

# Sampling variability and the changing ENSO–monsoon relationship

Benjamin A. Cash<sup>1</sup>  · Rondrotiana Barimalala<sup>3</sup> · James L. Kinter III<sup>1</sup> · Eric L. Altshuler<sup>1</sup> · Michael J. Fennessy<sup>1</sup> · Julia V. Manganello<sup>1</sup> · Franco Molteni<sup>2</sup> · Peter Towers<sup>2</sup> · Frederic Vitart<sup>2</sup>

Received: 17 March 2016 / Accepted: 15 August 2016 / Published online: 25 August 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** The impact of sampling variability on the correlation between all-India rainfall (AIR) and the El Niño–Southern Oscillation is investigated in a large ensemble of seasonal climate simulations made using the European Centre for Medium-Range Weather Forecasting Ensemble Prediction System at T319 (64 km). The analyzed runs consist of 51 ensemble members initialized each May 1 for the period 1980–2011 and integrated for 7 months. 10,000 pairs of 32-year timeseries of June–September (JJAS) mean AIR and NINO3 indices are created from this database by randomly drawing one of the 51 ensemble members for each year. The correlation between each pair of AIR and NINO3 series is then calculated, generating a distribution of AIR–NINO3 correlation values. The model is reinitialized with observations each May 1 and thus all members are drawn from the same background state by construction and any differences in correlation are attributable to sampling variability. The spread in the calculated correlation values and the differences between 32-year segments are sufficient to explain the observed variations in AIR–NINO3 correlation since the beginning of the 1900s, including the sharp decrease in correlation strength since the late 1970s. Sampling variability thus represents a strong null hypothesis for the observed changes and one that cannot be rejected at the 95 % level based on our simulations.

**Keywords** Monsoon · ENSO · Correlation · Sampling · Climate change

## 1 Introduction

The South Asian Summer Monsoon (SASM) represents one of the most important features of the global climate system. Interannual variations in the strength of the monsoon rains affect weather patterns across the globe and the lives and livelihoods of fully half the world’s population. Efforts to understand and predict (Blanford 1884; Walker and Bliss 1932) the year-to-year variations in the monsoon rains have been ongoing for well over a century. Many of the efforts to predict these variations have been statistical in nature, with a focus on identifying, understanding, and utilizing the time-lagged associations between rainfall over southeast Asia and remote sea surface temperature (SST) anomalies. These relationships have the potential to provide predictability even to quantities as notoriously chaotic as precipitation by linking them to more slowly and smoothly varying features of the climate system, such as ocean temperature (Charney and Shukla 1981).

Perhaps the most commonly studied relationship has been that of all-India rainfall and the El Niño–Southern Oscillation. All-India rainfall (AIR), defined as rainfall over the Indian landmass, and the El Niño–Southern Oscillation (ENSO) were significantly negatively correlated throughout much of the twentieth century. Lower than normal AIR is often associated with El Niño events. Likewise, AIR has tended to be above average during La Niña events (Rasmusson and Carpenter 1983). However, the strength of this relationship between ENSO and the monsoon, as measured by correlation coefficients calculated over a sliding window, has not been stationary with time (e.g.,

---

✉ Benjamin A. Cash  
bcash@gmu.edu

<sup>1</sup> George Mason University/Center for Ocean–Land–Atmosphere Studies, Fairfax, VA, USA

<sup>2</sup> European Centre for Medium-Range Weather Forecasts, Reading, England, UK

<sup>3</sup> Shanghai Jiao Tong University, Shanghai, China

Kumar et al. 1999; Gershunov et al. 2001). Values of the 32-year correlations were between  $-0.4$  and  $-0.45$  at the beginning of the 1900s, before increasing in magnitude to roughly  $-0.55$  with peak values near  $-0.6$  for the period 1935–1980 (see discussion of Fig. 6 below). The correlation then decreased sharply in magnitude in recent decades to roughly  $-0.2$ , a value that is not significantly different from 0 at the 90 % level. At the same time the association between ENSO and the East Asian Summer Monsoon (EASM) has been strengthening (e.g., Wu et al. 2012).

Thus, after peaking in the mid-twentieth century, the AIR–ENSO correlation has since weakened dramatically. One notable instance of this apparent breakdown in the negative relationship between the ISM and El Niño was the 1997 monsoon season, in which AIR was near normal despite the presence of a strong El Niño event. This led to significant errors in the seasonal monsoon forecast for that year (Kumar et al. 2006).

Multiple theories have been proposed for the observed trends in the ENSO–monsoon relationship, ranging from the influence of the North Pacific (Kinter et al. 2002), the tropical South Atlantic (Kucharski et al. 2009), the Indian Ocean Dipole (Ashok et al. 2004), a strengthening of the jet-stream over the North Atlantic (Chang et al. 2001), and global warming (Kumar et al. 1999). In addition to these hypotheses, it has been noted that the observed changes could arise simply due to sampling variability (Gershunov et al. 2001; DelSole and Shukla 2012). Gershunov et al. (2001) (hereafter G2001) in particular showed that the association between ENSO and the monsoon had undergone decadal-scale fluctuations in strength since the late 1800s in the observations, and that these fluctuations were actually lower in magnitude than what would be expected from simply correlating two white noise time series.

