

Response to comments

Editor comments

1. The language has been pointed as a weakness of the paper and I strongly urge the authors to perform a thorough spell and grammar check before it can be published. If this is not carried out I will not be able to suggest the paper for publication.

Thanks for your suggestion. We have carefully checked the revised manuscript for grammatical errors.

2. Please provide references to statements on for example farmers committing suicide.

We have removed the sentence from the revised manuscript as it was based on just media news.

3. P2, L3. Please change “reason” to “driver”

Done

4. You use the term lead time incorrectly. You speak of “7, 15, 30 and 45 days lead times”, but you mean accumulation periods or accumulation times. Lead time is the time at which a forecast is compared with observations, the day in which it verifies. If you accumulate a forecast from day 1 to 7 you are get the accumulated forecast for day 7, which strictly speaking is not the same as the lead time of 7 days. I would suggest that you change lead time to accumulation period or accumulation time throughout the manuscript.

Thanks. We have replaced the term ‘lead time’ with ‘accumulation period’ as per your suggestion.

5. Generally, sentences like:

“Fig. 4 shows spatial pattern of CSI averaged overall forecast dates for precipitation, daily mean temperature, runoff, and soil moisture for 45 day lead time.” and

“Figure 5 shows anomalies of total runoff and root-zone soil moisture predicted on 15th July 2009 for 45 day lead time using the VIC model with the bias corrected IITM-ensemble forecast, which were compared against the observed anomalies (the VIC model was forced with the observed data)”

are superfluous and should be deleted. This information should be fully contained in the figure caption.

Thanks. We have modified the sentences as suggested.

Interactive comment on “Short-to-medium range hydrologic forecast to manage water and agricultural resources in India” by Reepal Shah et al.

S. Prakash (Referee)

sprakash@citytech.cuny.edu

Received and published: 1 November 2016

The manuscript entitled “Short-to-medium range hydrologic forecast to manage water and agricultural resources in India” by R. Shah, A. K. Sahai, and V. Mishra evaluates precipitation and air temperature reforecasts/forecasts from CFSv2, GEFSv2 and IITM models over India for the period of 2001 to 2009. The evaluation is performed at lead times ranging from 7 to 45 days for the southwest monsoon season. The focus of this study is to assess the performance of operational numerical model forecasts for water resources and agricultural practices in India. The topic of research is of broader interest and vital in Indian perspective. The authors have also used bias correction method to precipitation and air temperature forecasts and integrated them in VIC model to assess total runoff and soil moisture. Before getting to my specific comments, I hope that my inputs are not taken as criticisms, but as constructive suggestions.

Thanks. We appreciate your constructive comments and suggestions.

Specific comments: 1. The improvement after bias-correction should be explicitly mentioned quantitatively in “Abstract” and “Conclusion” sections.

Thank you. We have added following text in the Abstract:

“Bias corrected precipitation forecast showed an improvement of 2.1 mm (on all-India median MAE) while bias corrected temperature forecast was improved by 2.1°C for 45 days accumulation period” on page # 1 in lines#17-18.

The following text has been added in the Conclusions:

“Bias correction of precipitation and air temperatures resulted an improvement of about 2.1 mm and 2.1 °C, respectively in all-India median mean absolute error. Total runoff and root-zone soil moisture forecasts obtained using the corrected IITM-ensemble showed higher skill as compared to the CFSv2 and raw IITM-ensemble for accumulation period up to 45 days. We found that all-India median CSI for runoff forecast was improved from 0.63 to 0.71 after bias correction while CSI of soil moisture forecast was improved from 0.6 to 0.67 for 45 days accumulation period” on page #13 in lines #23-27.

2. Please mention the spatial resolutions of the IITM forecast products as well.

We have included “Forecast ensemble members from IITM are available at 1° resolution” on page# 4 in line # 12

3. The spatial resolution of IMD gridded air temperature is 0.5_ and also all the model products are available at coarser spatial resolution. But, the assessment is performed at finer spatial resolution of 0.25. It is suggested to discuss about the propagation of errors due to resampling from coarser to finer spatial resolution with at least one example.

Thanks. we added the following text (page #4, Lines #29-30):

“We, however, carefully evaluated all the products at their original spatial resolution and at 0.25 ° to make sure that datasets are consistent at both resolutions for spatial and temporal variability. We found that the bias in the forecast products at coarser and higher resolution was consistent.”

4. The use of mean absolute error (MAE) alone for error quantification might be misleading eventually (Ref: Chai and Draxler, 2014, Geosci. Model Dev., 7, 1247-1250). The use of any normalized error metric would be more appropriate to better understand the error characteristics.

We compared Normalized RMSE and MAE for lead-7 & 15 days to find errors in different products as shown in figure below. Spatial patterns and overall results obtained from both the matrices were similar. Therefore, we kept MAE in the revised manuscript for further discussion.

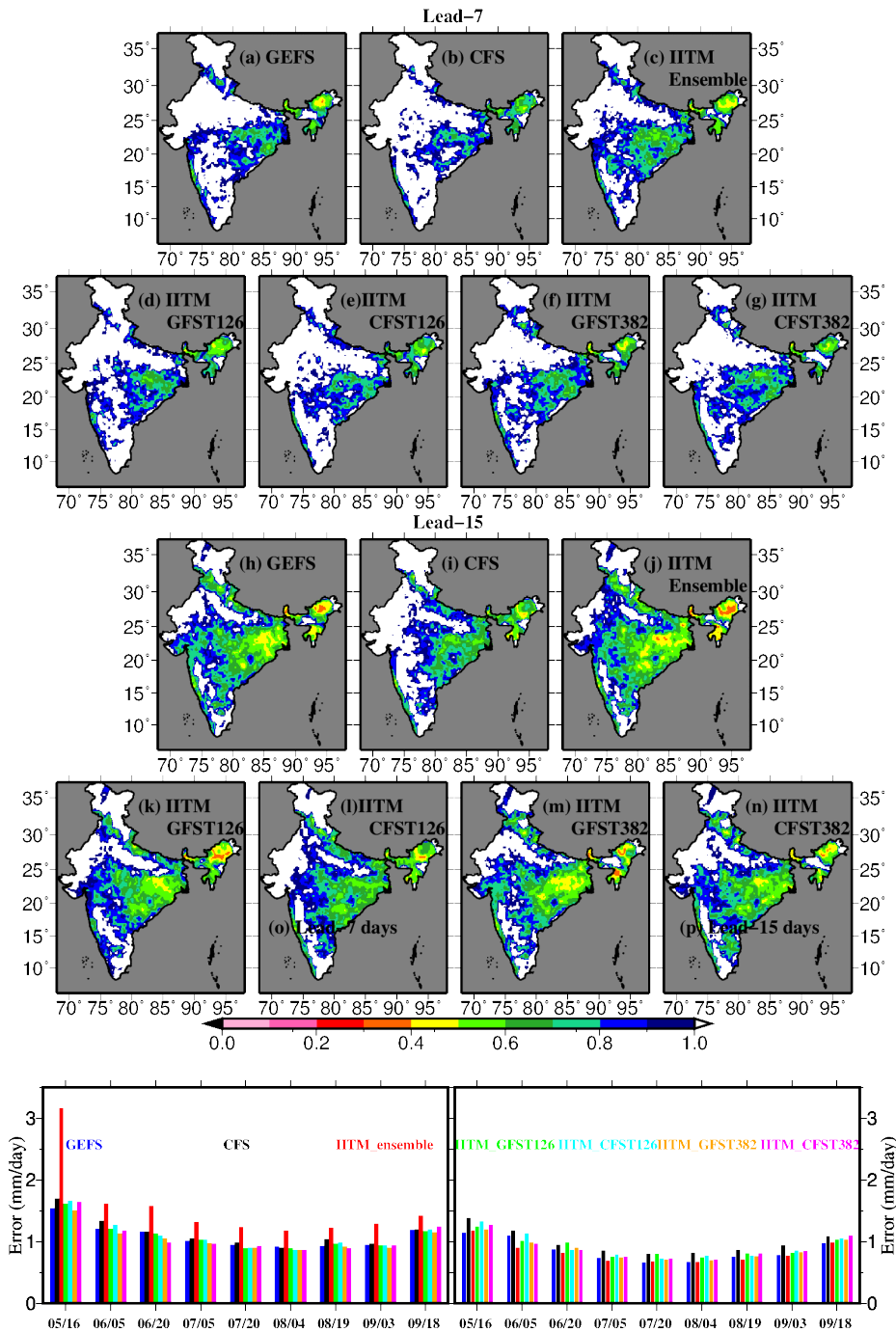


Figure 1: Normalized RMSE in precipitation forecast as compared to observed (OBS) precipitation. (a) Error in precipitation forecast accumulated up to 7-days from GEFSv2 as compared to OBS, (b) same as (a) but with CFSv2 (c) same as (a) but with IITM (multimodel, multiresolution) ensemble (d) same as (a) but with IITM GFST126, (e) same as (a) but with IITM CFST126, (f) same as (a) but with IITM CFST382, (g) same as (a) but with IITM GFST382. (h-n) same as (a-g) but for lead 15 days. (o) area-weighted error in different forecast accumulated up to 7-days as compared to OBS for forecast initiated on different monsoon season dates (p) same as (o) but for lead 15 days. (Period: 2001-2009).

5. Again, CSI is not an equitable categorical metric to evaluate the performance of any numerical model. It is surprising why authors selected CSI alone for this study, even though several better skill metrics are now well-documented.

Thanks. We used Equitable Threat score (ETS) as given below where a_c represents number of forecast events captured by chance.

$$ETS = (hit - a_r) / (hit + miss + false - a_r)$$

Where $a_r = (hit + miss) * (hit + false) / n$, n is sample size.

ETS estimated for CFSv2, IITM ensemble, and bias-corrected IITM ensemble mean's performance in capturing dry anomalies is shown in Figure below.

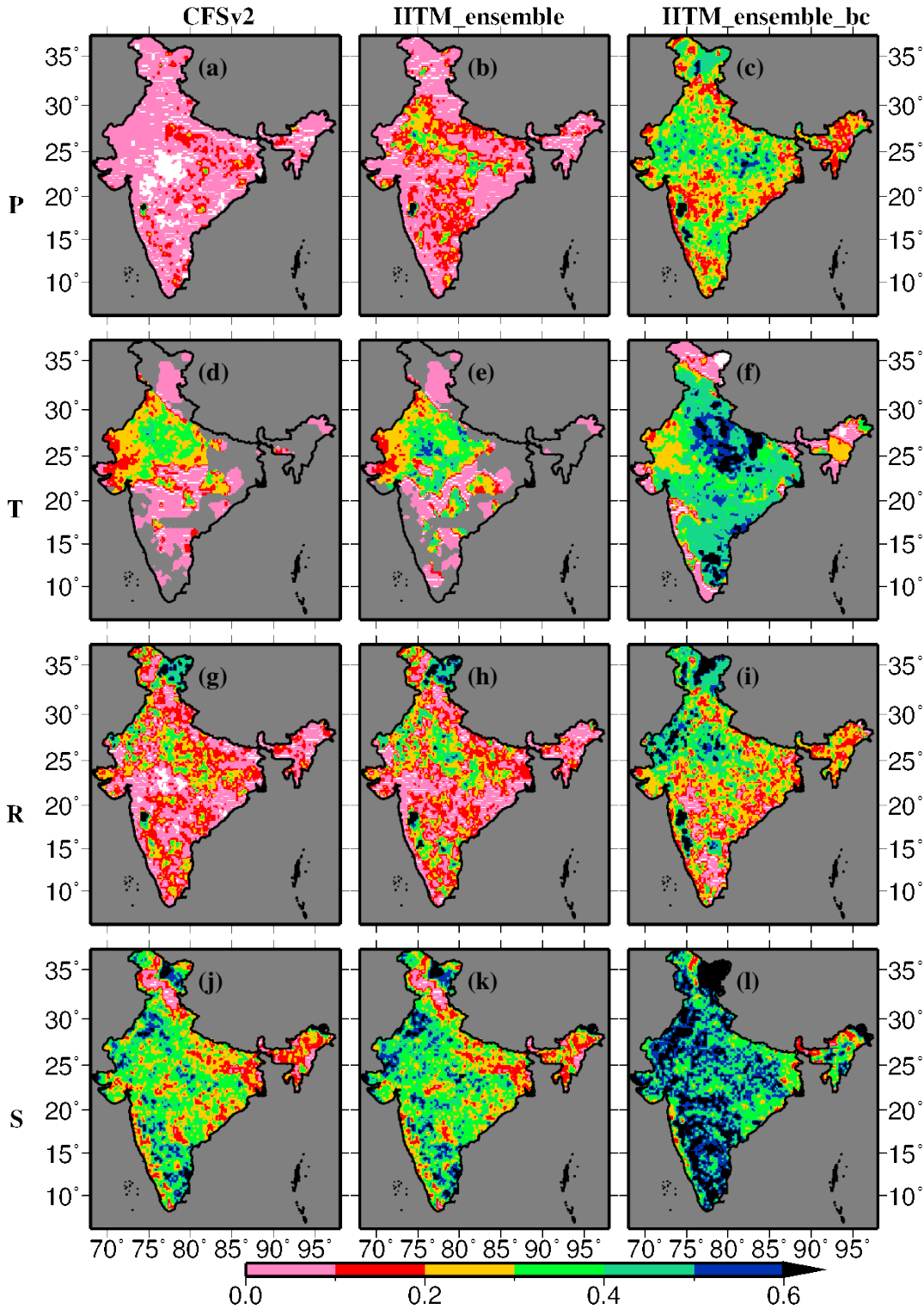


Figure 2 Equitable Threat Score (ETS, averaged for forecast dates) of predicting precipitation (a-c), temperature (d-f), runoff (g-i), and soil moisture (j-l) anomalies with respect to the observed anomalies for CSFv2, IITM-ensemble, and bias corrected IITM-ensemble (IITM-ensemble_bc).

