Clim Dyn (2016) 46:1991–2023
DOI 10.1007/s00382-015-2688-9

Credibility of statistical downscaling under nonstationary climate

Kaustubh Salvi[1] · Subimal Ghosh[1,2] · Auroop R. Ganguly[2,3]

Received: 11 May 2014 / Accepted: 25 May 2015 / Published online: 10 June 2015

Abstract Statistical downscaling (SD) establishes empirical relationships between coarse-resolution climate model simulations with higher-resolution climate variables of interest to stakeholders. These statistical relations are estimated based on historical observations at the finer resolutions and used for future projections. The implicit assumption is that the SD relations, extracted from data are stationary or remain unaltered, despite non-stationary change in climate. The validity of this assumption relates directly to the credibility of SD. Falsifiability of climate projections is a challenging proposition. Calibration and verification, while necessary for SD, are unlikely to be able to reproduce the full range of behavior that could manifest at decadal to century scale lead times. We propose a design-of-experiments (DOE) strategy to assess SD performance under nonstationary climate and evaluate the strategy via a transfer-function based SD approach. The strategy relies on selection of calibration and validation periods such that they represent contrasting climatic conditions like hot-versus-cold and ENSO-versus-non-ENSO years. The underlying assumption is that conditions such as warming or predominance of El Niño may be more prevalent under climate change. In addition, two different historical time periods are identified, which resemble pre-industrial and the most severe future emissions scenarios. The ability of the empirical relations to generalize under these proxy conditions is considered an indicator of their performance under future nonstationarity. Case studies over two climatologically disjoint study regions, specifically India and Northeast United States, reveal robustness of DOE in identifying the locations where nonstationarity prevails as well as the role of effective predictor selection under nonstationarity.

Keywords Statistical downscaling · Stationarity · Climate change

1 Introduction

The latest generation of climate (or earth system) models [General Circulation Models (GCMs), or Earth System Models (ESMs)] at global scale, specifically, the ones in the Coupled Model Intercomparison Project phase 5 (CMIP5), incorporates more detailed physics and generates more ensemble runs at higher resolutions than the previous phase 3 (CMIP3) generation (Knutti and Sedláček 2012; Kumar et al. 2014). However, despite specific improvements in large-scale drivers of precipitation (Ryu and Hayhoe 2013), CMIP5 models have not resulted in drastic improvements over CMIP3 for impact-relevant variables (Knutti and Sedláček 2012; Kumar et al. 2014). At regional scales, this is exemplified by the lack of improvement in simulations of the Indian Summer Monsoon Rainfall (ISMR), as shown by Shashikanth et al. (2013) and Ramesh and Goswami (2014). However, as shown in Sperber et al. (2012), certain multi-model diagnostics do show improvements in CMIP5,

Electronic supplementary material The online version of this article (doi:10.1007/s00382-015-2688-9) contains supplementary material, which is available to authorized users.

✉ Subimal Ghosh
subimal@civil.iitb.ac.in

[1] Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India

[2] Interdisciplinary Program in Climate Studies, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India

[3] Sustainability and Data Sciences Laboratory, Civil and Environmental Engineering, Northeastern University, Boston, MA 02115, USA

while specific models improve on certain dynamical processes such as northward propagation and tilted band of convection over the Indian monsoon region.

Two downscaling approaches (obtaining climate projections at high resolution), specifically, (1) dynamical downscaling (DD), which involves developing and running high resolution regional climate models (RCMs) forced with GCM simulated climate variables, and (2) statistical downscaling (SD), composed of data driven models, are used to add value to GCM simulations in terms of regional climatic attributes. While falsification is a challenge in climate projections in general, particular care is needed when evaluating downscaling approaches. The latter is especially true because so called scientific intuitions and educated guesses may be misleading. Thus, while much has been made about SD relations not holding true under a nonstationary climate, it is important to distinguish between the stationarity of the statistical relations versus the non-stationarity of global temperature anomalies. The basic relations among atmospheric variables governed by conservation laws are not expected to change because of global warming. However, if SD relations are calibrated under conditions dominated by stratiform precipitation but regional warming leads to an increase in the convective precipitation fraction, then the statistical relations leading to downscaled precipitation may indeed change. Conversely, DD processes are not necessarily automatically robust to changes in physical processes (e.g. changes in convective fraction of precipitation) unless the relevant physical parameters in the RCMs can capture the change.

The ability of RCMs to add value over GCMs has been debated in the literature (Racherla et al. 2012) and even in news articles (Kerr 2013). However, Laprise (2014) has argued that the results of Racherla et al. (2012) may be influenced by the presence of internal variability and sampling error. Recent literature on DD includes discussions about performance (Pinto et al. 2014; Xue et al. 2014) and usefulness (Glotter et al. 2014). The SD (Wilby 1994; Wilby and Wigley 1997; Wilby et al. 1998, 1999, 2004; Mehrotra and Sharma 2005, 2006; Raje and Mujumdar 2009; Hewitson and Crane 1992; Ghosh and Mujumdar 2006; Anandhi et al. 2008; Wang et al. 2009; Ghosh 2010; Groppelli et al. 2011; Kannan and Ghosh 2013; Salvi et al. 2013) involves establishing empirical relations between variables that are relatively better projected by GCMs (or even RCMs), usually at lower resolutions, with impact-relevant variables at higher resolutions. Figure 1 shows a comparison between mean of GCM simulated (re-gridded at observed rainfall resolution) rainfall (Fig. 1b) and mean of statistically downscaled rainfall at daily temporal resolution (Fig. 1c) with a reference to mean of observed data (Fig. 1a) over 1979–2005. Here, we use the fourth generation GCM simulations from Canadian Centre for Climate Modelling and Analysis (CCCmaCanESM2, CGCM4; resolution ~2.8°). These simulations possess improved features as compared to its third generation counterpart in terms of (1) improved resolution of spectral representation, (2) updated radiative transfer scheme (correlated-k distribution model (Li 2002; Li and Barker 2002, 2005) and more general treatment of radiative transfer in cloudy atmospheres (Pincus et al. 2003; Barker et al. 2008), (3) accounting for direct and indirect radiative effects of aerosols, and (4) inclusion of prognostic bulk aerosol scheme, single-moment cloud microphysics scheme, statistical approach