In this study we analyze a suite of integrations made using a state-of-the-art seasonal prediction model (Molteni et al. 2011) to expand upon these previous works. We show that the observed changes in the strength of the monsoon–ENSO relationship are consistent with fluctuations due simply to random sampling variations of internal climate variability. We find that nearly identical background states can produce values of the monsoon–ENSO correlation ranging from  $-0.73$  to  $0.47$ , and differences between 32-year segments range from  $-0.89$  to  $0.85$ . The observed changes in the AIR–NINO3 correlation over the past 100 years cannot be distinguished from changes due to sampling at the 90 % confidence level. Thus, based on our model results, we cannot reject the hypothesis that sampling variability is responsible for the observed variations in the ENSO–monsoon relationship at the 90 % level in favor of the impact of climate change.

## 2 Materials and methods

We analyze a suite of 7-month retrospective forecasts (hindcasts) made using a model based on the European Centre for Medium-Range Weather Forecasting (ECMWF) Ensemble Prediction System as part of *Project Minerva* (e.g., Zhu et al. 2015; Manganello et al. 2016). Runs were initialized May 1 and November 1 for each year in the period 1980–2011. Integrations were performed using the ECMWF Integrated Forecast System (IFS) coupled to Nucleus for European Modeling of the Ocean (NEMO) in a configuration similar to that of System 4 (Molteni et al. 2011). Integrations were performed for three different atmospheric resolutions: T319L91 (64 km horizontal resolution on the reduced Gaussian grid, 51 ensemble members), T639L91 (32 km horizontal resolution, 15 ensemble members), and T1279L91 (16 km horizontal resolution, 15 ensemble members) each with 91 levels in the vertical and coupled to the same 1-degree ocean resolution (refined to  $1/3$  degree near the equator).

In this work we focus on the boreal summer monsoon season (June–September; JJAS) in the 7-month 51-member T319 experiments initialized each May 1. Given that the length of the individual integrations is 7 months the JJAS means for 1 year are only related to values from the previous or following year in that they are generated by the same model. They are not part of a single, continuous integration. Instead each May 1 the model state is reset to the best estimate of the observed climate state, plus small perturbations for each ensemble member (see Molteni et al. 2011). As a result, any multi-year time series of seasonal (JJAS) means represents a set of values that are drawn from related model states but are not part of a single continuous model trajectory. While ensemble members are numbered 1 through 51 for convenience there is no more direct relationship between ensemble member 1 in 1980 and 1981 than there is between ensemble members 1 and 51 in those same years.

We exploit this semi-independent aspect of the ensemble members to construct multiple possible time series of 1980–2011 JJAS values via bootstrapping by randomly selecting one ensemble member for each year (see Fig. 1 for an illustration). In the simplest approach the same ensemble member is drawn for each year. However, as noted above, it is equally valid to draw any combination of members, leading to  $51^{32}$  ( $\sim 10^{54}$ ) potential combinations for the 32-year period. For this work we create 10,000 32-year samples from the T319 simulations and analyze the statistics of the monsoon and its relationship to other features in the simulations, specifically ENSO.

The fact that all 51 members are generated via small perturbations around the same observed state and run for less

**Fig. 1** Schematic for random draws used to create sample time series from Minerva data. *Black and red arrows* denote possible draw pathways

|     | May-Nov<br>1980 | May-Nov<br>1981 | May-Nov<br>1982 | ..... | May-Nov<br>2011 |
|-----|-----------------|-----------------|-----------------|-------|-----------------|
| 0   | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |
| 1   | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |
| 2   | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |
| 3   | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |
| 4   | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |
| 5   | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |
| ... |                 |                 |                 |       |                 |
| 50  | JJAS            | JJAS            | JJAS            | JJAS  | JJAS            |

than a year ensures that all runs experience nearly identical external forcing and background state. The deep ocean, land surface, sea surface, sea ice, and other components of the system with the potential to vary on decadal timescales are all initialized with small perturbations around the same basic state. All runs are also initialized from equally possible initial conditions. Thus differences between the various constructed time series arise purely from sampling different instances of climate noise, rather than decadal-scale variations.

