, Morrissey S, Curtis R, Joyce B, McGavock, Susskind J (2001) Global precipitation at one-degree daily resolution from multi-satellite observations. J Hydrometeor 2:36–50
- Huffman GJ, Adler RF, Bolvin DT, Gu GJ, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB (2007) The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol 8:38–55
- Levine RC, Turner AG (2012) Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature biases. Climate Dyn 38:2167–2190
- Levine RC, Turner AG, Marathayil D, Martin GM (2013) The role of northern Arabian Sea surface temperature biases in CMIP5 model simulations and future projections of Indian summer monsoon rainfall. Clim Dyn 41:155–172. https://doi.org/10.1007/s0038 2-012-1656-x
- Meehl GA, Arblaster JM, Lawrence DM, Seth A, Schneider EK, Kirtman BP, Min D (2006) Monsoon regimes in the CCSM3. J Clim 19:2482–2495
- Peters O, Neelin JD (2006) Critical phenomena in atmospheric precipitation. Nat Phys 2:393–396. https://doi.org/10.1038/nphys314
- Peters O, Neelin JD, Nesbitt SW (2009) Mesoscale convective systems and critical clusters. J Atmos Sci 66:2913–2924. https://doi. org/10.1175/2008JAS2761.1

- Rienecker MM et al (2011) MERRA—NASA's modern-era retrospective analysis for research and applications. J Clim 24:3624–3648
- Saha S et al (2010) The NCEP climate forecast system reanalysis. Bull Am Meteor Soc 91(8):1015–1057. (https://doi.org/10.1175/2010B AMS3001.1)
- Schumacher C, Zhang MH, Ciesielski PE (2007) Heating structures of the TRMM field campaigns. J Atmos Sci 64:2593–2610
- Sobel AH, Nilsson J, Polvani LM (2001) The weak temperature gradient approximation and balanced tropical moisture waves. J Atmos Sci 58:3650–3665
- Sperber KR, Annamalai H, Kang I-S, Kitoh A, Moise A, Turner A, Wang B, Zhou T (2013) The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. Clim Dyn 41:2711–2744
- Sugiyama M (2009) The moisture mode in the quasi-equilibrium tropical circulation model. Part I: Analysis based on the weak temperature gradient approximation. J Atmos Sci 66:1507–1523
- Sun Y, Ding Y, Dai A (2010) Changing links between South Asian summer monsoon circulation and tropospheric land-sea thermal contrasts under a warming scenario. Geophys Res Lett 37:L02704. https://doi.org/10.1029/2009GL041662
- Turner AG, Slingo JM (2009) Uncertainties in future projections of extreme precipitation in the Indian monsoon region. Atmos Sci Let 10:152–158
- Webster PJ (1987) The elementary monsoon. In: Fein JS, Stephens PL (eds) Monsoons. Wiley, New York, pp 5–30
- Yang S, Lau K-M (2006) Interannual variability of the Asian monsoon. In: Wang B (ed) The Asian monsoon. Praxis, pp 268–280
- Zhang C, Hagos S (2009) Bi-model structure and variability of largescale diabatic heating in the tropics. J Atmos Sci 66:3621–3640
- Zhang C, Nolan DS, Thorncroft CD, Nguyen H (2008) Shallow meridional circulations in the tropical atmosphere. J Climate 21:3453–3470