The spatial pattern shown by ETS is similar to that of CSI (Figure 4). This supports our finding that bias-corrected IITM ensemble performs better in capturing dry anomalies. Hence, we used CSI for our analysis.

6. It is suggested to discuss about the impact of sample size at significance of the evaluation in the “Conclusion” section.

We have discussed the influence of sample size in the revised manuscript and added the following text:

“One of the limitations of evaluation of the products in this study is small sample size. The evaluation of all the forecast products was based on 10 common years and 9 forecast dates during the monsoon season. Increasing the sample size in future based on the availability of forecasts for longer period may further improve evaluation and the bias correction.” on page #13 lines #9-11.

7. The authors have appreciably used VIC model here to assess one of the droughts in India. Better prediction of floods is also equally important during the monsoon in India. It would be great if authors demonstrate the same for one flood case too.

The focus of the present study was to evaluate hydrologic prediction for drought assessment (page #2, line#18-20), which has been mentioned in the last paragraph of Introduction. Assessment of floods is an important topic, which requires significant work related to observed data collection and development of robust routing models. Therefore, that will be considered in a separate manuscript.

8. A careful language check is recommended. For instance, first sentence of page 2 needs to be re-written. Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-504, 2016

The manuscript has been carefully checked for possible errors related to grammar.

Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-504-RC2, 2016
© Author(s) 2016. CC-BY 3.0 License.

Interactive comment on “Short-to-medium range hydrologic forecast to manage water and agricultural resources in India” by Reepal Shah et al.

Anonymous Referee #2

Received and published: 14 November 2016

The analysis contributes to a very important problem in Hydro climatology of Indian subcontinent and provides very useful information towards creating an operational subseasonal hydro-meteorological forecasts. The results show a distinct improvement by the IITM forecasts over the NCEP version of CFS 2.0. I have few minor comments, which the authors may address:

1. The authors may highlight, what are the reasons behind such improvements by the IITM model over NCEP CFS v2.0. This should come with some bullet points clearly highlighting the need for any model to be successfully applied for monsoon forecasts.

We appreciate your suggestion and added the following text in the revised manuscript (page #13, lines #14-18)

“The major factors that might have contributed in the improvements in the IITM forecast are:

- i. Ensemble members of IITM forecast are generated by perturbing initial atmospheric conditions to improve simulation of the northward propagation*
- ii. Improvements in the boundary conditions with bias corrected SST result in an improved precipitation prediction*
- iii. Higher spatial resolution of the IITM forecast can better resolve orographic rainfall”*

2. Please, provide some details on the lead-time dependant bias correction. Can this be applied to the CFS2.0 forecasts of precipitation?

Thanks for your valuable suggestion. We have provided following (page #11 Lines #25-28)

“However, the bias in the forecast products may have temporal variability and may not be constant for the entire period of 45 days. Therefore, bias correction approaches based on the variable lead time (Stockdale, 1997) need to be evaluated in future when IITM forecast for long-term retrospective period is available. The bias correction approach that we presented can be applied to evaluate seasonal forecast skill.”

3. I could not understand the sources of the observed soil moisture and runoff data. The authors may mention the same or they may provide a table on the details of the data used with their sources. This will help others to reproduce the results and validate the same.

Thanks. The VIC model was calibrated and evaluated using observed streamflow and satellite soil moisture and evapotranspiration (Shah and Mishra, 2016a,b). In this study, we used calibrated VIC model forced with observed IMD data to simulate soil moisture and runoff, which was considered as a reference to evaluate the forecast of soil moisture and runoff.

We have provided following (page#11 in lines#30-35)

“The VIC model was calibrated and evaluated using observed streamflow and satellite soil moisture and evapotranspiration (Shah and Mishra, 2016a and Shah and Mishra, 2016b).In this study, we used calibrated VIC model forced with observed IMD data to simulate soil moisture and runoff, which was considered as a reference to evaluate the forecast of soil moisture and runoff. Forecast of root-zone soil moisture and runoff was simulated using the VIC model forced with the forecast products (IITM-ensemble-bc, IITM-ensemble, and CFSv2), which were evaluated against the soil moisture/runoff obtained from the VIC model simulation using

the observed forcing from IMD (Supplemental Fig. S17)."

4. During the low rainfall periods, the human intervention is quite high in terms of irrigation. To the best of my knowledge, VIC does not have the capability of doing the same in a way that is applicable to Indian condition. I do not really blame the authors for the same as there is as such no way out, given the status of latest version of VIC. But this should be explicitly mentioned as limitation.

We appreciate your valuable comment. The present study considers only the monsoon season for analysis, however, we acknowledge that the role of irrigation during the dry season.

We have mentioned this limitation and added the following text (page #6 Lines 8-11) in the revised manuscript:

"The VIC model's version that was used in this study does not explicitly represent groundwater, rather it only accounts for baseflow. We acknowledge that India specific soil and vegetation parameters along with the representation of irrigation, reservoir, and groundwater can improve the water budget; however, these were not considered in the present study due to unavailability of either observations or the model version that has the representation of human interventions."

5. Similarly, the crop parameters, which are used in VIC are mostly based on Maize and Soyabean and this is different from Indian crop conditions. The authors may correct me if I am wrong. If I am correct, this should also be mentioned as a limitation. VIC also have limitation of not having a good ground water model. This should also come as a limitation.

Thanks. The point is valid.

We included the following text to address this (page #6, Line 4-8):

"The vegetation parameters used in this study were developed using 1-km Advanced Very High Resolution Radiometer (AVHRR) global land cover information. We used vegetation library that was developed at University of Washington. The vegetation parameters were not specifically developed to incorporate crops that are grown in India. However, the existing parameters were successfully used in the model application over India (Shah and Mishra, 2015; Shah and Mishra, 2016)."

6. Is the model calibrated or does it consider the recommended values of parameters of VIC from global data set? The authors may also publish the sensitive parameter values for VIC as supplementary dataset so that the readers will be able to reproduce and apply the work.

We appreciate your suggestion. However, the model calibration and evaluation is a part of the previously published manuscript (Shah and Mishra (2015, 2016) and Shah and Mishra (2016)), therefore, we are unable to provide the calibrated values of parameters in this study. Authors will be happy to share the parameters and data, if someone is interested.

We have mentioned following text in the revised manuscript (page #6 lines 11-13):

"The VIC model setup used in this study is well calibrated and evaluated against observed streamflow and satellite based evapotranspiration and soil moisture in Shah and Mishra (2016a) and Shah and Mishra (2016b)."

Short- to-medium-range Sub-Seasonal hydrologic forecast to manage water and agricultural resources in India

Reepal Shah¹, A. K. Sahai², Vimal Mishra¹

5 ¹Civil Engineering, Indian Institute of Technology (IIT) Gandhinagar and ITRA Project: Measurement to Management (M2M): Improved Water Use Efficiency and Agricultural Productivity through Experimental Sensor Network.

²Indian Institute of Tropical Meteorology (IITM) Pune, India

Correspondence to: Vimal Mishra (vmishra@iitgn.ac.in)

Abstract.

10 Water resources and agriculture are often affected by the weather anomalies in India resulting in a disproportionate damage. While short to ~~medium-range~~sub-seasonal prediction systems and forecast products are available, a skilful hydrologic forecast of runoff and root-zone soil moisture that can provide timely information has been lacking in India. Using precipitation and air temperature forecasts from the Climate Forecast System v2 (CFSv2), Global Ensemble Forecast System (GEFSv2) and four products from Indian Institute of Tropical Meteorology (IITM), here we show that the IITM ensemble mean (mean of all four products from IITM) can be used operationally to provide hydrologic forecast in India at 7-45 days
15 lead-accumulation period. The IITM ensemble mean forecast was further improved using bias correction for precipitation and air temperature. ~~Forecast based on the IITM ensemble mean showed better skills in majority of India for all the lead times (7-45 days) in comparison to the other forecast products.~~Bias corrected precipitation forecast showed an improvement of 2.1 mm (on all-India median MAE) while all-India median bias corrected temperature forecast was improved by 2.1°C for
20 45 days accumulation period. Moreover, the VIC simulated forecast of runoff and soil moisture successfully captured the observed anomalies during the severe ~~droughts~~drought years. The findings reported herein have strong implications for providing timely information that can help farmers and water managers in decision making in India.

1. Introduction

25 ~~Droughts in India have enormous implications for water resources and agriculture (Mishra *et al* 2014, Shah and Mishra 2015). Many regions in India face drought risks due to lack of monsoon season rainfall. In 2015, a large part of India was under drought which affected agriculture and water resources (Times of India, 18 June 2016). Moreover, in 2015, about 33 million people were affected by the drought that covered 256 districts, 10 states, and caused an estimated loss of 650,000 crore Indian rupee (Indian Express, 11 May, 2016). Drought in 2015 had socio-economic effects on people as agriculture based industries were closed, water was supplied using trains, and at a few places farmers committed suicide. The major driver of hydrological (based on runoff) or agricultural (based on soil moisture) droughts in India remains the Indian summer monsoon (Mishra *et al* 2014, Shah and Mishra 2015), which accounts for about 80% of the mean annual rainfall and has 10%~~

year-to-year variability (Rajeevan *et al* 2005, 2006, Rahman *et al* 2009). However, during the recent decades, increased air temperature has affected hydrologic and agricultural droughts in India (Mishra *et al* 2014) and in the other parts of the World (Dai *et al* 2004, Park Williams *et al* 2012, Shukla *et al* 2015, Livneh and Hoerling 2016).

Droughts in India have enormous implications for water resources and agriculture (Mishra *et al.*, 2014; Shah and Mishra, 2015). Many regions in India face drought risks due to lack of monsoon season rainfall. In 2015, a large part of India was under drought which affected agriculture and water resources (Mishra *et al.*, 2016). Moreover, in 2015, about 33 million people were affected by the drought that covered 256 districts, 10 states, and caused an estimated loss of 650,000 crore Indian rupee (Indian Express, 11 May, 2016). The major driver of hydrological (based on runoff) or agricultural (based on soil moisture) droughts in India remains the Indian summer monsoon (Mishra *et al.*, 2014; Shah and Mishra, 2015; Mishra *et al.*, 2016), which accounts for about 80% of the mean annual rainfall and has 10% year-to-year variability (Rahman *et al.*, 2009; Rajeevan *et al.*, 2005, 2006). However, during the recent decades, increased air temperature has affected hydrologic and agricultural droughts in many regions of the World (Dai *et al.*, 2004; Livneh and Hoerling, 2016; Park Williams *et al.*, 2012; Shukla *et al.*, 2015).

One of the relatively well known reasons/drivers of drought/drought occurrence in India is the positive sea surface temperature anomaly (El Nino) in the Pacific Ocean (Kumar *et al* 1999, 2006)(Kumar *et al.*, 1999, 2006) and in the Indian Ocean (Mishra *et al* 2012, Roxy *et al* 2015). However, in the absence of hydrologic forecast at appropriate lead (Mishra *et al.*, 2012; Roxy *et al.*, 2015). However, in the absence of hydrologic forecast at appropriate lead time, planning of agricultural and water resources sectors are often adversely affected. For instance, many times the cost of seeds, field preparation, and transplantation ~~can not~~ cannot be recovered due to prolonged anomalies of soil moisture or rainfall. Furthermore, water resources, reservoir operations, and irrigation planning is affected in the absence of a skilful forecast at sufficient lead time. Prediction of anomalies in meteorological and hydrological conditions well in advance can assist timely decision-making to minimize impact on agricultural and water resources sectors. Shah and Mishra (2016b)-showed potential of the Global Ensemble Forecast System (GEFS; Hamill *et al* 2013) for hydrologic predictions in India with a lead time up to 7 days. They reported that up to 7 days lead, major skills in hydrologic prediction are derived from initial hydrologic conditions (i.e.(GEFS; Hamill *et al.*, 2013) for hydrologic prediction in India with a lead time up to 7 days. They reported that up to 7-days lead time, major skill in hydrologic prediction is derived from initial hydrologic conditions (i.e. initial soil moisture content) as shown in Shukla and Lettenmaier (2011). Yuan *et al.*(2011) reported that soil moisture forecast from the CFSv2 (Saha *et al* 2014) provides useful information to predict droughts in the tropical region.(CFSv2; Saha *et al.*, 2014) provides useful information to predict droughts in the tropical region. Moreover, Yuan *et al* (2012a) showed that the CFSv2 can provide a better seasonal hydroclimatic forecast than ensemble streamflow prediction in the USA.