### 3 Results

#### 3.1 Model fidelity

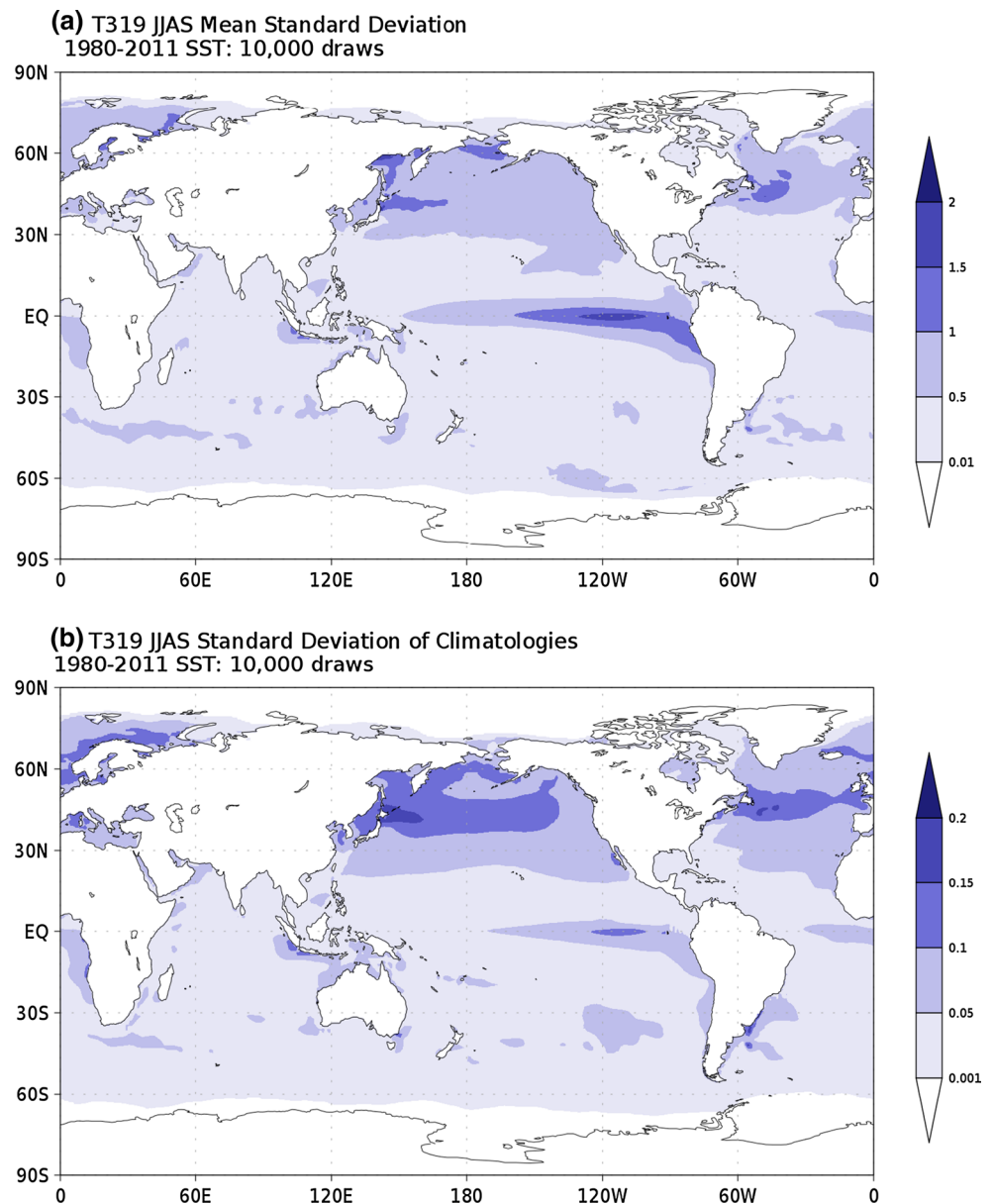
As a measure of the consistency of the different climatological means, and a check on the assertion that each time-series is drawn from the same 30-year base state, we first calculate the 30-year climatology of SST for each of our 10,000 samples. We then calculate the standard deviation of these 10,000 climatologies at each grid point. As shown in Fig. 2, the regions where the standard deviation of the climatologies is largest closely corresponds to the areas with the highest interannual variability, suggesting that differences in sampling the intraseasonal variability of the different ensemble members can affect the 30-year climatology. However, the standard deviation of the climatologies is small relative to the interannual variability, with relative magnitudes of generally 10–15 %. Thus differences between any two climatologies at any given point

are significantly less than their respective interannual variations. Similar ratios hold for precipitation (Fig. 3) and for 700 m heat content (not shown).

As with all models, the climatological precipitation in the Minerva simulations include a measure of systematic error (Fig. 4). The model produces excess precipitation over the dry northwest region of India, as well as across the central plains. Rainfall is underestimated directly over the Western Ghats and overestimated to the east of the mountains. Rainfall near Bangladesh and the southeast region of peninsular India is also underestimated. Interestingly, there is only minimal reduction in the mean bias as atmospheric horizontal resolution increases from 64 km (Fig. 4a) to 16 km (Fig. 4b). The model also tends to produce too little interannual variability in the seasonal mean rainfall total at both 64 km (Fig. 5) at 16 km (not shown), particularly over the sharp topography of the Western Ghats.

Despite the presence of these systematic biases, the ensemble of T319 simulations generally succeeds in bracketing the interannual variability of the AIR (Fig. 6). While the correlation between the ensemble mean and observed AIR is relatively low (0.29), and not significantly different from zero, there are only a small number of years (2002, 2004, 2009) where the observations fall entirely outside the envelope of the ensemble. In the majority of years the model successfully brackets the observations. This includes the much-discussed 1997 monsoon (noted above). While the ensemble mean suggests a drought, the strongest one in the model record, there are multiple ensemble members with comparable or even higher AIR than the observations. This raises the question as to whether difficulty in

**Fig. 2** Comparison of **a** interannual variations in sea surface temperature (SST) and **b** variations between individual climatologies from the 10,000 samples generated from the random draw. Both panels depict standard deviations. Note reduction in scale by factor of 10 in **b**