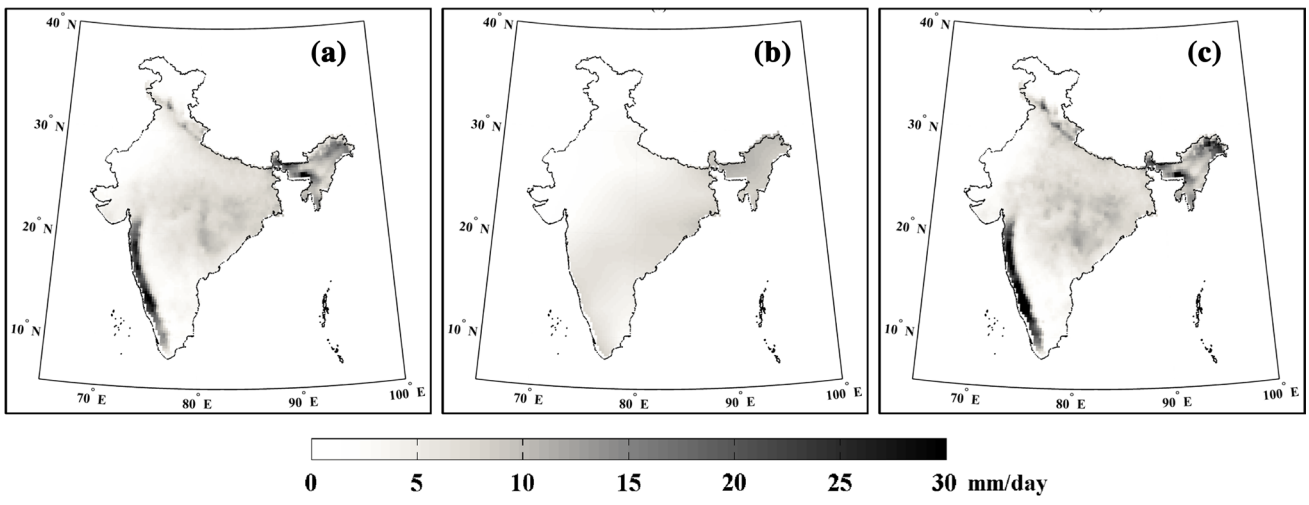


Fig. 1 Value-Addition by statistical downscaling in capturing statistical properties of observed rainfall data (1979–2005). **a** Mean of APHRODITE gridded rainfall data at 0.25° resolution (JJAS), **b** mean of GCM simulated rainfall (Canadian Centre for Climate Modeling and Analysis, CCCma-CanESM2), re-gridded at 0.25° resolution, and **c** mean of statistically downscaled rainfall at 0.25° resolution using kernel regression based downscaling methodology

for macro-physical properties of layer clouds (Chaboureau and Bechtold 2005). SD is performed with kernel regression (Kannan and Ghosh 2013; Salvi et al. 2013). Gridded observed rainfall for India at a 0.25° spatial resolution is obtained from APHRODITE (Asian Precipitation Highly Resolved Observational Data Integration towards the Evaluation of Water Resources project, Japan). Figure 1 shows that GCM simulations fail to capture the magnitude and spatial pattern of mean observed rainfall. Whereas, the downscaled rainfall shows absolute similarities in all respect with observed mean rainfall. The reasons behind the failure of the GCM in simulating spatial variability of rainfall are its coarse resolution, which cannot address fine resolution factors affecting precipitation process such as orography, choice of parameterization schemes, representation of vegetation etc.

SD approaches can be grouped into three categories: transfer functions, weather typing and weather generators. Transfer function methods develop projection models relating the predictors (climate variables at lower resolution) to the variables of interest (Hewitson and Crane 1992; Crane and Hewitson 1998; Wilby et al. 2004), while weather typing relies on grouping local, meteorological variables in relation to different classes of atmospheric circulation. Future regional climate scenarios are constructed either by resampling from the observed variable distribution, conditional on circulation patterns produced by a GCM or first generating synthetic sequences of weather patterns using Monte Carlo techniques and then resampling from the generated data (Hay et al. 1991; Bogardi et al. 1993; Özelkan and Duckstein 1996). Weather generators simulate synthetic sequences of weather variables, which are statistically consistent with the observed characteristics of the historical record in time (Richardson 1981; Hughes et al. 1993; Hughes and Guttorp 1994a, b; Wilks 1999; Khalili and Brissette 2009). The present study uses a state-of-the-art SD approach (Kannan and Ghosh 2013) that essentially combines weather typing with transfer functions. It involves establishing relationship between climate predictors and rainfall state (which is a qualitative representation of rainfall occurring in a given region at different grids). With the established relationship and future climate predictors, future rainfall states are estimated. Rainfall for future is obtained using kernel regression, conditioned on the projected rainfall states.