Despite the utility of the various forecast products that can provide useful skill in hydrologic predictions, efforts have largely been limited to evaluate the potential of these products to provide forecast at 7-45 days ~~lead-time~~accumulation period that can be used for agricultural and water resources planning in India. Here we provide an assessment of ~~skills~~skill in hydrologic forecast ~~that can be utilized for drought forecast~~ at 7-45 days ~~lead-time~~accumulation period using data from GEFSv2, CFSv2, and IITM to improve management of water and agricultural resources in India.

2. Data and Methodology

2.1 Observed data

Forecast products were evaluated against observed data from India Meteorological Department (IMD). We used the 0.25° daily gridded precipitation product from IMD which was developed based on ground observations from 6995 stations across India using an inverse distance weighing scheme (Shepard 1984) and is available for the period of 1901-2015 (Pai *et al* 2015). The IMD precipitation captures spatial variability of the monsoon season rainfall and features related to orographic rainfall in the Western Ghats and foothills of Himalaya. We used 0.5° daily observed maximum and minimum temperatures from IMD, which were developed based on 395 stations across India (Srivastava *et al* 2009). Gridded air temperature dataset is available for the period of 1951-2013 and have been used in many previous studies (Mishra *et al* 2014, Shah and Mishra 2014, 2015, 2016).

2. Data and Methodology

2.1 Observed data

Forecast products were evaluated against observed data from India Meteorological Department (IMD). We used the 0.25° daily gridded precipitation product from IMD which was developed based on ground observations from 6995 stations across India using an inverse distance weighing scheme (Shepard, 1984) and is available for the period of 1901-2015 (Pai *et al.*, 2015). The IMD precipitation captures spatial variability of the monsoon season rainfall and features related to orographic rainfall in the Western Ghats and foothills of Himalaya. We used 0.5° daily observed maximum and minimum temperatures from IMD, which were developed based on 395 stations across India (Srivastava *et al.*, 2009). Gridded air temperature dataset is available for 1951-2013 and have been used in many previous studies (Mishra *et al.*, 2014; Shah and Mishra, 2015, 2016b, 2014; Mishra *et al.* 2016).

2.2 Forecast Products

We evaluated prediction ~~skills~~skill of precipitation, maximum and minimum temperatures from the CFSv2 reforecast (Saha *et al* 2014)(Saha *et al.*, 2014), GEFSv2 reforecast (Hamill *et al* 2013)(Hamill *et al.*, 2013) and forecast products from IITM. Reforecast from the CFSv2 are based on a dynamical coupled model and are available at every 5th day from the start of year from the National ~~Center~~Centre of Environmental Prediction (NCEP). Moreover, 6-hourly forecasts at every 5th day from

CFSv2 are available with up to nine months lead time and at 1° resolution for ~~the period of 1982-2009 from the CFSv2.~~ Climate forecast System (CFS) model's atmospheric component is operational at T126 spectral truncation (~100 Km horizontal resolution) and 64 sigma-pressure hybrid vertical resolution. Shukla and Lettenmaier (2014)(2011) using CFSv2 reported that initial hydrologic conditions dominate skill of hydrologic prediction in the continental United States (CONUS) up to 1-month lead time, beyond which skill from meteorological ~~forcings~~ forcing dominated. McEvoy *et al* (2016)McEvoy et al. (2016) recently demonstrated higher skill for potential evapotranspiration than precipitation using the CFSv2. Moreover, Yuan et al. (2014)(2011) reported that ~~the~~ CFSv2 performs better than CFSv1 for prediction of precipitation and air temperature in the United States. Mo and Lettenmaier (2014)(2014) found that for shorter lead times (about 1 month) ~~the~~ CFSv2 forecast has higher skill for soil moisture prediction than the benchmark forecast (climatological mean). Moreover, Tian et al. (2016) evaluated ~~the~~ CFSv2 for the CONUS and found that extreme indices based on temperature were better predicted than that of precipitation.

Other than CFSv2, we compared precipitation and temperature forecast from ~~the~~ GEFSv2 reforecast (Hamill *et al* 2013)(Hamill et al., 2013), which is based on the Global Forecast System (GFS) model, for 7 and 15 days lead time. Ensemble members are generated in GEFS by making perturbations in initial atmospheric conditions which leads lead to 11 ensemble members. The GEFS model runs at T254L42 resolution (~40 Km horizontal resolution) for the first 8 days lead time and at T190 (~54 km) for lead time beyond 7.5 days. The GEFS reforecast are available at 1° resolution for lead time up to 16 days and at 0.5° for 8 days lead from 1985 to present. Shah and Mishra (2016b)(2016b) evaluated skill of ~~the~~ GEFSv2 reforecast for drought prediction in India up to lead 7 for accumulation period of 7 days and found that the GEFS reforecast showed correlation of more than 0.5 against drought estimates from the observed data.

We obtained four forecast products from IITM. The forecast products of IITM are generated from the same CFSv2 model that has been described above. Abhilash et. al. (2014) ~~have~~ developed an ensemble prediction system using CFSv2 at T126 horizontal resolution (~100km) [called as hereafter: IITM-CFST126] for prediction of monsoon intraseasonal oscillations (MISO) over the Indian monsoon region 15-20 days in advance. They found that though the skill was reasonable, ~~but~~ there was a significant dry bias over the Indian land. Sharmila *et al.* (2013)(2013) reported that CFSv2 simulates the northward propagating MISO reasonably well but it has cold bias in SST and tropospheric temperatures. Thus, Abhilash et al (2013)(2013) implemented a lead time dependent SST bias correction and forced the GFS (Atmospheric component of CFSv2) with slightly different physics and showed that it has improved skill over Indian Land India compared to the CFSv2 (called as hereafter: IITM-GFST126). Subsequently, Sahai et. al. (2015a) ~~have also~~ (2015a) implemented a high resolution version of CFSv2 (at T382 horizontal resolution ~35km; called as hereafter: IITM-CFST382) and showed that it has better skill in steep orographic regions. Although these three individual models show similar prediction skill and their errors saturate at about the same lead time of around 25 days, there are many instances where the three models disagree in predicting particular events, such as the amplitude and phase of monsoon intraseasonal oscillation (MISO) propagation.

Considering these facts, Abhilash et al. (2015) have (2015) proposed a CFS based multimodel ensemble mean (MME), which improved the spread errors spread error relationship and added value to both the deterministic and probabilistic forecasts. Real time skills skill for these models have has been reported in various communications. (Sahai et al 2013, Borah et al 2015, Joseph et al 2015b, 2015a, Sahai et al 2015b) the previous studies (Borah et al., 2015; Joseph et al., 2015b, 2015a, Sahai et al., 2013, 2015b). Subsequently, bias corrected SST forced GFS was also run at T382 resolution (called as hereafter: IITM-GFST382). Thus IITM's forecasts are available for four models, named IITM-CFST126, IITM-GFST126, IITM-CFST382, and IITM-GFST382, respectively. Model integrations for the year starting since 2001 to 2015 are carried out from 16th May and continued up to 28th September at every 5 day interval (16th May, 21st May, 26th May, ..., 23rd Sep, 28th Sep) for the next 45 days period. Forecast ensemble members from IITM are available at 1° resolution. Ensemble mean of all four IITM products (hereafter: IITM-ensemble) and individual products were compared with the CFSv2 and GEFSv2 to evaluate the hydrologic prediction skill. The aim of this comparison was to evaluate if IITM forecast products provide better prediction skills skill than CFSv2 and GEFSv2. Moreover, the product that provide the best hydrologic prediction skills skill in India can be used operationally to forecast hydrologic conditions and rainfall and temperature anomalies that can help in decision making in agricultural and water resources.

We used ensemble mean (of all available ensemble members) of individual forecast products for evaluation. We selected forecasts at every 15th day and, which was evaluated by accumulating for 7, 15, 30, and 45 day accumulation period using accumulated precipitation and average temperature. We use the term accumulation period instead of lead time as forecast evaluation was performed for accumulated precipitation and mean temperature for 7, 15, 30, and 45 days lead. We selected forecasts starting from 16th May till end of September as currently IITM provides forecast during the monsoon season. However, IITM will extend forecast into the non-monsoon season in near future. We aggregated all the observed and forecast variables (precipitation, maximum and minimum air temperatures) to daily-scale (if they were available at sub-daily time period) and regridded them to 0.25° horizontal resolution to make them consistent with the spatial resolution of observed data. We regridded precipitation and air temperature using Maurer et al.(2002) which uses the Synergraphic Mapping System (SYMAP) algorithm (Shepard 1984) (Shepard, 1984) for precipitation and lapse rate based on elevation data for air temperature. We, however, carefully evaluated all the products at their original spatial resolution and at 0.25° to make sure that datasets are consistent at both resolutions for spatial and temporal variability. We found that the bias in the forecast products at coarser and higher resolution was consistent.

We considered a common period of 2001-2009 for comparison and evaluation of different forecast products against the observed gridded data from IMD.

2.3 Forecast Evaluation

For evaluation of the forecast ~~from each product against the observations~~, we prepared yearly time-series of precipitation and temperature forecast for each forecast ~~dates; for lead 7, 15, 30, and 45 days; from each products and corresponding OBS; and for each grids,date~~ by accumulating precipitation and averaging temperature for a given lead time- ~~(7-45 days)~~. For instance, if the date of forecast was 1st June and lead time 15 days, accumulated precipitation and mean temperature for 15 days from June 1st for each of the products were estimated for the period 2001-2009. As the period for evaluation was 2001-2009, sample size was 10, and we acknowledge that a larger sample size with data for a longer retrospective record will help us to better categorize uncertainty in forecast ~~skillsskill~~. We used coefficient of correlation, mean absolute error (MAE), and critical success index (CSI) to evaluate performance of the forecast products. A non-parametric Spearman Rank Correlation coefficient ~~(Wilks 2006)(Wilks, 2006)~~ was used to evaluate performance of forecast products in capturing temporal relationship with OBS. For this the forecast product and corresponding OBS are assigned ranks and then correlation was estimated using following equation (2.1)

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

(2.1)

Where ~~r_s~~ r_s is Spearman Rank Correlation coefficient; ~~d_i~~ d_i is difference in rank between paired forecast and OBS; and ~~n~~ n is sample size (here 10). Significance of correlation was tested using the exact permutation distribution test ~~(Robson 2002)(Robson, 2002)~~. Observed samples were permuted and rank correlations were estimated. Estimated correlation is significant if it rejects the null hypothesis at 5% significance level.

Mean absolute error (MAE) was used to estimate error in the forecast products as compared to OBS. Absolute error was estimated in all the forecast products for each year as compared to OBS and then mean of all years was taken to estimate MAE. ~~Critical success index (CSI; Wilks 2006) was used to evaluate anomalies predicted using forecast products as compared to OBS. CSI is ratio of hit (i.e. forecast predicts OBS dry anomaly;) and sum of hit, miss (i.e. forecast misses OBS dry anomaly), and false (i.e. forecasts predicts dry anomaly whereas OBS don't; CSI= hit/(hit+miss+false)).~~ Critical success index (CSI; (Wilks, 2006)) was used to evaluate anomalies predicted using forecast products as compared to OBS similar to (AghaKouchak and Mehran, 2013). CSI is ratio of hit events and sum of hit and miss, and false events (hit+miss+false).

2.4 The Variable Infiltration Capacity (VIC) model

We used the Variable Infiltration Capacity (VIC, version 4.1.2) (Liang *et al* 1994, 1996) model to simulate hydrologic variables (total runoff and soil moisture) using forcing (daily precipitation, and maximum and minimum temperatures) from IMD and the forecast products. The VIC model simulates water and energy fluxes at each grid cell and sub-grid variability of precipitation, elevation, soil, and vegetation is well represented (Gao *et al* 2010). The VIC model setup used in this study

Formatted: Font: Not Bold

Formatted: Right, Line spacing: 1.5 lines

is well calibrated and evaluated against observed streamflow and satellite based evapotranspiration and soil moisture in Shah and Mishra (2016a) and Shah and Mishra (2016b). The soil and vegetation parameters used in this study are described in Shah and Mishra (2015). The VIC model has been widely used for hydrologic prediction at watershed and regional scales (Shukla and Lettenmaier 2011, Yuan and Wood 2012b, Mo and Lettenmaier 2014, Shah and Mishra 2016b).

We used the Variable Infiltration Capacity (VIC, version 4.1.2) (Liang et al., 1994, 1996) model to simulate hydrologic variables (total runoff and root-zone soil moisture) using meteorological forcing (daily precipitation, and maximum and minimum temperatures) from IMD and the forecast products. Soil moisture and runoff predicted using forecast products were evaluated against soil moisture and runoff simulated using the observed forcing from IMD. The VIC model simulates water and energy fluxes at each grid cell and sub-grid variability of precipitation, elevation, soil, and vegetation is well represented (Gao et al., 2010). The soil parameters used were developed based on Harmonized World Soil Database (HWSD) v1.2. The vegetation parameters used in this study were developed using 1-km Advanced Very High Resolution Radiometer (AVHRR) global land cover information. We used vegetation library that was developed at University of Washington. The vegetation parameters were not specifically developed to incorporate crops that are grown in India. However, the existing parameters were successfully used in the model application over India (Shah and Mishra, 2015; Shah and Mishra, 2016). The VIC model's version that was used in this study does not explicitly represent groundwater, rather it only accounts for baseflow. We acknowledge that India specific soil and vegetation parameters along with the representation of irrigation, reservoir, and groundwater can improve the water budget; however, these were not considered in the present study due to unavailability of either observations or the model version that has the representation of human interventions. The VIC model setup used in this study is well calibrated and evaluated against observed streamflow and satellite based evapotranspiration and soil moisture in Shah and Mishra (2016a) and Shah and Mishra (2016b). The VIC model has been widely used for hydrologic prediction at watershed and regional scales (Mo and Lettenmaier, 2014; Shah and Mishra, 2016b; Shukla and Lettenmaier, 2011; Yuan and Wood, 2012b).