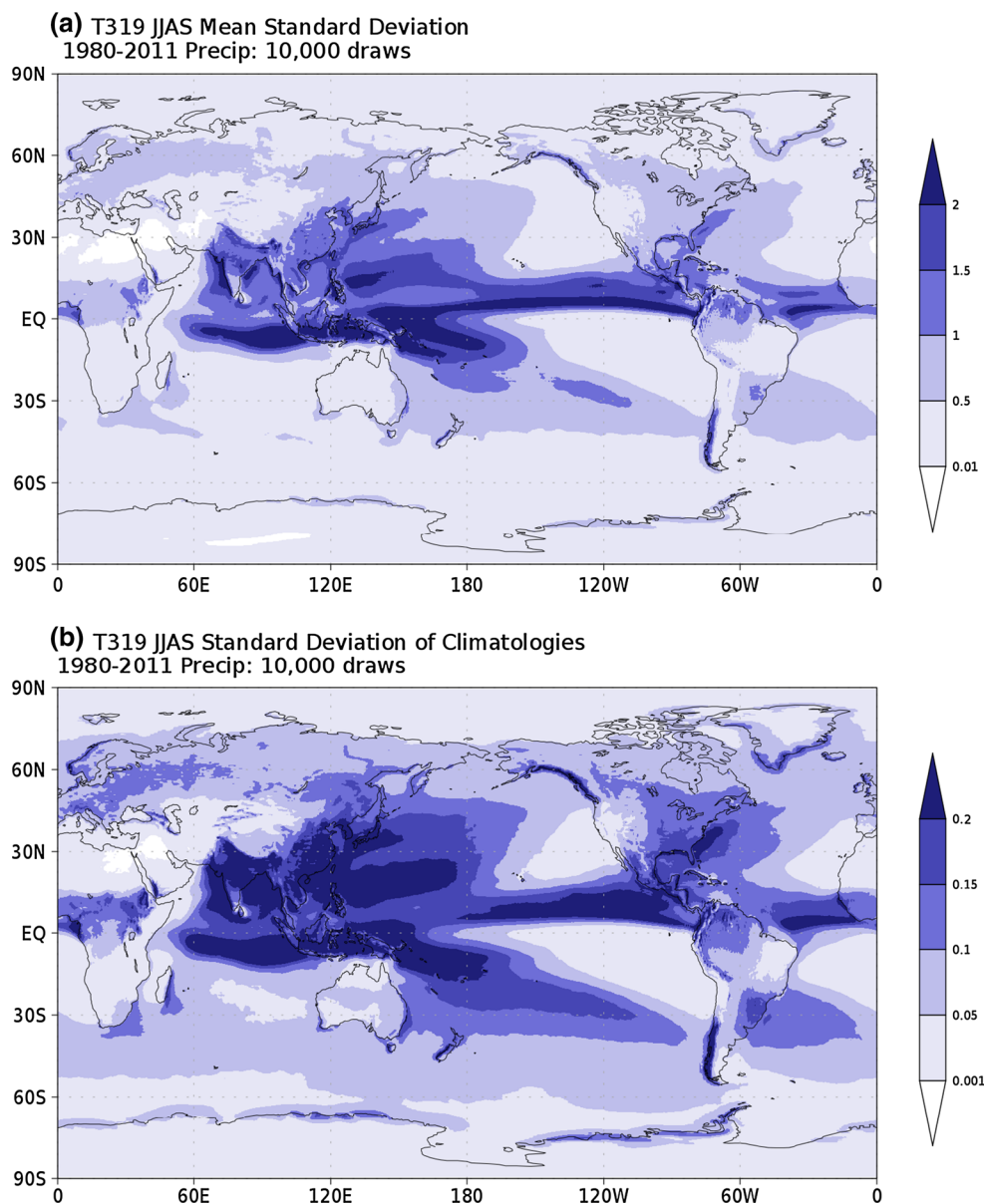
representing the 1997 monsoon, or any other unusual event, reflects model limitations or the likelihood of the event.

### 3.2 The ENSO–monsoon relationship

The weakening of the ENSO–monsoon relationship is frequently expressed as a reduction in the magnitude of the sliding correlation between AIR and an SST index representing ENSO activity as a function of time (e.g., Kumar et al. 1999; Gershunov et al. 2001; DelSole and Shukla 2012). Here we calculate the observed contemporaneous JJAS AIR–NINO3 correlation for two sliding windows of

different lengths, 21- and 32-years (Fig. 7) for the period 1901–2010. The sliding correlations are performed as follows. Taking the 21-year window as an example, we first calculate the correlation between AIR and NINO3 for the period 1901–1921, then 1902–1922, etc., proceeding the end of the analysis period. Each 21- or 32-year period is linearly detrended to focus on the association between interannual variations in AIR and NINO3. Rainfall data is taken from the Indian Meteorological Department (IMD) 0.25° gridded daily product (Pai et al. 2013) and SST data is taken from the Hadley Centre v1.1 product (Rayner et al. 2003).

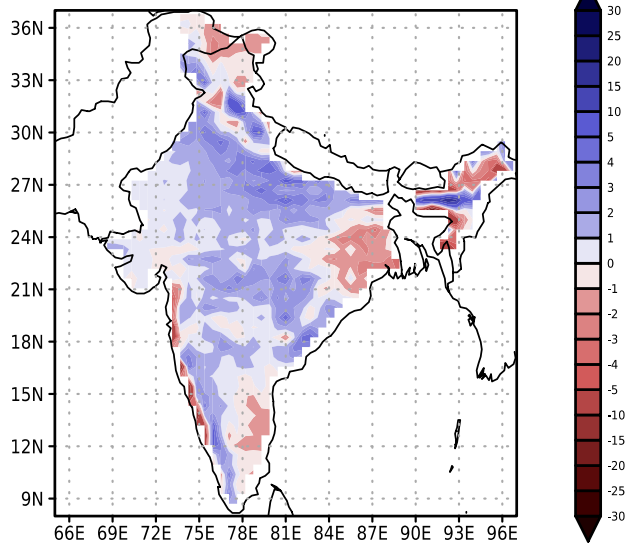
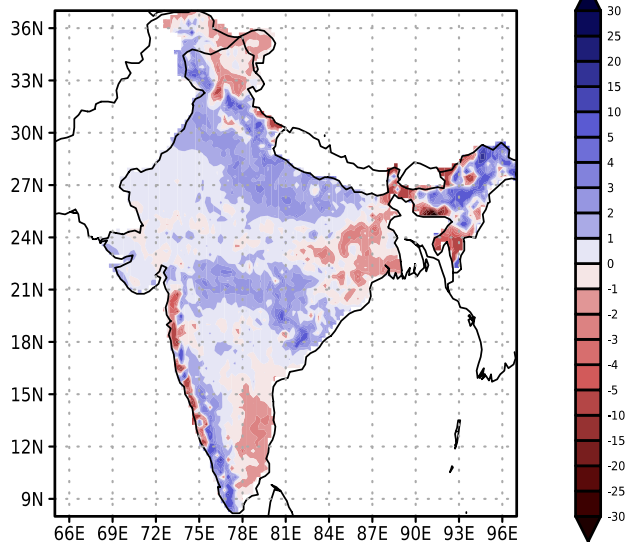
**Fig. 3** Comparison of **a** inter-annual variations in precipitation and **b** variations between individual climatologies from the 10,000 samples generated from the random draw. Both panels depict standard deviations. Note reduction in scale by factor of 10 in **b**