Despite of the skills, illustrated by SD models in capturing properties of observed data over historic period, these models suffer a major setback in the form of a limitation. A common perception is that SD approaches, being empirically based, are even more sensitive to nonstationarity in climate compared to their physically-based counterparts such as RCM-based DD. This is because, empirical relations are developed based on historical observations and assumed to be stationary over time. This is usually termed as assumption of stationarity. While this is perceived to be a strong assumption given expected nonstationarity under changing climate, surprisingly little literature is devoted in examining this assumption as a hypothesis. Nonstationarity is an inbuilt trait, which is embedded in climate systems at different spatio-temporal scales (Hertig and Jacobeit 2013). Different studies show the existence of non-stationarity in the climate system e.g. time variant circulation to weather links (Huth 1997) and significance of non-stationarity in relationship between circulation pattern and temperature (Slonosky et al. 2001). Wilby et al. (1998) proposed three underlying factors that can be associated with non-stationarity in SD models: an incomplete set of predictor variables, inadequate calibration periods, and situations in which the climate system structure(s) changes through time. With changing climate, the relationship between predictors and the predictand of interest could potentially change. Studies have proposed embedding potential effects of non-stationarity within statistical relationships in SD models. This has been tried either by imposing some modifications in the existing methods or proposing completely new methodologies. Busuioc et al. (1999) proposed a technique for verification of capability of empirical downscaling procedures under changing climatic conditions by comparing outputs of SD models with their parent GCMs. It is assumed that if a GCM shows the skills in simulating regional variable such as rainfall reasonably well, then the same GCM may reasonably simulate the changes in climatology on account of GHG emissions in future. If the changes in climatology, simulated by GCM are similar to those simulated by downscaling methodology, then it is safe to assume that the downscaling model is indirectly capturing the non-stationarity in statistical relationship. The major limitations in this approach are, GCMs are very unlikely to simulate precipitation well, and secondly good simulations of climatology do not ensure reliable simulations of changes (Racherla et al. 2012). Furthermore, the changes in precipitation are guided by local factors such as orography (Salvi et al. 2013), the choice of convective parameterization scheme (Litta et al. 2011), and regional land use land cover (Pielke et al. 2007). Charles et al. (1999) applied nonhomogeneous hidden Markov model as a downscaling technique to a network of 30 daily precipitation stations in Australia and found that inclusion of dew point temperature at 850 hPa plays crucial role in successful performance of downscaling model under climate change conditions. Schmith (2008) carried out a systematic study on the effect of assumption of stationarity on downscaling and showed that selection of predictors plays major role in avoiding non-stationarity. Underwood (2009) applied Generalized Additive Models (GAMs), which explore the non-linear relationships in the data and able to simulate seasonal patterns and long term trends of

observed data. GAMs provide ability to identify the type of relationship (linear or nonlinear) between covariates and response variable. Duan and McIntyre (2012) developed an approach for identification of non-stationarity in downscaling models using regression analysis, which was tested for 50 sites in south-east England over the period '1855–2008'. The regression coefficients for SD are obtained for each of '30 year moving windows' running through the above mentioned period. The existence of significant trends in regression coefficients is treated as an indicator for presence of non-stationarity. Hertig and Jacobeit (2013) developed an approach, which involved application of 'Generalized Linear Regression Model' (GLR) with Poisson distribution to establish relationship between different predictors and rainfall. The relationship is established over different sets of '31 years' period with a moving window approach. A bootstrap based validation is used to generate 1000 iterations of each calibration and validation period and biases are compared.

Climate conditions in the future, especially at multi-decadal to century scales, may alter significantly owing to changes in radiative forcing. Thus, the relative prevalence of convective versus large-scale precipitation generation may change under a warming environment. An SD approach that is capable of adjusting to these changes in states and captures physically-consistent relations has a better possibility of generalizing to future climate. This paper proposes a design-of-experiment based framework for examining the ability of SD models to generalize under approximate non-stationarity. An implicit assumption,