2.5 Bias-correction of Precipitation and Temperature Forecast

We evaluated improvements in hydrologic prediction due to prediction can be achieved by post-processing the forecast of meteorological variables (precipitation, maximum and minimum temperatures) from the selected forecast products. We corrected precipitation forecast using the linear scaling approach as described in Shah and Mishra (2015, 2016b). For each forecast date, we corrected accumulated precipitation for the selected (7, 15, 30 and 45 days) lead accumulation period. We first corrected accumulated precipitation due to extreme events (above 90th percentile) for each forecast date in the training period and a scaling factor was obtained for each forecast dates based on ratio of accumulated precipitation for 45 days lead accumulation period due to extreme events in the observed and forecast product products. In the second step, after the correction for extreme precipitation, scaling factors were obtained based on accumulated precipitation for 45 days lead accumulation period, for each forecast dates from the forecast products and OBS

for the entire training period. Scaling factors were estimated for the training period (Nine years), which were evaluated in the testing period (One year). More detailed information on this method can be obtained from Shah and Mishra (2016b).

Field Code Changed

To correct daily mean (of maximum and minimum) temperature from the forecast, we performed Quantile-Quantile (Q-Q) mapping (Wood 2002). ~~Initially, we prepared yearly time-series of 45-days lead~~(Wood, 2002). ~~Initially, we prepared yearly time-series of 45-days accumulation period~~ average temperature forecast for all the forecast dates along with corresponding observed time-series. For each forecast date and for each grid cell, we estimated quantiles of mean temperature for 45 days ~~lead accumulation period~~ for each year using the climatology of the entire period. To estimate quantiles, cumulative distribution functions (CDF) were fitted. Weibull plotting position was used to map cumulative distribution function when percentiles fall between $1/(N+1)$ and $N/(N+1)$; where N is number of climatological years during the training period. In case when percentiles fall beyond these limits, normal distribution was fitted and values were extrapolated. More details on the Q-Q mapping can be obtained from Shah and Mishra (~~2016b~~)(2016b). Similarly, quantiles were estimated for OBS temperature for corresponding time-series. Based on estimated quantiles, Q-Q mapping was done and forecast was replaced with corresponding value based on OBS. We estimated bias corrected mean temperature using Q-Q mapping. Bias (difference between corrected and uncorrected 45-day average mean temperature) was then added equally to daily raw Tmax and Tmin to get the corrected values of daily maximum and minimum temperatures. We did not bias correct Tmax and Tmin individually as that will affect the ~~daily~~diurnal temperature range (Tmax-Tmin). We adopted multifold validation approach leaving-one-year out for testing both precipitation and mean temperature (Shah and Mishra, 2016b).

Field Code Changed

Forecast of soil moisture and runoff is essential for planning and decision making in agriculture and water resources (Asoka and Mishra, 2015). Hence, we evaluated forecast skill of soil moisture and runoff simulated using meteorological variables from IITM-ensemble. Using the raw and bias corrected forecasts (precipitation, maximum, and minimum temperatures), the Variable Infiltration Capacity (VIC) model was run to obtain soil moisture and total runoff (surface runoff+baseflow) forecast. We evaluated improvements in correlation of runoff and soil moisture predicted using the bias-corrected precipitation and temperatures from the IITM ensemble (IITM-ensemble-bc) against uncorrected (raw) precipitation and temperatures from the IITM ensemble mean (IITM-ensemble) and CFSv2 (Supplemental Fig.S14). For simulating runoff and soil moisture, forcings from all the three products were used to run the VIC model at 0.25° and daily resolution while initial hydrologic conditions were generated using the observed forcing from the IMD. Forecast skill in hydrologic prediction was evaluated for mean total runoff and soil moisture for 7-45 day accumulation period. We considered 45 day accumulation period to evaluate the hydrologic prediction skill as for shorter lead times forecast skill are generally higher owing to persistence in initial hydrologic condition.

3. Results and discussion

3.1 Comparison of Forecast ~~Skills~~**Skill** for Precipitation and Temperature forecast:

3.1.1 Lead ~~time~~**7** and 15-days

We estimated forecast ~~skills~~**skill** (against observations, OBS hereafter) in precipitation and air temperature from all the forecast products for 7, 15, 30, and 45 ~~days lead time~~**day accumulation period**. Hydrologic forecast at these ~~lead times~~**accumulation periods** can be used for planning (field preparation, sowing, irrigation, water management, and reservoir operations) and decision making in water resources and agriculture. ~~To estimate forecast skills, correlation coefficient, mean absolute error (MAE), and critical success index (CSI) were used. Figure 1 shows Spearman Rank correlation between accumulated precipitation forecast and corresponding OBS (IMD) for 7 and 15 days lead for forecasts initiated during the monsoon months of the period 2001-2009. The 2001-2009 was selected for the analysis as forecast from all the products is available for this period.~~ All the forecast products showed significantly high (more than 0.75) **Spearman Rank** correlation (Fig. 1a-n) in the majority of India for ~~lead-accumulation period of~~**7** days indicating higher ~~skills~~**skill** for shorter lead time. We noticed that correlation declines as ~~lead-accumulation period~~**was** increased from 7 to 15-days especially in the central region (Fig. 1). Moreover, we find that ~~the~~-GEFSv2 and IITM-ensemble (correlation more than 0.6 for majority of India) perform better than ~~the~~-CFSv2 for 15 ~~days lead time~~**day accumulation period**. Correlations between observed and forecast were generally lower for forecast initiated during the months of July and August (Fig. 1o-p). Among all the forecast products, IITM products and their IITM-ensemble mean (mean of all four IITM forecast products) showed better correlations with OBS as compared to ~~the~~-GEFSv2 and CFSv2 for 7 and 15 ~~days lead~~**day accumulation period** (Fig. 1 and Supplementary Table S1). Among the IITM products, products with the atmospheric model operating at higher resolution (IITM-CFST382 and IITM-GFST382) showed relatively better performance as compared to the other two IITM products, which demonstrates that the models operating at higher resolution provide a better forecast skill (Duffy et al., 2003; Roebber et al., 2004).

We estimated MAE in precipitation forecast from all the products as compared to OBS for ~~lead-accumulation period of~~**7** and 15 days (Fig. S1). We find that MAE is proportional to ~~the~~ magnitude of precipitation as the monsoon season precipitation is higher in the core monsoon, northeastern, and Western Ghats regions (Fig. S1). Moreover, all the products showed a lower MAE in the arid and semi-arid regions of the western India during the monsoon season and MAE was higher during the months of July-September (Fig. S1o,p). MAE, however, decreases as forecast ~~lead time~~**accumulation period** was increased from 7 to 15 days, which is due to longer accumulation period for precipitation. We noticed that all India median MAE (~~median of all the grids~~) in the forecast products vary with the date of forecast, however, both ~~the~~-CFSv2 and IITM ensemble mean showed comparable MAE at all India scale for 7 day ~~lead-accumulation period~~**accumulation period** (Fig S1o and Table S1). However, for 15 day ~~lead-accumulation period~~**accumulation period**, and for most of the forecast dates (Fig S1p), the IITM-ensemble showed lower error compared to the other products. Overall, based on correlation and MAE, we find that the IITM ensemble performs better than the other forecast products for 7 and 15 day ~~lead-accumulation period~~**accumulation period** for precipitation prediction.

Field Code Changed

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Lower ~~skill~~skill in precipitation forecast in July and August can be attributed to high intraseasonal variability as a large fraction of total precipitation in the monsoon season occurs during these months. Intraseasonal variability can be characterized by spells of active-break periods of length 3-5 days (~~Rajeevan et al 2010~~)(Rajeevan et al., 2010). Active-break spells are dominated by SST, wind pattern, Maiden-Julian Oscillation (MJO), and ITCZ (~~Goswami and Ajayamohan 2000; Woolnough et al 2007; Rajeevan et al 2010~~)(Goswami and Ajayamohan, 2000; Rajeevan et al., 2010; Woolnough et al., 2007). Predictability of precipitation in India depends on the ability of models to capture intraseasonal and interannual variability in precipitation (~~Webster et al 1998~~)(Webster et al., 1998). Improvements in spatial resolution of the atmospheric model and bias corrected SST in the IITM forecast products ~~leads~~lead to enhancement in forecast ~~skill~~skill, which potentially can be used for decision making in water resources and agriculture in India.

Similar to precipitation for 7 and 15-days ~~lead~~accumulation period, we evaluated ~~skill~~skill in maximum (Tmax) and minimum (Tmin) temperatures from all the forecast products against observed air temperatures from IMD (Fig. S2). Tmax averaged for 7 day ~~lead~~accumulation period from all the forecast products showed a good correlation with OBS over the most of India (Fig. S2a-g). Similar to precipitation from the IITM-ensemble, Tmax showed the highest correlation with OBS (0.78; Table S1). However, correlation for 15 days ~~lead~~accumulation period was lower than that of 7 days ~~lead~~accumulation period (Fig S2h-n,p; Table S1). The IITM-ensemble showed correlation above 0.8 over most of the regions in India and generally ~~skill~~skill in Tmax forecast are better than that of precipitation. However, all the forecast products showed a negative correlation (OBS and forecast) in the Northern Himalayan region, which can be partially attributed to sparse gage stations in the complex regions of Himalayas (Mishra, 2015).

At 7 days ~~lead~~accumulation period, the forecast products showed higher MAE in the Northwestern arid region, Himalayan range, and Western Ghats (Fig. S3). The IITM products and ensemble mean showed improvement in MAE, which was contributed by enhancements in spatial resolution and bias corrected inputs (SST) in IITM models (Fig. S3a-g,o and Table S1). Overall, the IITM-ensemble showed lower MAE for most of the forecast dates during the monsoon season (Fig S3o and Table S1). Moreover, the IITM-ensemble showed lower all-India median MAE (1.2 °C) as compared to the GEFSv2 (2.0 °C) and CFSv2 (1.7 °C) for 15 days ~~lead~~accumulation period (Fig. S3h-n,p). Similar to 7 day ~~lead~~accumulation period, all India median MAE in Tmax was the lowest in IITM ensemble for 15 days ~~lead~~accumulation period. CFSv2 models showed better skill in Tmax than GEFSv2, which is consistent with the findings of Shah and Mishra (2016b).

Similar to precipitation and Tmax, forecast ~~skill~~ ~~was~~skill was estimated based on correlation and MAE for minimum temperature (Tmin). Tmin from all the forecast products showed lower correlation with OBS as compared to precipitation and Tmax in July-August (Fig. S4). For Tmin, ~~the~~GEFSv2 (correlations for ~~lead 7~~accumulation period of 7 day: 0.55 & ~~lead 15~~and accumulation period of 15 day: 0.52) and the IITM-ensemble (correlation 0.52 & ~~and~~ 0.48 for 7 and 15 day) showed comparable ~~skill~~skill (Table S1). For Tmax and Tmin forecasts, the IITM-ensemble showed lower all-India median

Field Code Changed

Field Code Changed

MAE as compared to ~~the~~ GEFsv2 (Figs. S3,S5 and Table S1). Predictions of Tmin from all the products showed weaker performance than Tmax, which was also reported in Shah and Mishra (~~2016b~~)(2016b). The difference in the performance of Tmax and Tmin can be explained as Tmax is mostly governed by partitioning of energy budget which can be simulated by land surface models) whereas Tmin depends on ~~nighttime~~night time boundary conditions and presence of clouds on infrared losses (which may be difficult to simulate) (~~Pattantyus-Abraham et al 2004, Pitman and Perkins 2009~~)(Pattantyus-Abraham et al., 2004; Pitman and Perkins, 2009). Overall, predictions of Tmin from all the forecast products showed higher errors in the Northwest and Himalayan range and for the most cases, the IITM ensemble outperformed the other forecast products (Fig. S5).

3.1.2 Lead ~~time~~ 30 & 45 -days

~~Hydrologic forecast at longer leads (30 to 45 days) are valuable for planning of water resources for irrigation. We, therefore, evaluated forecast skills at 30 and 45 day leads.~~ Since ~~the~~ GEFsv2 reforecast is available only up to lead ~~16~~ ~~time~~16 days, our comparison for the ~~lead~~accumulation period of 30 and 45 days was limited to the forecast products from the IITM and CFSv2. The four IITM products and their ensemble mean showed comparatively better (though not significant) correlations with OBS as compared to ~~the~~ CFSv2 (~~Figure~~Fig. S6, Table S1). We found that the correlations were higher than 0.5 in the majority of western and central India indicating a reasonable skill at 30 ~~days~~ ~~lead~~ ~~time~~ ~~day~~ accumulation period in the IITM ensemble. However, at 45-~~days~~ ~~lead~~ ~~day~~ accumulation period, satisfactory forecast ~~skills~~skill can only be seen in the arid and semi-arid regions where precipitation amount is substantially lower than the other regions in India (Fig. S6). These results indicate that based on correlations, a reasonable skill can be obtained in the precipitation forecast from the IITM products.