As expected, the 21-year window shows much greater variability and sensitivity to choice of period than the 32-year window, with changes in a single year producing changes in correlation magnitude by over 0.1. The 32-year sliding window is generally less sensitive to single-year changes in the window but jumps of 0.1 still occur. Given this sensitivity focusing on any specific value seems unwise; however, certain general patterns are apparent in both time series. Values tend to lie near  $-0.35$  for the earliest part of the record (1901–1930), increase in magnitude to between  $-0.5$  and  $-0.7$  from 1930 to 1970, and then the magnitude decreases rapidly to roughly  $-0.2$  in the current

period. The maximum change in the 32-year sliding correlation occurs between periods centered in the 1960s and recent decades, with a magnitude of 0.4.

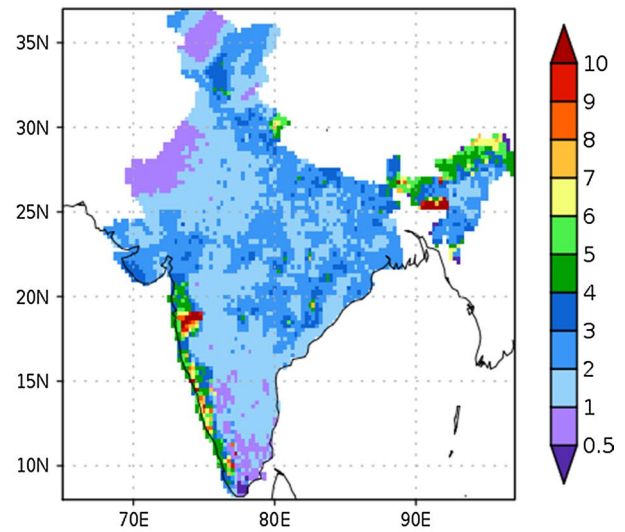
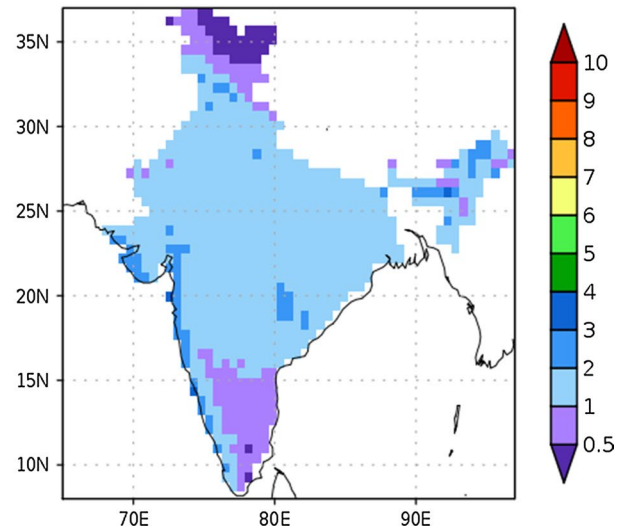
A key question in understanding these observed changes in the strength of the ENSO–monsoon relationship is how likely they are to arise in the absence of changes in the background state of the system. G2001 used a similar bootstrapping approach to the one presented here to show that correlating time series of white noise could generate a distribution of correlation coefficients consistent with the observed record. Building upon that approach, in this work we directly calculate a distribution of potential AIR–NINO3 correlations using

**(a)** T319 1980-2010 JJAS rainfall bias (mm/d)**(b)** T1279 JJAS rainfall bias (mm/d)

**Fig. 4** Ensemble mean bias in precipitation for June–September (JJAS) for the period 1980–2011 relative to the India Meteorological Department (IMD) 0.25 degree rainfall product for **a** T319/64 km and **b** T1279/16 km simulations