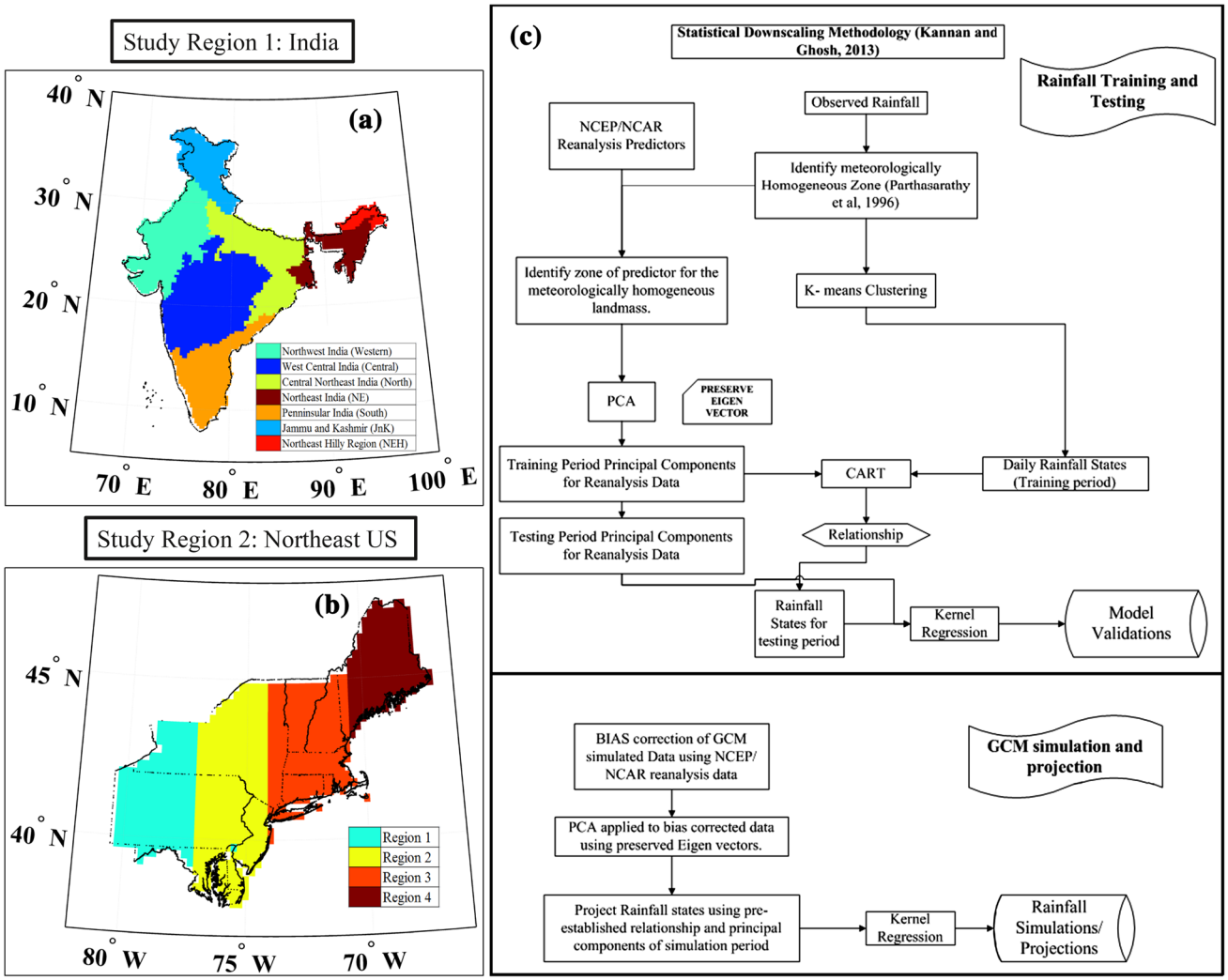


Fig. 2 Study regions and flowchart of statistical downscaling methodology. **a** 'India' with seven meteorologically homogeneous zones identified by India Meteorological Department (IMD) (Parthasarathy et al. 1996), **b** 'Northeast United States', divided into four regions of almost equal area, **c** flowchart, showing kernel regression based multisite statistical downscaling methodology by Kannan and Ghosh (2013)

which has been made by previous researchers (DelSole and Chang 2003), is that other than "drastic" changes or climate "surprises" (e.g. through runaway positive feedback), which are not usually considered by either models (GCMs or RCMs) or SD, climate conditions in the current or future will have "signatures" over time and space. For any specific region, future climate states and empirical relations may be assumed to have signatures in the past climate. Thus, if conditions get warmer (colder) or El Niño-s becomes more (less) prevalent in the future owing to global warming and different forcings such as volcanic eruptions, anthropogenic aerosols leading to changes in radiative forcing, the climate may begin to resemble those historical situations where similar conditions occurred owing to natural variability. This implicit assumption is at the core of the proposed design of experiments strategy.

For present study, guiding assumption for the strategy is that general applicability of empirical relations may be examined by carefully delineating the data into a statistical model-building or training phase followed by a test phase. The delineation will consider contrasting climate conditions at aggregate scales, such as hot versus cold years, or El Niño, La Niña, and non-ENSO years. In view of the long lead times of interest in climate studies, and the fact that climate happens to be a nonlinear dynamical system, a rigorous evaluation needs to consider the downscaled changes under plausible future scenarios. Here we identify two different time periods from observations which have resemblance in predictors to those of pre-industrial conditions and RCP 8.5, the most severe among the projected greenhouse-gas emissions scenarios. We use the Euclidean distance, between the multi-dimensional predictors of historical and PI/RCP 8.5 scenario as a measure of resemblance. SD is evaluated based on the premise that an ability to simulate and generalize under these differences is likely to translate to better ability to generalize under anticipated future non-stationarity.

The manuscript is organized as follows. Section 2 provides complete details about the downscaling methodology, study regions, and data used. Section 3 provides the details of experiments designed. Section 4 presents the results that are obtained by executing these experiments over the study regions. Section 5 summarizes the results and provides critical discussion on them. Section 6 provides the important conclusions of the present study and the future scope.

2 Data and downscaling methodology

Designing of experiments for validation of SD models, needs an established downscaling approach, to which the experiments may be applied. Here we use a recently developed non-parametric regression approach for SD (Kannan and Ghosh 2013), which has been successfully applied to Indian Summer Monsoon Rainfall (Salvi et al. 2013). The SD model with the proposed experiments is applied to two different regions, India (Fig. 2a) and North Eastern US (NEUS) (Fig. 2b). These two regions belong to two different climatic zones with different climatology. For India, the model is applied to summer monsoon (June, July, August and September) rainfall, while for NEUS it is applied to summer precipitation (June–August).

2.1 Data

The SD model develops statistical relationship between predictor (large scale or synoptic circulation) and predictand (rainfall, here). Here, we use reanalysis data for predictors and observed gridded rainfall data for the predictand. Reanalysis data consists of climatic variables, expressed in gridded format for the entire globe. In the present study, National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) data are used as predictors. NCEP/NCAR are the research organizations that use data assimilation system, which includes the NCEP global spectral model and a three-dimensional analysis scheme to obtain reanalysis data for various climatic variables. NCEP/NCAR climatic variables are categorized in three types (Kalnay et al. 1996). Category 'A' variables are those which are strongly influenced by observations e.g. zonal and meridional wind. These are highly accurate. Category 'B' variables are influenced by the model and observations but not up to the extent of category 'A' level variables e.g. specific humidity. Category 'C' variables are completely determined by the model e.g. rainfall. Additional information on types of NCEP/NCAR reanalysis variables can be found in Kalnay et al. (1996).