Precipitation forecast at ~~lead~~ ~~accumulation period of~~ 30 and 45 days showed ~~similar~~ spatial patterns ~~similar to that~~ of MAE as were observed for the ~~lead~~ ~~accumulation period of~~ 7 ~~&~~and 15 days (Fig. S7).The IITM ensemble showed an improvement in error over ~~the~~ CFSv2 in the majority of India (Fig. S7). The IITM ensemble mean showed lower error for ~~lead~~ ~~accumulation period of~~ 30 and 45 days (Fig. S7m,n). This improvement in correlation and MAE can be attributed to finer resolution of the models and bias corrected SSTs, as shown by the IITM-CFST382, and IITM-GFST382 in comparison to IITM-GFST126, IITM-CFST126, GEFsv2, and CFSv2.

Prediction of Tmax from the IITM ensemble showed significant and higher correlation with OBS at 30 ~~days~~ ~~lead~~ ~~day~~ ~~accumulation period~~, with major contribution from the IITM-GFST382 product (Fig. S8). We notice that the IITM ensemble showed correlations more than 0.6 for the majority of India between OBS and predicted Tmax at 30 day ~~lead~~ ~~accumulation period~~. At 45 day ~~lead~~ ~~accumulation period~~, correlation decreases (in comparison to 30 day ~~lead~~ ~~accumulation period~~), however, predictions of Tmax from the IITM ensemble mean showed better ~~skills~~skill than CFSv2 with OBS. Spatial patterns of MAE in Tmax prediction for ~~lead~~ ~~accumulation period of~~ 30 and 45 days were consistent with spatial patterns ~~at~~ ~~lead~~ ~~for~~ ~~accumulation period of~~ 7 and 15 days indicating larger errors in predicted Tmax in the northern and western parts of

the country (Fig S9). Predictions of Tmin showed lower correlation as compared to Tmax (similar to shorter lead times), especially in the Northwest region, where correlations were negative (Fig S10). Predictions of Tmin from the IITM-GFST126 and GFST382 showed better correlation in the southern peninsula. Spatial patterns of MAE in Tmin predictions at ~~lead accumulation period of~~ 30 and 45 days were consistent with spatial pattern ~~at lead for~~ 7 and 15 days ~~accumulation period~~ (Fig S11). Predictions from ~~the~~ IITM-CFST382 product showed lower errors as compared to the all other products (Table S1). Predictions of Tmin from the IITM-ensemble mean showed lower error (~~lead-30 day accumulation period:~~ 0.9 and ~~lead-45 day accumulation period:~~ 1.1 °C) as compared to ~~the~~-CFSv2 (1.2 and 1.2 °C ~~for 30 and 45 days accumulation periods~~) [Table S1]. Overall, the IITM ensemble performs better than ~~the~~-GEFSv2 and CFSv2 for all the ~~lead-time~~~~accumulation~~ ~~periods~~ (7-45 days). Moreover, the IITM ensemble mean also outperforms other products from the IITM in most of the cases in terms of their individual performance.

Since the IITM ensemble performed better than the other forecast products from IITM, the performance of the IITM ensemble was compared against ~~the~~-CFSv2 ~~at for~~ 7-45 day ~~leads~~ ~~45 days accumulation periods~~ (Fig.2). Since the forecast skill declines with the lead time, we discuss forecast ~~skills~~ ~~skill~~ at 45 day ~~lead~~~~accumulation period~~ in details and results for the other leads are presented in supplemental ~~information~~ Fig. S12. At 45 day ~~lead~~~~accumulation period~~, correlation in precipitation forecast from ~~the~~-CFSv2 is more than 0.2 only in a few regions (mainly ~~entered~~~~centred~~ in northern and western India) [Fig.2a]. The IITM-ensemble showed correlation (~0.3) higher than ~~the~~-CFSv2 (Fig 2b) in most of ~~region~~~~the~~ ~~regions~~, especially during July-August months (Fig 2c). For Tmax and Tmin forecasts, the IITM-ensemble showed higher correlations than ~~the~~-CFSv2 in majority of India (Fig 2d,e,g,h). We found that the difference in forecast ~~skills~~ ~~skill~~ from the IITM-ensemble and CFSv2 is higher ~~at for~~ longer ~~lead-times:~~~~accumulation periods~~. At 7 day ~~lead~~~~accumulation period~~, precipitation forecast from ~~the~~-CFSv2 and IITM-ensemble showed correlation more than 0.6 in most regions in India, therefore, ~~at for~~ shorter ~~leads~~~~accumulation periods~~, difference in the forecast ~~skills~~ ~~skill~~ of CFSv2 and IITM-ensemble is moderate (Fig S12a). For 15 and 30 day ~~leads~~~~accumulation period~~, difference in correlations shown by ~~the~~-CFSv2 and IITM ensemble was higher than for 45 days- ~~accumulation period~~ (Fig S13 and S14). These results show that the IITM ensemble forecast of precipitation, Tmax, and Tmin have better ~~skills~~ ~~skill~~ than ~~the~~-CFSv2 for ~~the~~ majority of India, which can be used for hydrologic prediction of runoff and soil moisture that can be ~~used~~ ~~valuable~~ for decision making of water resources and agriculture. Moreover, for 30 and 45 day ~~leads~~~~accumulation periods~~, the IITM-ensemble showed relatively better forecast ~~skills~~ ~~skill~~ than that of ~~the~~ CFSv2.

3.2 Performance of bias-corrected IITM-ensemble

~~Since the forecast products have bias against observations, we evaluated forecast skills after the bias correction, which may improve forecast skills in precipitation and temperature anomalies. Bias in mean air temperature and precipitation was corrected using Q-Q mapping and linear scaling, respectively as described in Shah and Mishra (2016b). We corrected mean~~

air temperature using Q-Q mapping and estimated the bias between corrected and uncorrected mean temperature, which was added to Tmax and Tmin to obtain daily maximum and minimum temperatures. Since the sample size was small, we evaluated the effectiveness of the Q-Q mapping using pooled data for all the selected forecast dates and found that results were consistent with corrected temperature for specific forecast dates. Figure 3 shows improvements in MAE in predictions of precipitation and mean air temperature after the bias correction. Our results show that the bias correction resulted in reduction in all-India median MAE in precipitation predictions for all the forecast dates during the monsoon season months (Fig. 3c) especially in the Himalayan range and Northeast region (Fig 3a,b). We found substantial improvements in MAE of maximum and minimum temperatures after the bias correction (Figs 3d,e). For instance, all India median MAE was reduced for all the forecast dates after the bias correction (Fig. 3f). Median reduction in MAE for all dates was observed as 2.1 °C. We found that the bias correction substantially improved temperature forecast from the IITM ensemble. This improvement in temperature forecast can be valuable for hydrologic applications. For instance, air temperature influences energy budget in hydrologic models and therefore can affect the partitioning of evapotranspiration and runoff. Due to high intraseasonal variability in the monsoon season precipitation, bias correction resulted in only marginal improvements in the precipitation forecast.

Fig.S15 shows bias in precipitation forecast from the IITM ensemble mean, before and after linear scaling for multifold trials using leave one year out approach. It can be seen We find that linear scaling improved negative bias in precipitation forecast in central India and Western Ghats and positive bias in the Himalayan range and Southern peninsula. During the testing period (one year), improvement in bias is consistent with the training period (Nine year; Supplementary-Fig S15c,d). Improvements in correlation of all-India average precipitation predictions from the IITM-ensemble before and after bias-correction can be noticed (Fig. S16). At 45 days lead a substantial improvement was noticed as compare to other lead times (Fig. S16d). Overall, we noticed that the IITM ensemble mean showed improved forecast skills after the bias correction for most of the regions. At 45 days accumulation period a substantial improvement was noticed as compare to other accumulation periods (Fig. S16d). Overall, we noticed that the IITM ensemble mean showed improved forecast skill after the bias correction for most of the regions. We bias corrected the forecast products for the accumulation period of 45 days. However, the bias in the forecast products may have temporal variability and may not be constant for the entire period of 45 days. Therefore, bias correction approaches based on the variable lead time (Stockdale, 1997) need to be evaluated in future when IITM forecast for long-term retrospective period is available. However, the bias correction approach that we presented can be applied to evaluate seasonal forecast skill

3.3 Prediction of Soil moisture and total runoff

The VIC model was calibrated and evaluated using observed streamflow and satellite soil moisture and runoff is essential for planning and decision making in agriculture and water resources. Hence, we evaluated forecast skills of evapotranspiration (Shah and Mishra,2016a and Shah and Mishra,2016b). In this study, we used calibrated VIC model

forced with observed IMD data to simulate soil moisture and runoff simulated using meteorological variables from IITM ensemble (Asoka and Mishra 2015). Using the raw and bias corrected forecasts (precipitation, maximum, and minimum temperatures), the Variable Infiltration Capacity (VIC) model was run to obtain soil moisture and total runoff (surface runoff+baseflow) forecast. We evaluated improvements in correlation of runoff and soil moisture predicted using the bias-corrected precipitation and temperatures from the IITM ensemble (IITM-ensemble-bc) against uncorrected (raw) precipitation and temperatures from the IITM ensemble mean (IITM-ensemble) and the CFSv2 (Fig S17). For simulating runoff and soil moisture, forcings from all the three products were used to run the VIC model at 0.25° and daily resolution while initial hydrologic conditions were generated using the observed forcing from the IMD. Forecast skills in hydrologic prediction were evaluated for mean total runoff and soil moisture for 7-45 day leads. We, which was considered 45 day lead times as a reference to evaluate the hydrologic prediction skills as for shorter lead times forecast skills are generally higher owing to persistence in initial hydrologic condition-forecast of soil moisture and runoff. Forecast of root-zone soil moisture and runoff was simulated using the VIC model forced with the forecast products (IITM-ensemble-bc, IITM-ensemble, and CFSv2), which were evaluated against the soil moisture/runoff obtained from the VIC model simulation using the observed forcing from IMD (Fig. S17). For all the forecast dates predicted root-zone soil moisture (top 60 cm soil moisture; Fig. S17S14) showed higher correlation than total runoff (Supplementary Fig S17), which is due to higher persistence in soil moisture as compared to runoff (Shah and Mishra 2016b). Improvements in all India median (Shah and Mishra, 2016b). The bias-corrected IITM-ensemble showed higher correlations can be observed with bias corrected IITM-ensemble than the uncorrected IITM-ensemble and the CFSv2, with higher improvements in runoff forecast.

We evaluated forecast skill of the IITM-ensemble for hydroclimatic anomalies using Critical Success Index (CSI). For estimation of CSI, models ability to predict dry anomaly were considered for precipitation, runoff, and soil moisture (i.e. standardized anomaly < 0) and warm anomaly for air temperature (i.e. temperature anomaly > 0). Observed climatology was used to estimate anomalies of precipitation, air temperature, soil moisture and runoff. Fig. 4 shows spatial pattern of CSI averaged overall forecast dates for precipitation, daily-mean temperature, runoff, and soil moisture for 45 day lead-. CSI of predicting dry anomaly in precipitation using the IITM-ensemble were higher in the Northwestern region whereas lower in Himalayan range and southern peninsula as compared to the CFSv2, which is consistent with the results based on correlation and MAE (Fig. 4). Bias corrected IITM-ensemble showed an improved CSI in comparison to the raw forecast from the IITM-ensemble and CFSv2 for the majority of regions in India. However, CSI of predicting warm temperature anomalies was lower than that of CSI of predicting dry precipitation anomalies (Fig. 4), especially in Himalayan ranges. This can be an account of due to higher uncertainty among observations in this region (Mishra 2015); (Mishra, 2015). CSI in runoff and soil moisture is higher as compared to precipitation and temperature due to persistence in initial hydrologic conditions (Fig. 4). For 7, 15 and 30 day leads, accumulation periods CSI is higher than that of 45 day lead accumulation period (Fig S18). We observed that as lead accumulation period was increased from 7 to 45 days, CSI of runoff declines in the arid and semi-arid

regions of northwest. Overall, we found that the bias correction of forecast improves CSI of precipitation, temperature, total runoff, and soil moisture anomalies in India.

5 To show the utility of bias corrected forecast in hydrologic prediction in India, we analyzed the forecast for one of the recent drought years in India. ~~Figure 5 shows anomalies~~ Anomalies of total runoff and root-zone soil moisture predicted on 15th July 2009 for 45 day ~~lead~~ accumulation period using the VIC model with the bias corrected IITM-ensemble forecast, ~~which~~ were compared against the observed anomalies (~~the VIC model was forced with the observed data~~ Fig. 5). Forecast of these hydroclimatic anomalies at sufficient lead time can be helpful in decision making related to water resources and agriculture.