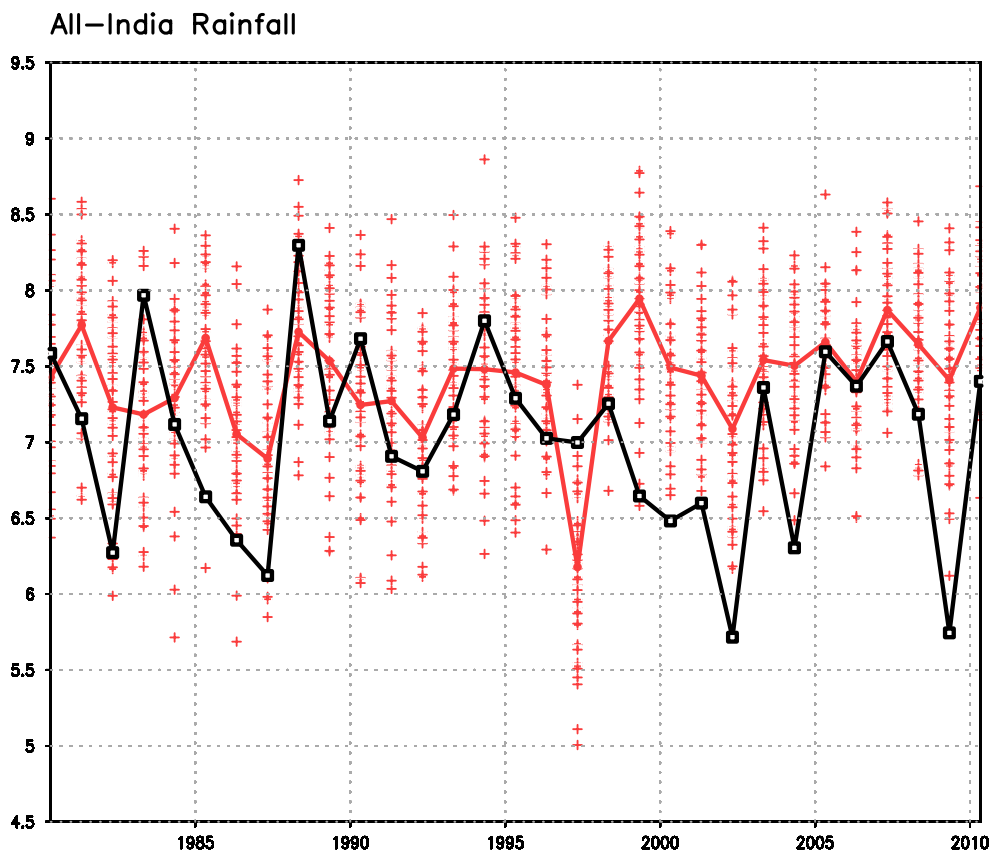
our 10,000 constructed Minerva time series (Fig. 8). This allows us to not only further assess the fidelity of the Minerva simulations but also the impact of sampling a more realistic estimate of variability on the ENSO–monsoon relationship.

We first consider the case where the impact of internal variability is suppressed by taking the ensemble mean.

**(a)** Observed standard deviation of JJAS rainfall**(b)** T319 standard deviation of JJAS rainfall

**Fig. 5** Interannual standard deviation of JJAS precipitation for the period 1980–2011 from **a** IMD and **b** T319/64 km simulation

The correlation between NINO3 and AIR in this case is  $-0.52$ , which is significantly stronger than the observed value of about  $-0.2$  for the recent period. However, it is relatively close to the value of  $-0.45$  for the full 109-year observational period. Both  $-0.45$  and  $-0.52$  are significantly different from 0 at the 99 % confidence level, and indicate that the ENSO–monsoon relationship is negative and statistically significant when the influence of climate noise is reduced in both the model and observations.



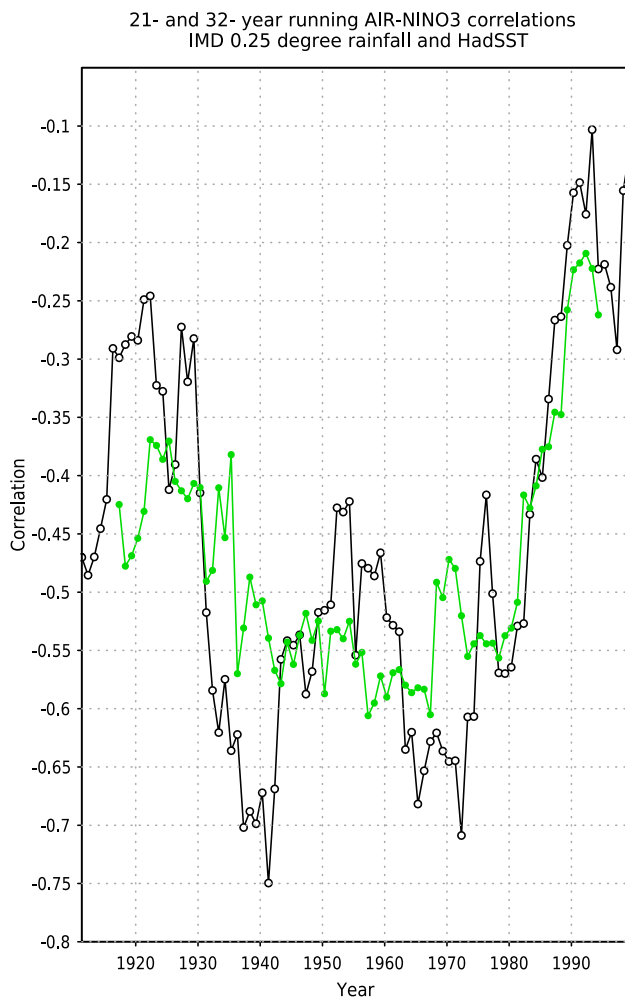
**Fig. 6** JJAS all-India rainfall (AIR) in mm/day for the period 1980–2011 from the T319/64 km Minerva simulations (*red line* denotes ensemble mean, *red crosses* denote individual ensemble members) and observations from IMD (*black line*)

The median of distribution is  $-0.27$ , placing the observed value of  $-0.2$  well within one standard deviation of the peak. The minimum observed value of about  $-0.6$  also falls within the model distribution, and its magnitude is exceeded in  $\sim 1\%$  of the samples. Thus, even the strongest correlations observed are present in the Minerva simulations, although as one would expect they occur relatively infrequently. This is despite the fact that the background state for simulations is for the current period only. It is also despite the fact that, as indicated by Fig. 5, the variability in the model is actually lower than observed and may be underestimating the spread in the distribution. Interestingly the Minerva simulations suggest a significant probability of observing a positive correlation between AIR and NINO3, although this value is not observed in the relatively short observational record.