NCEP/NCAR reanalysis data possess positive assets such as fixed state-of-the-art assimilation scheme, inclusion of more observations, better quality control (Bromwich and Fogt 2004), better accessibility and spatial coverage across the regions of large observational data voids. However, the same data suffers drawbacks such as limited skills and reliability because of shortage of observations in high southern latitudes, presence of artificial trends in the mean sea level pressure fields near Antarctica (Hines et al. 2000), lack of sufficient skills in representing atmospheric state for certain regions and time (Stickler and Brönnimann 2011), absence of aerosol component, overestimation of seasonal precipitation, runoff, evaporation, and surface water variations resulting in poor representation of hydrologic cycle (Road and Betts 1999; Maurer et al. 2001) etc. These limitations and biases are likely to affect the results of statistical experiments, designed in the present study. However, most of the climate variables (NCEP/NCAR reanalysis data), used in the present study for obtaining rainfall simulations (refer Table 1 for list of climate variables used in

Table 1 Data used in the present study

Section 1: data used for experiment series 1

	Source	Resolution
Predictors		
Surface level specific humidity (huss)	NCEP-NCAR reanalysis data	Spatial: 2.5°, temporal: daily
Specific humidity (500 hPa) (hus)		
Sea level pressure (psl)		
Geopotential height (500 hPa) Zg		
Temperature (surface) (tas)		
Temperature (500 hPa) (ta)		
Uwind (near surface) (uas)		
Uwind (500 hPa) (ua)		
Vwind (near surface) (vas)		
Vwind (500 hPa) (va)		
Predictand		
APHRODITE rainfall data for Asia	Asian precipitation highly resolved observational data integration towards evaluation project	Spatial: 0.25°, temporal: daily
	Monsoon Asia V1101	Lat: 6.875°–37.125°
		Lon: 68.125°–97.125°
National Oceanic and Atmospheric Administration Climate Prediction Center	http://www.esrl.noaa.gov/psd/data/gridded/data.unified.html	0.25°
		Lat: 37.875°–47.375°
		Lon: 279.625°–293.125°

Time slices	Years	Comments
Training period	1951–1980	Past
Testing period	1981–2005	Recent past

Regions	Extent of zone of predictors in terms of latitude and longitude latitude	Longitude
Zone of predictors (NE US)		
Region 1	35–47.5	275–287.5
Region 2	30–50	277.5–290
Region 3	35–50	280–295
Region 4	37.5–52.5	285–300
Zone of predictors (India)		
PI	5–22.5	70–90
WCI	12.5–30	70–87.5
NWI	17.5–35	65–80
CNI	15–32.5	72.5–92.5
NE	20–32.5	85–100
JAK	27.5–40	70–82.5
NEH	25–32.5	90–100

Section 2: data used for experiment series 2

GCM used	Details	Resolution
CCCmA, CanESM2	Canadian Centre for Climate modelling and Analysis, Second Generation Earth System Models	Spatial: ~2.8°, temporal: daily

Scenarios considered for experiment	Details	Time slice considered to check signature
Pre-industrial run	Pre-industrial control simulation	2121–2150
RCP8.5	Representative concentration pathways 8.5	2070–2099

Table 1 continued

Section 3: predictors used to improve performance of SD model under non-stationary conditions		
Predictors	Source	Resolution
Surface level specific humidity (huss)	NCEP-NCAR reanalysis data	Spatial: 2.5°, temporal: daily
specific humidity (500 hPa) (hus)		
Sea level pressure (psl)		
Geopotential height (500 hPa) Zg		
Temperature (surface) (tas)		
Temperature (500 hPa) (ta)		
Uwind (near surface) (uas)		
Uwind (500 hPa) (ua)		
Vwind (near surface) (vas)		
Vwind (250 hPa) (va_250 hPa)		
Specific humidity (300 hPa) (hus_300 hPa)		
Geopotential height (250 hPa) Zg_250 hPa		
Temperature (250 hPa) (ta_250 hPa)		
Uwind (250 hPa) (ua_250 hPa)		
Vwind (250 hPa) (va_250 hPa)		
Specific humidity (850 hPa) (hus_850 hPa)		
Geopotential height (850 hPa) Zg_850 hPa		
Temperature (850 hPa) (ta_850 hPa)		
Uwind (850 hPa) (ua_850 hPa)		
Vwind (850 hPa) (va_850 hPa)		

the present study) belong to category 'A' (Kalnay et al. 1996). Also, data availability for longer period is of prime importance for execution of experiments, designed in this study. Hence, for the present study, we have selected NCEP/NCAR reanalysis daily data as the source of gridded climate variables due to its availability for longer period (1948–2014).

Predictors are coarse resolution climate variables, which represent synoptic scale circulation pattern over a study region, well simulated by climate models and are associated with precipitation process (Wilby et al. 2004). Here, we consider five surface level climate variables, viz., specific humidity, mean sea level pressure, temperature, near surface zonal (U) wind, near surface meridional (V) wind, and five pressure level climate variables (at 500 hPa) specific humidity, geopotential height, temperature, Uwind, and Vwind, obtained from NCEP/NCAR reanalysis dataset as predictors for both the study regions (Table 1). For selection of the spatial extent of the predictors (zone of predictors), we consider seven meteorological subdivisions of India (Fig. 2a), and for each subdivision the spatial extent is mentioned in Table 1. Similarly for NEUS, we divide the study area into four regions (of almost equal size) (Fig. 2b) and the spatial extents of predictors are presented in Table 1, for each region. The rationale behind selecting this spatial extent is available in Salvi et al. (2013). Selection of zone of predictor (Salvi et al. 2013) is an important step in kernel regression based SD methodology, which is used in this study. Zone of predictors is a region which completely encompasses a hydro-meteorologically homogeneous unit. It is assumed that the rainfall in a particular zone [e.g. Western Central India (WCI) zone in Fig. 2a] is mainly influenced by the climate variables in the zone of predictors. For present study, we have used a correlation based methodology for identification of zone of predictors for each unit in study regions (Salvi et al. 2013). This methodology involves obtaining two dimensional contour plots of correlation between predictor at each grid and spatially averaged observed rainfall time series in a particular zone. Method of selection of the zone of predictors (Salvi et al. 2013) for a particular hydro-meteorologically homogeneous unit is explained here with an example of WCI zone (Fig. 2a). WCI is composed of 1273 grids at 0.25 degree resolution and its spatial extent is over latitude 15°–25° and longitude 72°–84° approximately. We select a bigger region (latitude 5°–40° and longitude 60°–120° here) in such a way that WCI is well within this region. All the NCEP/NCAR reanalysis data climate variables (at 2.5° resolution), which are used in the study as predictors, are extracted over this region. Hence, we get a 15 × 25 grid size over latitude 5°–40° and longitude 60°–120° at 2.5° resolution for each predictor. Following steps are executed to decide zone of predictors.