10 We found that the IITM-ensemble-bc successfully captured spatial pattern of observed anomalies, which demonstrates the utility of hydroclimatic forecast for various applications. Persistence in initial hydrologic conditions simulated using the observed forcing and ability of the IITM-ensemble-bc to capture anomalies in precipitation and temperature (Fig S19) resulted in an improved forecast of total runoff and root-zone soil moisture in the majority of regions in India. However, some overestimation in the areal extent and severity of hydroclimatic anomalies can be noted in central India. These results
15 show that the framework developed using the IITM-ensemble-bc forecast and the VIC model can be used to predict runoff and soil moisture up to 45 day ~~lead~~ accumulation period of forecast. Early-warning based on predictions can be helpful in decision making in water resources and agricultural sector so as to minimize risk.

4. ~~Discussions~~ Summary and Conclusions

20 Hydrologic forecast at 7-45 day ~~lead time~~ accumulation period is essential for decision making in agriculture and water resources. Considering the importance of hydrologic prediction in India, we evaluated ~~the~~ CFSv2, GEFSv2, and forecast products from the IITM. We found that meteorological variables predicted using the IITM products, especially the IITM-ensemble showed better forecast ~~skills~~ skill than the other two (CFSv2 and GEFSv2) products for all the ~~lead~~ time accumulation periods (7,15, 30, and 45 days) during the monsoon season. We observed improved skills for ~~lead~~ the accumulation period 30 and 45 days by using the IITM-ensemble in comparison to ~~the~~ CFSv2, which may be associated with the improvement in model resolution and initial condition used at IITM. For instance, Roxy et al. ~~(2015) reported that the CFSv2 has cold bias of 2-3°C in SSTs which may lead to dry bias in the monsoon season in India.~~ (Roxy et al., 2015) reported that CFSv2 has cold bias of 2-3°C in SSTs which may lead to dry bias in the monsoon season in India. Abhilash et al.(2014a) showed that forcings from GFS and CFS models with bias-corrected SSTs lead to improvement in predictability over the Indian region and that is due to improvement in ability to capture active and break spells. The IITM-ensemble
30 performs better than individual IITM products for most of the selected forecast dates. ~~This is consistent with findings of Palmer et al.(2004) and Kirtman et al (2014), where they reported that multimodel ensemble outperform individual model.~~

This is consistent with findings of Palmer et al.(2004) and Kirtman et al. (2014), where they reported that multimodel ensemble outperforms individual model. One of the limitations of evaluation of the forecast products in this study is small sample size. The evaluation of all the forecast products was based on 10 common years between all products and 9 forecast dates during the monsoon season. Increasing the sample size in future based on the availability of forecasts for longer period may further improve evaluation and the bias correction. Our results showed higher forecast skill in the IITM ensemble, which might be associated with its ability to capture intraseasonal variability of rainfall during the monsoon season. The major factors that might have contributed in the improvements in the IITM forecast are:

- i. Ensemble members of IITM forecast are generated by perturbing initial atmospheric conditions to improve simulation of northward propagation
- ii. Improvements in the boundary conditions with bias corrected SST result in improved precipitation prediction
- iii. Higher spatial resolution of the IITM forecast can better resolve orographic rainfall

We evaluated the performance of bias corrected forecast from the IITM-ensemble for lead time accumulation periods up to 45 days. Linear-scaling of precipitation forecast and Q-Q mapping of temperature forecast resulted in reduced errors and bias in forecast in India. Linear scaling precipitation with multifold validation showed an improvement in the Himalayan range and south central region. Bias correction of precipitation and air temperatures resulted in an improvement of about 2.1 mm and 2.1°C, respectively in all-India median of mean absolute error. Total runoff and root-zone soil moisture forecasts obtained using the corrected IITM-ensemble showed higher skill as compared to the CFSv2 and raw IITM-ensemble for lead time accumulation period up to 45 days. We found improvements in that all-India median CSI for runoff forecast was improved from 0.63 to 0.71 after bias-corrected IITM-ensemble in capturing anomalies correction while CSI of soil moisture and total runoff in comparison to the CFSv2 and raw IITM-ensemble, forecast was improved from 0.6 to 0.67 for 45 day accumulation period.

Using forcing from the IITM-ensemble and the VIC model, anomalies in precipitation, temperature, root-zone soil moisture, and total runoff were successfully predicted, which can be used in decision making in water resources and agriculture. The bias corrected forecast from the IITM ensemble, which outperforms the GEFSv2 and CFSv2, can be used to develop a hydrologic prediction platform for India. Information on forecast of anomalies in 7-45 day advance with the existing drought monitoring system in India (Shah and Mishra 2015)(Shah and Mishra, 2015) can be valuable for decision-making in water-resources and agriculture. The hydrologic prediction based on the IITM-ensemble and the VIC model can provide a basis to predict both meteorological and hydrological anomalies and the information can be provided to farmers and water managers. The forecast of root-zone soil moisture along with precipitation and temperature anomalies can be used for irrigation planning. Moreover, runoff forecast at 7-45 day accumulation period can be valuable for water managers in India.

Acknowledgement

Authors acknowledge the data availability from the Climate Forecast System (CFSv2), Global Ensemble Forecast System (GEFS), and Indian Institute of Tropical Meteorology (IITM). Financial assistance from the ITRA-Water project was greatly appreciated.

5

References

- Abhilash, S., Sahai, A. K., Borah, N., Chattopadhyay, R., Joseph, S., Sharmila, S., De, S. and Goswami, B. N.: Does bias correction in the forecasted SST improve the extended range prediction skill of active-break spells of Indian summer monsoon rainfall?: Bias correction in SST and extended range prediction, *Atmospheric Sci. Lett.*, 15(2), 114–119, doi:10.1002/asl2.477, 2013.
- Abhilash, S., Sahai, A. K., Borah, N., Chattopadhyay, R., Joseph, S., Sharmila, S., De, S. and Goswami, B. N.: Does bias correction in the forecasted SST improve the extended range prediction skill of active-break spells of Indian summer monsoon rainfall?: Bias correction in SST and extended range prediction, *Atmospheric Sci. Lett.*, 15(2), 114–119, doi:10.1002/asl2.477, 2014a.
- Abhilash, S., Sahai, A. K., Borah, N., Chattopadhyay, R., Joseph, S., Sharmila, S., De, S., Goswami, B. N. and Kumar, A.: Prediction and monitoring of monsoon intraseasonal oscillations over Indian monsoon region in an ensemble prediction system using CFSv2, *Clim. Dyn.*, 42(9–10), 2801–2815, doi:10.1007/s00382-013-2045-9, 2014b.
- Abhilash, S., Sahai, A. K., Borah, N., Joseph, S., Chattopadhyay, R., Sharmila, S., Rajeevan, M., Mapes, B. E. and Kumar, A.: Improved Spread–Error Relationship and Probabilistic Prediction from the CFS-Based Grand Ensemble Prediction System, *J. Appl. Meteorol. Climatol.*, 54(7), 1569–1578, doi:10.1175/JAMC-D-14-0200.1, 2015.
- ~~AsokaAghaKouchak, A. and Mishra, V.: Prediction of vegetation anomalies to improve food security~~Mehran, A.: ~~Extended contingency table: Performance metrics for satellite observations and water management in India: NDVI FORECAST IN INDIA, Geophys-climate model simulations: TECHNICAL NOTE, Water Resour. Res.-Lett., 42(13), 5290–5298., 49(10), 7144–7149, doi:10.1002/2015GL063991, 2015~~wrcr.20498, 2013.
- Borah, N., Sahai, A. K., Abhilash, S., Chattopadhyay, R., Joseph, S., Sharmila, S. and Kumar, A.: An assessment of real-time extended range forecast of 2013 Indian summer monsoon: AN ASSESSMENT OF REAL-TIME ER FORECAST OF 2013 INDIAN SUMMER MONSOON, *Int. J. Climatol.*, 35(10), 2860–2876, doi:10.1002/joc.4178, 2015.
- Dai, A., Trenberth, K. E. and Qian, T.: A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming, *J. Hydrometeorol.*, 5(6), 1117–1130, 2004.
- Duffy, P. B., Govindasamy, B., Iorio, J. P., Milovich, J., Sperber, K. R., Taylor, K. E., Wehner, M. F. and Thompson, S. L.: High-resolution simulations of global climate, part 1: present climate, *Clim. Dyn.*, 21(5–6), 371–390, doi:10.1007/s00382-003-0339-z, 2003.
- Gao, H., Tang, Q., Shi, X., Zhu, C., Bohn, T. J., Su, F., Sheffield, J., Pan, M., Lettenmaier, D. P., Wood, E. F. and others: Water budget record from Variable Infiltration Capacity (VIC) model, *Algorithm Theor. Basis Doc. Terr. Water Cycle Data Rec.* [online] Available from:

Formatted: Space After: 0 pt

https://www.researchgate.net/profile/Xiaogang_Shi/publication/268367169_6_Water_Budget_Record_from_Variable_Infiltration_Capacity_VIC_Model/links/55715dee08aee701d61cc286.pdf (Accessed 2 August 2016), 2010.

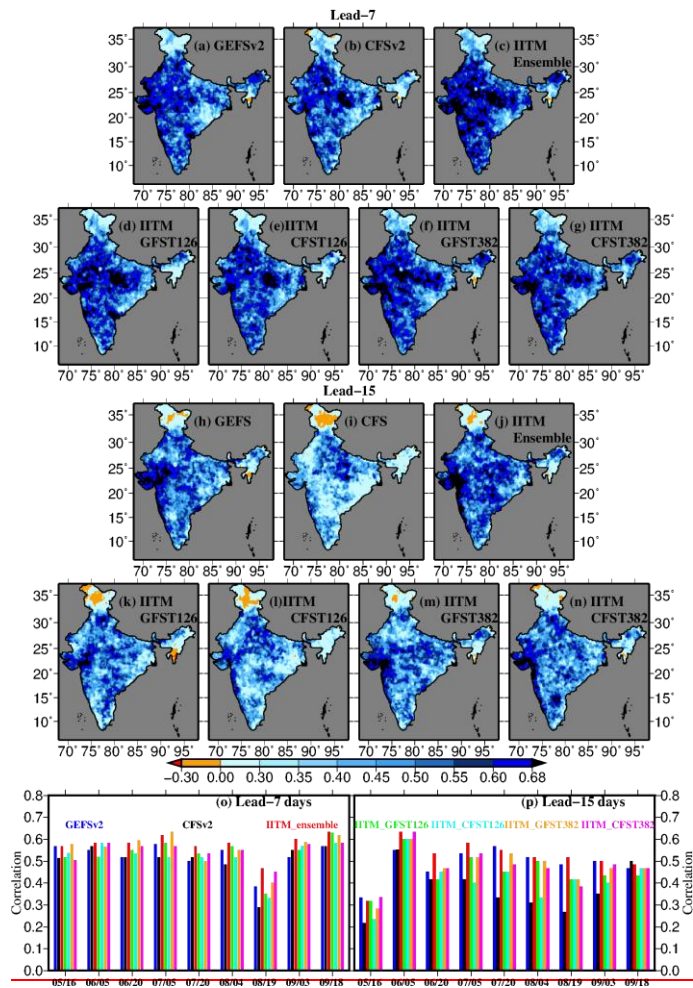
- Goswami, B. N. and Ajayamohan, R. S.: Goswami_AjayaMohan_ISO_IAV.pdf, *Journal of Climate*, 14, 1180–1198, 2000.
- 5 | Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau Jr, T. J., Zhu, Y. and Lapenta, W.: NOAA's second-generation global medium-range ensemble reforecast dataset, *Bull. Am. Meteorol. Soc.*, 94(10), 1553–1565, 2013.
- | Joseph, S., Sahai, A. K., Abhilash, S., Chattopadhyay, R., Borah, N., Mapes, B. E., Rajeevan, M. and Kumar, A.: Development and Evaluation of an Objective Criterion for the Real-Time Prediction of Indian Summer Monsoon Onset in a Coupled Model Framework, *J. Clim.*, 28(15), 6234–6248, doi:10.1175/JCLI-D-14-00842.1, 2015a.
- 10 | Joseph, S., Sahai, A. K., Sharmila, S., Abhilash, S., Borah, N., Chattopadhyay, R., Pillai, P. A., Rajeevan, M. and Kumar, A.: North Indian heavy rainfall event during June 2013: diagnostics and extended range prediction, *Clim. Dyn.*, 44(7–8), 2049–2065, doi:10.1007/s00382-014-2291-5, 2015b.
- | Kirtman, B. P., Min, D., Infanti, J. M., Kinter III, J. L., Paolino, D. A., Zhang, Q., Van Den Dool, H., Saha, S., Mendez, M. P., Becker, E. and others: The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2
15 | toward developing intraseasonal prediction, *Bull. Am. Meteorol. Soc.*, 95(4), 585–601, 2014.
- | Kumar, K. K., Rajagopalan, B. and Cane, M. A.: On the Weakening Relationship Between the Indian Monsoon and ENSO, *Science*, 284(5423), 2156–2159, doi:10.1126/science.284.5423.2156, 1999.
- | Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G. and Cane, M.: Unraveling the Mystery of Indian Monsoon Failure During El Niño, *Science*, 314(5796), 115–119, doi:10.1126/science.1131152, 2006.
- 20 | Liang, X., Lettenmaier, D. P., Wood, E. F. and Burges, S.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, 99(D7), 14415–14428, doi:10.1029/94JD00483, 1994.
- | Liang, X., Wood, E. F. and Lettenmaier, D. P.: Soil Moisture Simulation Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, *Glob. Planet. Change*, 13(1), 195–206, doi:10.1016/0921-8181(95)00046-1, 1996.
- | Livneh, B. and Hoerling, M. P.: The Physics of Drought in the U.S. Central Great Plains, *J. Clim.*, doi:10.1175/JCLI-D-15-0697.1, 2016.
- 25 | Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P. and Nijssen, B.: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States*, *J. Clim.*, 15(22), 3237–3251, 2002.
- | McEvoy, D. J., Huntington, J. L., Mejia, J. F. and Hobbins, M. T.: Improved seasonal drought forecasts using reference evapotranspiration anomalies, *Geophys. Res. Lett.*, 43(1), 377–385, doi:10.1002/2015GL067009, 2016.
- 30 | Mishra, V.: Climatic uncertainty in Himalayan water towers: Climatic uncertainty, *J. Geophys. Res. Atmospheres*, 120(7), 2689–2705, doi:10.1002/2014JD022650, 2015.
- | Mishra, V., Smoliak, B. V., Lettenmaier, D. P. and Wallace, J. M.: A prominent pattern of year-to-year variability in Indian Summer Monsoon Rainfall, *Proc. Natl. Acad. Sci.*, 109(19), 7213–7217, doi:10.1073/pnas.1119150109, 2012.
- | Mishra, V., Shah, R. and Thrasher, B.: Soil Moisture Droughts under the Retrospective and Projected Climate in India, *J. Hydrometeorol.*, 15, 2267–2292, doi:10.1175/JHM-D-13-0177.1, 2014.
- 35