The distribution of correlation values from the Minerva samples can be closely approximated as the distribution of a white noise process. To show this we calculate a 10,000-sample distribution of  $\rho_{(x,y)} = rx + w\sqrt{(1-r^2)}$ , where  $w$  and  $x$  are independent identically distributed

samples of length 32 drawn from white noise processes. We find that the two correlations distributions are nearly identical if  $r$  is chosen as the peak Minerva value of  $-0.27$  (Fig. 8). Thus in this metric the differences between the different ensemble members, which arise due to the chaotic internal variability of the model, are consistent with the differences introduced by sampling white noise. Examination of SST and precipitation composites (not shown) associated with the 50 strongest positive and negative correlations found no notable differences, further supporting the idea that sampling noise is responsible for the differences.

In addition to calculating the distribution of correlation values generated by sampling variability we also calculate the distribution of changes between different 32-year periods (Fig. 9). As we would expect the distribution peaks near zero, and has a standard deviation of 0.22. If we take the maximum change in 32-year correlations from the observed record of  $\sim 0.4$ , we find that changes of this magnitude occur in 10% of the cases for random samples drawn from the Minerva simulations. Thus, while a change

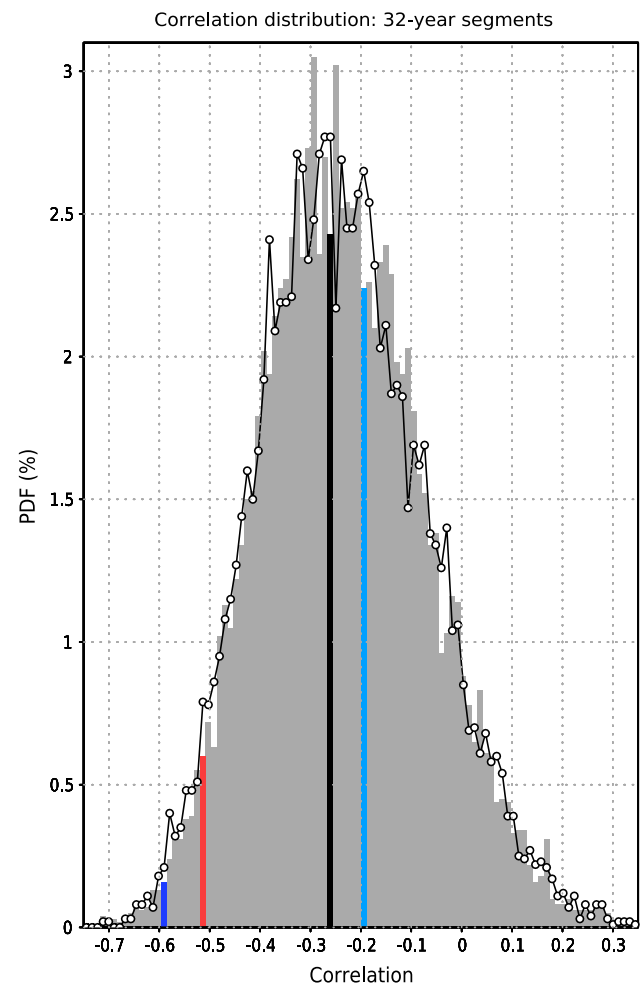


**Fig. 7** 21-Year (black) and 32-year (green) sliding correlation between JJAS AIR and NINO3 index

of this magnitude is a somewhat unusual occurrence, it falls comfortably within the range expected from sampling. This distribution is almost identical to the one produced by taking 32-year samples from two separate white noise processes (Fig. 9), consistent with G2001. Thus, our results demonstrate that observed changes in the ENSO–monsoon relationship can be plausibly explained by random sampling of internal climate variability in the absence of any changes in the background state.

#### 4 Discussion

In this work we investigate the impact of sampling and internal variability on the diagnosed relationship between all-India rainfall and ENSO in a large suite of hindcasts made using a world-class seasonal forecast model as part of *Project Minerva*. We find that we can reproduce the full