1. Data, used to carry out the analysis (1) observed rainfall data matrix over WCI zone (dimensions: N × 1273, where N is the sample size and 1273 are number of grids in WCI zone), (2) NCEP/NCAR climate predictor e.g. temperature data matrix (dimension: N × 15 × 25);
2. Obtain spatially averaged rainfall time series 'Y' (average of 1273 rainfall values for each day) for WCI zone (Dimensions of spatially averaged rainfall time series will be N × 1).
3. Compute the correlation coefficient between 'Y' and temperature time series at each node from temperature data matrix. Outcome of this step will be correlation matrix with dimensions 15 × 25.
4. Plot the matrix in the form of contours.
5. Repeat the same procedure with each NCEP/NCAR climate predictor, used in the study.
6. Visually compare all the contour plots and decide the region as 'zone of predictors' in such a way that for most of the climate predictors, high correlation contour regions lie inside the zone of predictors.

Supplementary Figures 1 and 2 (refer Online Resource 1, SF1 and SF2) show the correlation contour plots for one representative zone from each study region. The predictand, which is the gridded rainfall data, is obtained for India, from APHRODITE (Yatagai et al. 2009). The downloaded data is for 'Monsoon Asia (MA)' V1101, from 60°E to 150°E (longitude) and 15°S to 55°N (latitude) and over a period of 1951–2005. APHROTIDE data is a widely used gridded rainfall product (Wang and Gillies 2013; Han and Zhou 2012; Gillies et al. 2012; Tojo et al. 2011). For NEUS, the gridded rainfall data, provided by National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC), is used as predictand. The data is at 0.25° resolution and it is available for entire US.

2.2 Statistical downscaling methodology

Here, we follow the SD model developed by Kannan and Ghosh (2013), which is based on Classification and Regression Tree (CART) and kernel regression. This model is reported to simulate well, the climatology, statistics of rainfall and cross-correlation between rainfalls at multiple stations. The steps involved in developing the predictor-predictand relationship are Principal Component Analysis, K-means clustering, CART and kernel regression. Figure 2c shows the flowchart of the methodology. The mathematical operations shown in the Figure are elaborated briefly in the following text. For detailed methodology, Kannan and Ghosh (2013), Salvi et al. (2013) can be referred.

2.2.1 *K-means clustering*

In the SD model (Kannan and Ghosh 2013); 'k means clustering' technique is used to present the spatial pattern of rainfall in a region, with a cluster. Each of the meteorologically homogeneous regions is considered as individual unit and k mean clustering is applied over all the grids, located in that region. The primary purpose is to categorize the rainfall in a single zone (at different grids) on a given day, into a state, which represents a specific spatial rainfall pattern. The number of classes/categories is identified based on Dunn's index and Silhoutte index. This operation is applied to the gridded rainfall.

2.2.2 *Principal Component Analysis (PCA)*

While establishing statistical relationship between predictors and predictand, use of correlated multi-dimensional predictors may lead to multicollinearity. In the present study, Principal Component Analysis (PCA) is used to tackle this problem. It consists of orthogonal transformations which transforms the predictor matrix into principal components such that there is no/less correlation among principal components. Based on the percentage of total variability of predictors, explained by principal components, they are selected for regression. Here, we consider the principal components, which collectively present ~95 % of the total variability.

2.2.3 *Classification and regression tree (CART)*

In the present study, CART analysis is used to establish statistical relationship between principal components (obtained by performing PCA on climate variables) and rainfall state (obtained by applying k means clustering on zonal rainfall matrix). It comes under the category of decision tree learning technique. The details regarding CART analysis are presented in Kannan and Ghosh (2013). The developed relationship, when applied to principal components from daily predictor field, results into the rainfall class of the region for that day.

2.2.4 *Kernel regression*

Kernel regression is a non-parametric statistical regression technique, which is generally a smoothening filter that simulates the predictand for a desired predictor data point by applying weights to the other points lying in the neighboring region of the desired one. Generally a weight function is deployed to fulfill this task, which assigns heavy weights to the nearby data points and very low weights to the points which are far away. Here, we use Nadaraya and Watson estimator (Nadaraya 1964; Watson 1964) for this purpose.