- Mo, K. C. and Lettenmaier, D. P.: Hydrologic Prediction over the Conterminous United States Using the National Multi-Model Ensemble, *J. Hydrometeorol.*, 15(4), 1457–1472, doi:10.1175/JHM-D-13-0197.1, 2014.
- Pai, D. S., Sridhar, L., Badwaik, M. R. and Rajeevan, M.: Analysis of the daily rainfall events over India using a new long period (1901–2010) high resolution ($0.25^\circ \times 0.25^\circ$) gridded rainfall data set, *Clim. Dyn.*, 45(3), 755–776, doi:10.1007/s00382-014-2307-1, 2015.
- Palmer, T. N., Doblas-Reyes, F. J., Hagedorn, R., Alessandri, A., Gualdi, S., Andersen, U., Feddersen, H., Cantelaube, P., Terres, J.-M., Davey, M., Graham, R., Délecluse, P., Lazar, A., Déqué, M., Guérémy, J.-F., Díez, E., Orfila, B., Hoshen, M., Morse, A. P., Keenlyside, N., Latif, M., Maisonnave, E., Rogel, P., Marletto, V. and Thomson, M. C.: DEVELOPMENT OF A EUROPEAN MULTIMODEL ENSEMBLE SYSTEM FOR SEASONAL-TO-INTERANNUAL PREDICTION (DEMETER), *Bull. Am. Meteorol. Soc.*, 85(6), 853–872, doi:10.1175/BAMS-85-6-853, 2004.
- Park Williams, A., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., Swetnam, T. W., Rauscher, S. A., Seager, R., Grissino-Mayer, H. D., Dean, J. S., Cook, E. R., Gangodagamage, C., Cai, M. and McDowell, N. G.: Temperature as a potent driver of regional forest drought stress and tree mortality, *Nat. Clim. Change*, 3(3), 292–297, doi:10.1038/nclimate1693, 2012.
- Pattantyus-Abraham, M., Kiraly, A. and Janosi, I.: Nonuniversal atmospheric persistence: difference scaling of daily minimum and maximum temperatures, *Phys. Rev. E*, 69(2114021110), 1–7, 2004.
- Pitman, A. J. and Perkins, S. E.: Global and Regional Comparison of Daily 2-m and 1000-hPa Maximum and Minimum Temperatures in Three Global Reanalyses, *J. Clim.*, 22(17), 4667–4681, doi:10.1175/2009JCLI2799.1, 2009.
- Rahman, S. H., Sengupta, D. and Ravichandran, M.: Variability of Indian summer monsoon rainfall in daily data from gauge and satellite, *J. Geophys. Res.*, 114(D17), doi:10.1029/2008JD011694, 2009.
- Rajeevan, M., Bhate, J., Kale, J. D. and Lal, B.: Development of a high resolution daily gridded rainfall data for the Indian region, *Met Monogr. Climatol.*, 22, 2005, 2005.
- Rajeevan, M., Bhate, J., Kale, J. D. and Lal, B.: High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells, *Curr. Sci.*, 91(3), 296–306, 2006.
- Rajeevan, M., Gadgil, S. and Bhate, J.: Active and break spells of the Indian summer monsoon, *J. Earth Syst. Sci.*, 119(3), 229–247, 2010.
- Robson, A. J.: Evidence for trends in UK flooding, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, 360(1796), 1327–1343, doi:10.1098/rsta.2002.1003, 2002.
- Roebber, P. J., Schultz, D. M., Colle, B. A. and Stensrud, D. J.: Toward improved prediction: High-resolution and ensemble modeling systems in operations, *Weather Forecast.*, 19(5), 936–949, 2004.
- Roxy, M. K., Ritika, K., Terray, P., Murtugudde, R., Ashok, K. and Goswami, B. N.: Drying of Indian subcontinent by rapid Indian Ocean warming and a weakening land-sea thermal gradient, *Nat. Commun.*, 6, 7423, doi:10.1038/ncomms8423, 2015.
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M. and Becker, E.: The NCEP Climate Forecast System Version 2, *J. Clim.*, 27(6), 2185–2208, doi:10.1175/JCLI-D-12-00823.1, 2014.

- | Sahai, A. K., Sharmila, S., Abhilash, S., Chattopadhyay, R., Borah, N., Krishna, R. P. M., Joseph, S., Roxy, M., De, S., Pattnaik, S. and others: Simulation and extended range prediction of monsoon intraseasonal oscillations in NCEP CFS/GFS version 2 framework, *Curr. Sci.*, 104(10), 1394–1408, 2013.
- | Sahai, A. K., Abhilash, S., Chattopadhyay, R., Borah, N., Joseph, S., Sharmila, S. and Rajeevan, M.: High-resolution operational monsoon forecasts: an objective assessment, *Clim. Dyn.*, 44(11–12), 3129–3140, doi:10.1007/s00382-014-2210-9, 2015a.
- | Sahai, A. K., Chattopadhyay, R., Joseph, S., Mandal, R., Dey, A., Abhilash, S., Krishna, R. P. M. and Borah, N.: Real-time performance of a multi-model ensemble-based extended range forecast system in predicting the 2014 monsoon season based on NCEP-CFSv2, *Curr. Sci.*, 109(10), 1802–1813, 2015b.
- 10| Shah, H. L. and Mishra, V.: Hydrologic Changes in Indian Sub-Continental River Basins (1901-2012), *J. Hydrometeorol.*, doi:10.1175/JHM-D-15-0231.1, 2016a.
- | Shah, R. and Mishra, V.: Evaluation of the Reanalysis Products for the Monsoon Season Droughts in India, *J. Hydrometeorol.*, 15(4), 1575–1591, doi:10.1175/JHM-D-13-0103.1, 2014.
- | Shah, R. D. and Mishra, V.: Development of an Experimental Near-Real-Time Drought Monitor for India*, *J. Hydrometeorol.*, 16(1), 327–345, doi:10.1175/JHM-D-14-0041.1, 2015.
- 15| Shah, R. D. and Mishra, V.: Utility of Global Ensemble Forecast System (GEFS) Reforecast for Medium-Range Drought Prediction in India, *J. Hydrometeorol.*, 17(8), 1781–1800, doi:http://dx.doi.org/10.1175/JHM-D-15-0050.1, 2016b.
- | Sharmila, S., Pillai, P. A., Joseph, S., Roxy, M., Krishna, R. P. M., Chattopadhyay, R., Abhilash, S., Sahai, A. K. and Goswami, B. N.: Role of ocean–atmosphere interaction on northward propagation of Indian summer monsoon intra-seasonal oscillations (MISO), *Clim. Dyn.*, 41(5–6), 1651–1669, doi:10.1007/s00382-013-1854-1, 2013.
- 20| Shepard, D. S.: Computer Mapping: The SYMAP Interpolation Algorithm, in *Spatial Statistics and Models*, edited by G. L. Gaile and C. J. Willmott, pp. 133–145, Springer Netherlands, Dordrecht. [online] Available from: http://dx.doi.org/10.1007/978-94-017-3048-8_7, 1984.
- | Shukla, S. and Lettenmaier, D. P.: Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill, *Hydrol. Earth Syst. Sci. Discuss.*, 8(4), 6565–6592, doi:10.5194/hessd-8-6565-2011, 2011.
- 25| Shukla, S., Safeeq, M., AghaKouchak, A., Guan, K. and Funk, C.: Temperature impacts on the water year 2014 drought in California: TEMPERATURE IMPACTS ON THE 2014 DROUGHT, *Geophys. Res. Lett.*, 42(11), 4384–4393, doi:10.1002/2015GL063666, 2015.
- 30| Srivastava, A. K., Rajeevan, M. and Kshirsagar, S. R.: Development of a high resolution daily gridded temperature data set (1969-2005) for the Indian region, *Atmospheric Sci. Lett.*, n/a-n/a, doi:10.1002/asl.232, 2009.
- | [Stockdale, T. N.: Coupled ocean-atmosphere forecasts in the presence of climate drift. *Mon. Weather Rev.*, 125\(5\), 809–818, 1997.](#)
- 35| Tian, D., Wood, E. F. and Yuan, X.: CFSv2-based sub-seasonal precipitation and temperature forecast skill over the contiguous United States, *Hydrol. Earth Syst. Sci. Discuss.*, 1–24, doi:10.5194/hess-2016-376, 2016.

Formatted: Space After: 0 pt

- | Webster, P. J., Magana, V.O., Palmer, T. N., Shukla, J., Tomas, R. A., Yanai, M. and Yasunari, T.:
webster_etal_1998_intraseasonal.pdf, *J. Geophys. Res.*, 103(C7), 14451–14510, doi:10.1029/97JC02719/, 1998.
- | Wilks, D. S.: *Statistical Methods in the Atmospheric Science*, Academic Press, New York., 2006.
- 5 | Wood, A. W.: Long-range experimental hydrologic forecasting for the eastern United States, *J. Geophys. Res.*, 107(D20),
doi:10.1029/2001JD000659, 2002.
- | Woolnough, S. J., Vitart, F. and Balmaseda, M. A.: The role of the ocean in the Madden–Julian Oscillation: Implications for
MJO prediction, *Q. J. R. Meteorol. Soc.*, 133(622), 117–128, doi:10.1002/qj.4, 2007.
- | Yuan, X. and Wood, E. F.: Downscaling precipitation or bias-correcting streamflow? Some implications for coupled general
circulation model (CGCM)-based ensemble seasonal hydrologic forecast: DOWNSCALING PRECIPITATION OR BIAS-
10 CORRECTING STREAMFLOW, *Water Resour. Res.*, 48(12), n/a-n/a, doi:10.1029/2012WR012256, 2012a.
- | Yuan, X. and Wood, E. F.: On the clustering of climate models in ensemble seasonal forecasting: CLUSTERING OF
SEASONAL FORECAST MODELS, *Geophys. Res. Lett.*, 39(18), n/a-n/a, doi:10.1029/2012GL052735, 2012b.
- | Yuan, X., Wood, E. F., Luo, L. and Pan, M.: A first look at Climate Forecast System version 2 (CFSv2) for hydrological
seasonal prediction: A FIRST LOOK AT CFSv2, *Geophys. Res. Lett.*, 38(13), n/a-n/a, doi:10.1029/2011GL047792, 2011.