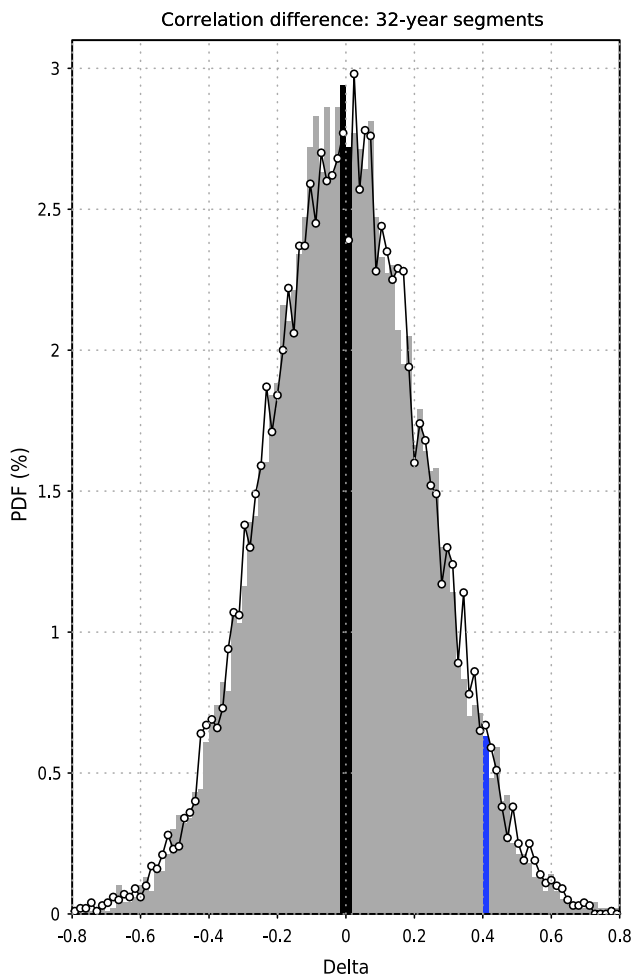


**Fig. 8** Distribution of AIR–NINO3 correlation coefficients. *Black bar* denotes median value from the 10,000 Minerva samples. *Red bar* denotes value from ensemble mean. *Light blue bar* denotes recent observed value, *dark blue bar* denotes observed value with greatest magnitude. *Black line* denotes values from 10,000 samples from a white noise process with a correlation of  $-0.27$

range of correlation coefficients observed since the beginning of the twentieth century simply by sampling different representations of internal climate variability during the period 1980–2011. By calculating the differences between 32-year samples we can also reproduce the full range of observed changes in correlation strength as well.

The results of our experiments indicate that a changing background state is not a necessary condition for producing a wide range of values in the correlation between ENSO and the monsoon, a contribution from climate change in the observed record cannot be ruled out using this approach. We further show that, consistent with previous studies, the distribution of correlation coefficients from the Minerva samples can be closely approximated as a white noise





**Fig. 9** Distribution of change in AIR–NINO3 correlation coefficients. *Black bar* denotes median value from the 10,000 Minerva samples. *Dark blue bar* denotes observed value with greatest magnitude. *Black line* denotes values drawn from 10,000 samples of a white noise process with a correlation of  $-0.27$

process. Our results strongly support the hypothesis that sampling variability, rather than a changing background state, is responsible for the observed fluctuations in the AIR–NINO3 correlation during the past hundred years.

**Acknowledgments** Funding of COLA for this study is provided by Grants from NSF (AGS-1338427), NOAA (NA09OAR4310058 and NA14OAR4310160), and NASA (NNX14AM19G). We would like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR’s Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

**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

- Ashok K, Guan Z, Saji NH, Yamagata T (2004) Individual and combined influences of ENSO and the Indian Ocean dipole on the Indian summer monsoon. *J Clim* 17:3141–3155
- Blanford MF (1884) On the connections of Himalayan snowfall with dry winds and seasons of drought in India. *Proc R Soc Lond* 34:3
- Chang CP, Harr P, Ju J (2001) Possible roles of Atlantic circulations on the weakening Indian Monsoon Rainfall–ENSO relationship. *J Clim* 14:2376–2380
- Charney JG, Shukla J (1981) Predictability of monsoons. Cambridge University Press, Cambridge, pp 99–109
- DelSole T, Shukla J (2012) Climate models produce skillful predictions of Indian summer monsoon rainfall. *GRL* 39:L09703. doi:10.1029/2012GRL051279
- Gershunov A, Schneider N, Barnett T (2001) Low-frequency modulation of the ENSO–Indian monsoon rainfall relationship: signal or noise? *J Clim* 14:2486–2492
- Kinter JL, Miyakoda K, Yang S (2002) Recent change in the connection from the Asian monsoon to ENSO. *J Clim* 15:1203–1215
- Kucharski F et al (2009) A Gill–Matsuno-type mechanism explains the tropical Atlantic influence on African and Indian monsoon rainfall. *Q J R Meteorol Soc* 135:569–579
- Kumar KK, Rajagopalan B, Cane M (1999) On the weakening relationship between the Indian monsoon and ENSO. *Science* 25:2156–2159
- Kumar KK, Rajagopalan B, Hoerling M, Bates G, Cane M (2006) Unraveling the mystery of Indian monsoon failure during El Niño. *Science* 314:115–119
- Manganello JV, Hodges KI, Cash BA, Kinter JL, Altshuler EL, Fennessy MJ, Vitart F, Molteni F, Towers P (2016) Seasonal forecasts of tropical cyclone activity in a high atmospheric resolution coupled prediction system. *J Clim* 29:1179–1200
- Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitart F (2011) The new ECMWF seasonal forecast system (System 4). ECMWF Technical Memorandum, No. 656
- Pai DS et al (2013) Development and analysis of a new high spatial resolution ( $0.25 \times 0.25$ ) long period (1901–2010) daily gridded rainfall dataset over India. NCC Research Report No. 1/2013, National Climate Centre, India Meteorological Department, Pune, p 70
- Rasmusson EM, Carpenter TH (1983) The relationship between eastern equatorial Pacific sea surface temperatures and rainfall over India and Sri Lanka. *Mon Weather Rev* 111:517–528. doi:10.1175/1520-0493(1983)111<0517:TRBEEP>2.0.CO;2
- Rayner NA et al (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res*. doi:10.1029/2002JD002670
- Walker GT, Bliss EW (1932) World weather V. *Q J R Meteorol Soc* 4:53–84
- Wu Z, Li J, Jiang Z, He J, Zhu X (2012) Possible effects of the North Atlantic Oscillation on the strengthening relationship between the East Asian summer monsoon and ENSO. *Int J Climatol* 32:794–800. doi:10.1002/joc.2309
- Zhu J et al (2015) ENSO prediction in Project Minerva: sensitivity to atmospheric horizontal resolution and ensemble size. *J Clim* 28:2080–2095