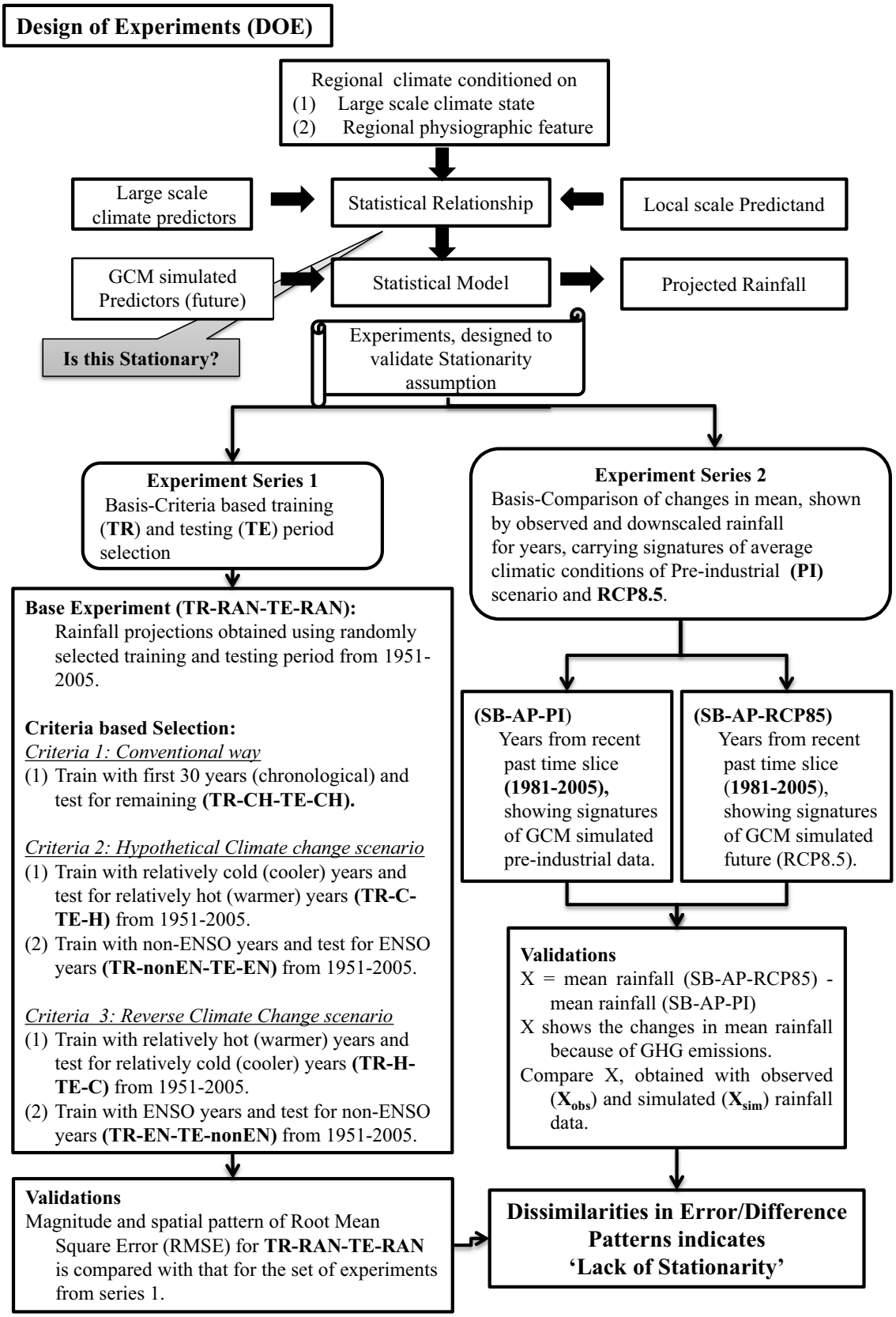


Fig. 3 Generic design of experiments for testing validity of assumption of stationarity in statistical relationship under changing climatic conditions

The kernel regression is performed on the principal components that are obtained from predictors, conditional on the CART derived class and rainfall values are obtained at multiple sites. The cluster/class represents the spatial pattern of rainfall in a zone and hence, it is possible to preserve the cross correlation across rainfall at multiple sites (Kannan and Ghosh 2013).

2.2.5 Rainfall projections using GCM simulations

The establishment of statistical relationship is achieved using NCEP-NCAR reanalysis data as predictors and observed rainfall as predictand. When the rainfall needs to be simulated for historic period or projected for future, the pre-established relationship is applied to GCM simulated predictors. GCM simulated predictors cannot be directly used due to presence of bias with respect to reanalysis data. It is necessary to remove bias from GCM simulations before using GCM data. The developed relationship is then applied to bias corrected predictors for historical and future simulations. In this work, we are not simulating downscaled rainfall for future, but focusing on designing the validation experiments with historical downscaled simulations and this should be performed before applying the relationship to future.

3 Designs of experiments

In this present study, we design two set of experiments, for validating the assumption of stationarity, in predictor-predictand relationship, in a SD model. These designs of experiments (DOE) are based on the performance of SD model, for the period, when the synoptic scale circulation pattern is different from those, which are present in calibration/training data. The satisfactory performance is expected by a model, when its formulation is robust to capture future relationship or the climate change signal is weak (future conditions are not significantly different from the past training period). This is the basis for the first series of experiments. Second rationale behind DOE is based on differences between the rainfall projections for years, which are similar to the most severe GCM future projections (RCP8.5) and GCM preindustrial simulations. These differences may be considered as the possible signatures of climate change on rainfall. Here, in the second series of experiments, we evaluate the SD models based on its performances in simulating these differences. Motivation for second series of experiments lies in gauging the capability of SD models to capture changes in mean rainfall because of GHG emissions.

The two rationales discussed, form the basis of designing two different series of experiments viz., (1) *Experiment Series 1* Training period selection based on predefined criteria, and (2) *Experiment Series 2* Comparison of differences in mean rainfall between the years similar to preindustrial run and future period. Figure 3 shows the pictorial representation of design of experiments.