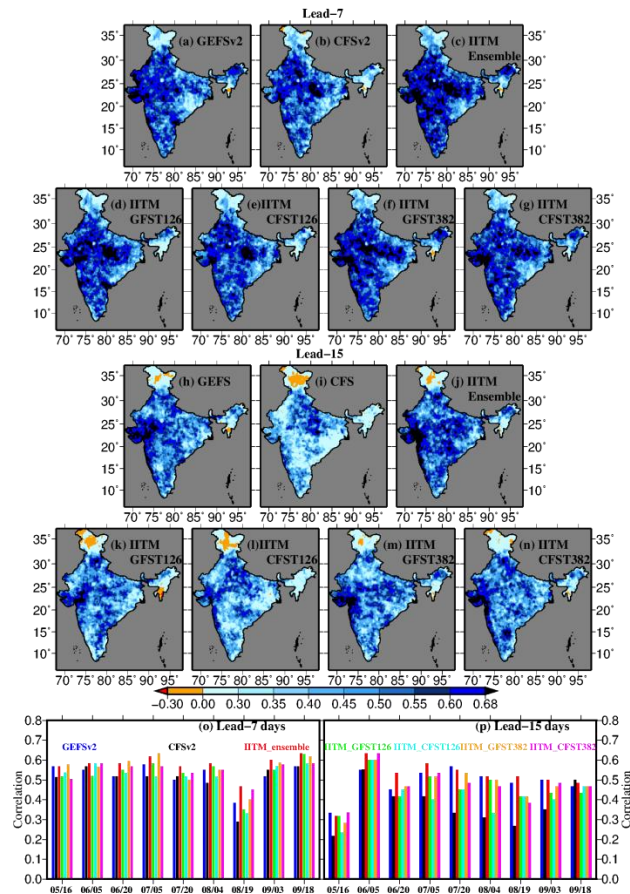


Figure 1: Correlation between precipitation forecasts and observed precipitation (OBS). (a) Correlation between precipitation forecast from the GEFSv2 accumulated up to 7-days **lead accumulation period** and corresponding OBS, (b) same as (a) but for the CFSv2 (c) same as (a) but for the IITM Ensemble, (d) same as (a) but for the IITM GFST126 (e) same as (a) but for the IITM CFST126 (f) same as (a) but for the IITM GFST382 (g) same as (a) but for the IITM CFST382 (h-n) same as (a-f) but for **lead-up-to-15 accumulation period of 15** days. (o) All-India median correlation between different days precipitation forecasts at 7-day **lead accumulation period** and corresponding OBS for the forecasts initiated on different dates (p) same as (o) but for **lead-15 accumulation period of 15** days (period: 2001-2009).

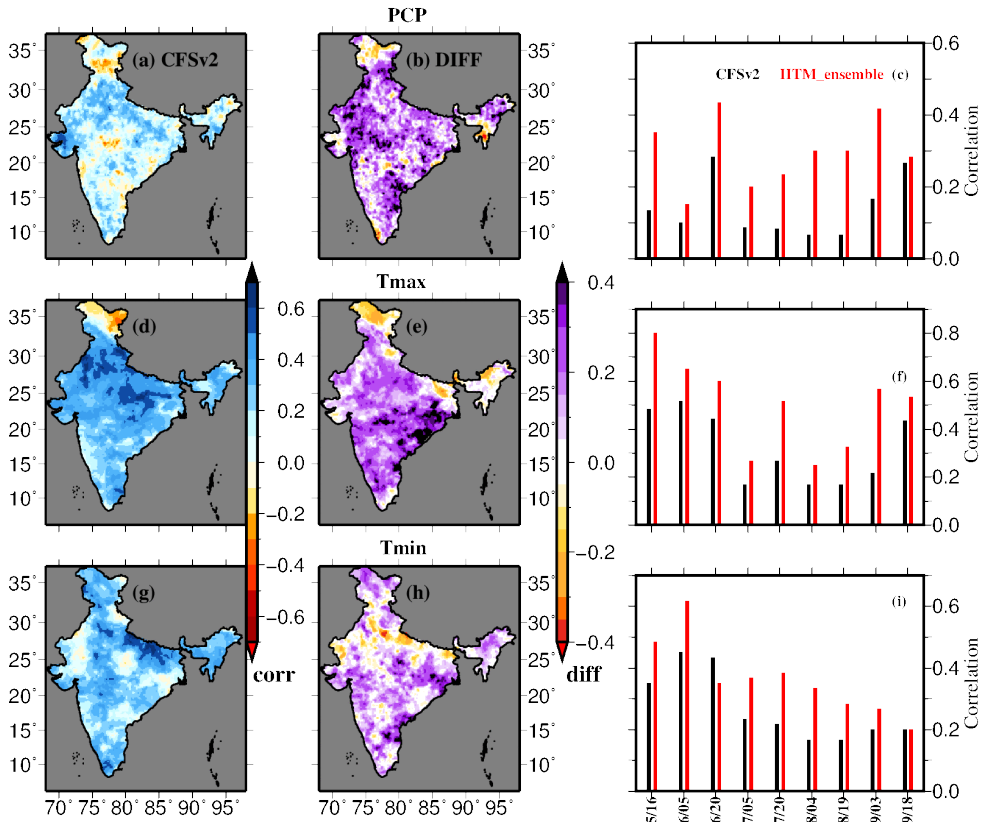


Figure 2: Improvements in correlations in the IITM-ensemble forecast in comparison to the CFSv2 for 45 day lead-accumulation period. (a) correlation between precipitation forecast from the CFSv2 and OBS (b) change in correlation coefficient of precipitation forecast from the IITM-ensemble and OBS as compared to (a). Correlations in (a) and (b) are median of correlations for the different forecast dates during the monsoon season. (c) All-India averaged median correlation for forecast initiated on different forecast dates. (d-f) is same as (a-c) but for daily maximum temperature and (g-i) is same as (a-c) but for daily minimum temperature.

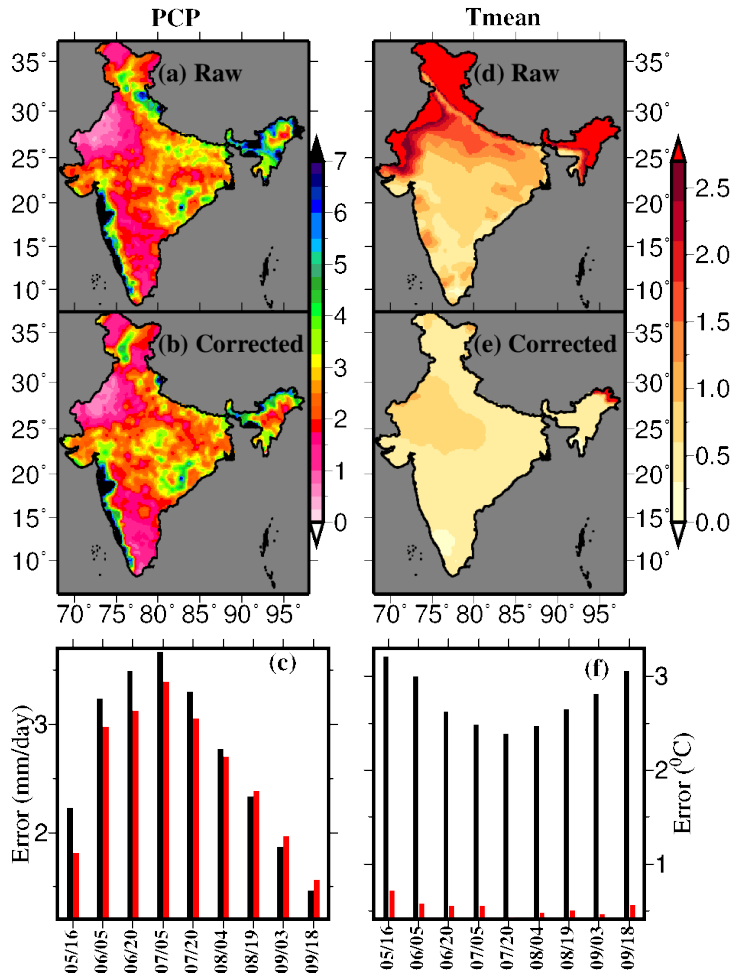


Figure 3: Median absolute error (MAE) in forecast at 45 days lead day accumulation period from the IITM-ensemble before and after bias correction. (a and b) Median (of all forecast dates) MAE (mm/day) in precipitation forecast before and after bias correction. (c) Comparison of all-India median MAE for each forecast dates (d-f) same as (a-c) but for daily mean temperature in °C.

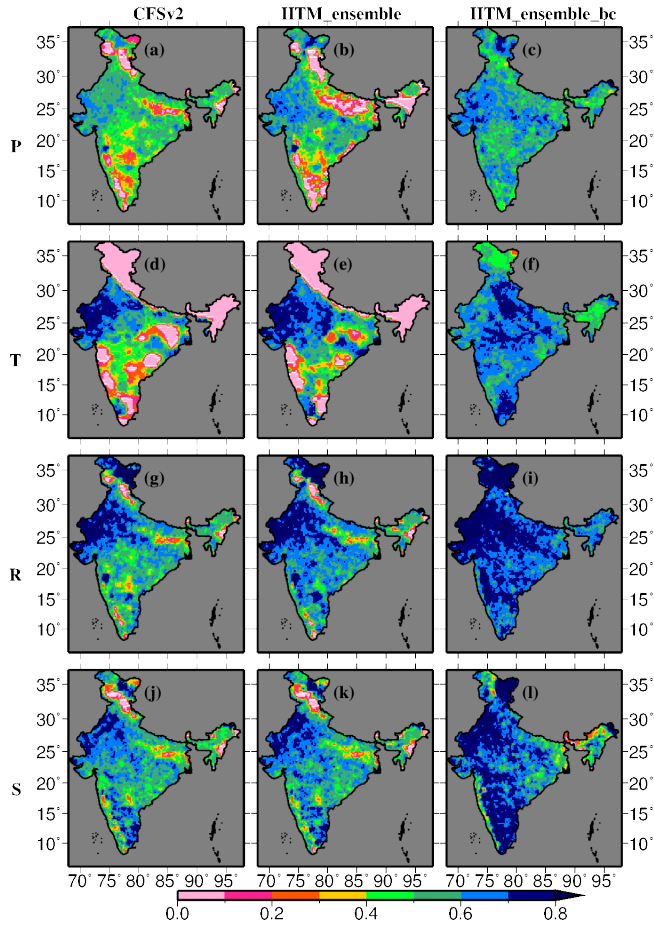


Figure 4: Critical Success Index (CSI, averaged for forecast dates) of predicting precipitation (a-c), temperature (d-f), runoff (g-i), and soil moisture (j-l) anomalies with respect to the observed anomalies for CFSv2, IITM-ensemble, and bias corrected IITM-ensemble (IITM-ensemble_bc).

5

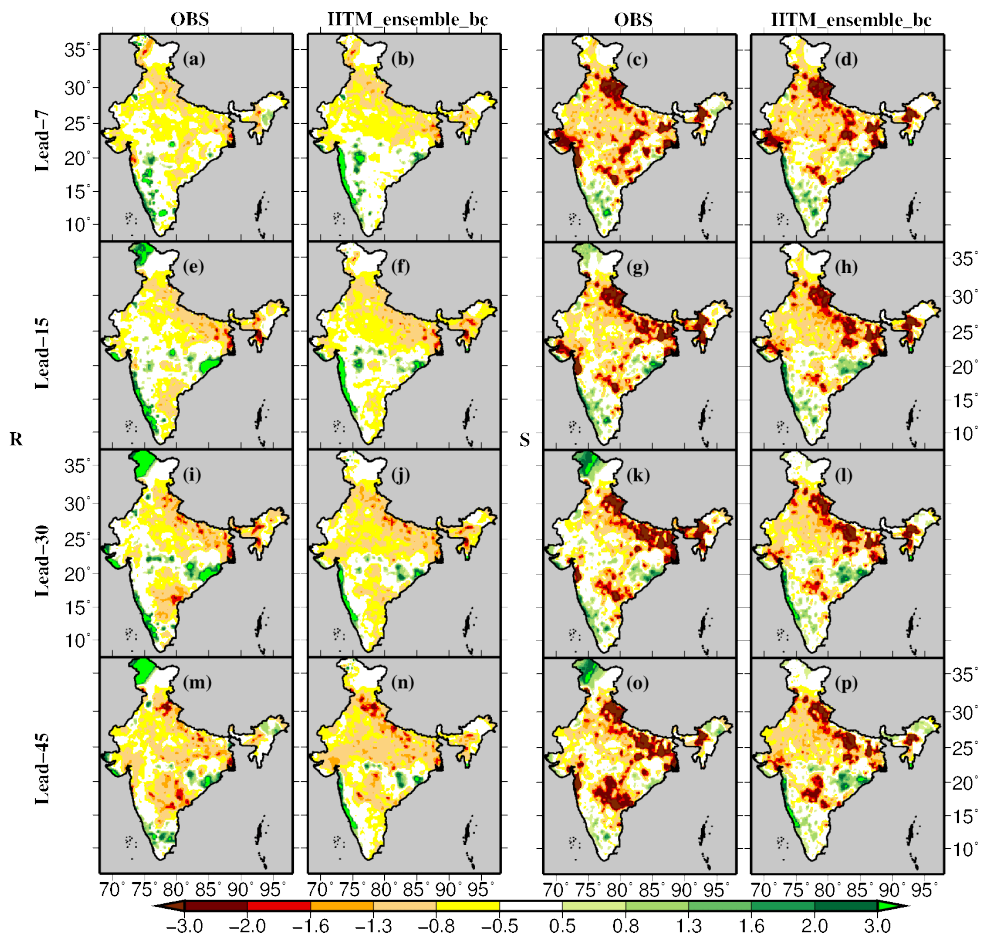


Figure 5: Predicted anomalies of hydrologic variables for forecast initiated on 15th July, 2009 for **lead-accumulation periods of 7, 15, 30, and 45 days**. (a) Observed (standardized) anomalies in (VIC-simulated) runoff at **lead-time 7 days** (b) anomalies in (VIC-simulated) runoff using bias-corrected IITM-ensemble for **lead-7-accumulation period of 7 days**. (c and d) same as (a and b) but for root-zone soil moisture (e-h), (i-l), (m-p) same as (a and b) for **lead-accumulation period of 15, 30 and 45 days**, respectively.

Formatted: Font: Bold