3.1 Experiment series 1: criteria based training period selection

This category consists of experiments, which are designed to validate the performance of statistical relationship, established for different training conditions. The conventional way of identifying training and testing period for a model is simply based on chronology i.e. first 'x' years are treated as training and next 'y' years are testing period (Kannan and Ghosh 2013; Salvi et al. 2013). Validation of the statistical models is important and the established statistical relationships are validated during a period independent from the calibration period. Different validation procedures such as split-sampling (Busuioc et al. 1999), cross validation (Murphy 1999), stratified validation (Wilks 1999), statistical model ensembles (Hertig and Jacobeit 2008) etc. are proposed in literature. Here, three additional (additional to the conventional way of chronological selection of training and testing period) ways of identifying training and testing period are proposed viz. random selection of training and testing period, using cold years for training and hot years for testing (and vice versa), and using non El-Niño years as training and El-Niño years as testing (and vice versa). The criteria for random selection of training period, ensures complete mixing of circulation patterns from different climate conditions and hence the model gets better exposure to a wide range of climate variability. This experiment is treated as 'base experiment' and is referred as 'TR-RAN-TE-RAN' (training and testing days are selected randomly) in the present study. The results of all other experiments (belonging to series 1) will be compared with the results of TR-RAN-TE-RAN to identify the regions which lack stationarity. The last two criteria are derived from the hypothetical expected climate change signals and reverse of climate change signal respectively. The rise in global temperature leads to a warmer future (IPCC 2007). The experiment of training the model with colder years and testing it for hot years is a way of evaluating a model under changed climatic conditions, where the model is calibrated with historic climate scenario (relatively colder) and validated with changed conditions (warmer conditions).

Similarly the possible increase in the frequency of El-Niño Southern Oscillations (ENSO) under the influence of GHG emissions (Timmermann et al. 1999) led to idea of third experiment in which training period consists of non El-Niño years and validation period consists of El-Niño years. Recent report by Intergovernmental Panel on

Climate Change (IPCC 2014) mentions different aspects of ENSO phenomenon such as inter-decadal modulations in amplitude and spatial pattern within the instrumental record. There are little consensus on whether the observed changes in ENSO are due to external forcing or natural variability. This questions the outcome of the study by Timmermann et al. (1999), which proposes increase in the frequency of ENSO events under GHG emissions. Nevertheless, ENSO is the dominant mode of inter-annual variability and likely to influence regional scale rainfall due to changes in moisture availability (IPCC 2014). The premise behind all the 'series 1' experiments is to test the ability of the model for those climatic conditions, which differ from training period climatic conditions to a great extent. Hence, even if the changes in the frequency of ENSO events under GHG emissions is debatable, considering possible influences of ENSO on regional scale rainfall and disjoint climatic conditions at the time of ENSO and non-ENSO event, we include this as part of series 1 experiment. The details about each experiment are discussed in following sections.

3.1.1 Base experiment (TR-RAN-TE-RAN)

TR-RAN-TE-RAN represents categorizing the overall time slice (1951–2005) into training and testing period at daily scale randomly. A uniform random number is generated for each day in the time slice. If the generated random number value is less than or equal to ~0.55, that day will be categorized as training period day and if random number is greater than ~0.55, the day will be treated as testing period day. Threshold value for the random number (~0.55) is fixed in such a way that all the days, which are categorized under the heading of 'training period', together will form a time slice of 30 years and the validation period data will form a time slice of 25 years. In the experiment TR-RAN-TE-RAN, the classification of data as training and testing period ensures complete mixing of synoptic scale atmospheric circulation patterns.

3.1.2 Training period selection based on chronological order (TR-CH-TE-CH)

Conventionally, in a SD model, first 'X' year time slice is recognized as training period and the remaining is considered as testing period. The experiment TR-CH-TE-CH represents selection of training period based on chronological order. For this experiment, first 30 years (1951–1980) are considered as training period and next 25 years (1981–2005) are considered as testing. Here onwards, the period of first 30 years will be referred to as 'Past' and next 25 years will be referred as 'Recent Past'.

3.1.3 Training period selection based on hot and cold years (TR-C-TE-Hand TR-H-TE-C)

This set of experiments involves identifying relatively warmer/colder years from the time slice 1951-2005 as training set. Its complimentary subset from 1951 to 2005, i.e. relatively colder years (in case of warmer years as training) or warmer year (in case of colder years as training) is considered as testing period. General methodology to identify relatively warmer/colder years is described below. Consider a study region, where temperature data is available (at any temporal resolution e.g. daily/monthly) for 'T' years and at 'N' number of stations or 'N' grids (in case of gridded data).

1. Obtain spatio-temporal average of temperature data for each year. In case of daily data, the data available for '1 year' will have dimensions 365 × N and in case of monthly data, the dimension will be 12 × N.
2. Above step will provide a single spatio-temporally averaged temperature value for a particular year. Repeat this step for each year and generate annual temperature time series (dimensions: No. of years × 1).
3. Arrange this time series in ascending order. First X years will be relatively colder years as compared to other. These years will be considered as training period for experiment TR-C-TE-H. The same step can be repeated by arranging the time series in descending order. First X years will be relatively warmer years and will be considered as training period for experiment TR-H-TE-C. This methodology is followed for obtaining relatively warmer and colder years for the two study regions viz. India and NEUS. For the present study, relatively warmer/colder years over the study regions are not decided based on the temperature data for entire year (as discussed in the methodology). We rather use temperature data over summer months for deciding whether a particular year is relatively warmer/colder. Premise behind selecting summer months for India [March, April, and May (MAM)] and NEUS (Jun, July, and Aug) is that, these months together constitute relatively high temperature periods as compared to the other months of the year and can be considered as a realistic criteria for categorizing a year as relatively warmer/colder. In addition to this, 'MAM' months for India are considered as pre-monsoon season. Literature suggest that pre monsoon temperature over land region affects land sea thermal gradient which is a causal factor for monsoon variability (Gautam et al. 2009) and hence hydrological cycle (Bernett et al. 2005). In case of NEUS, rainfall projections are obtained over summer months June, July, and August, hence, it is ideal to use the average temperature over these months for identifying relatively warmer/colder years.

For India, relatively warmer and colder years are decided based on spatially and temporally averaged pre-monsoon temperature over the months MAM (source: Indian Institute of Tropical Meteorology, Pune, India, website: ftp://www.tropmet.res.in/pub/data/txtn/NEW-TNREGION.TXT and ftp://www.tropmet.res.in/pub/data/txtn/NEW-TXREGION.TXT). The two links, mentioned above provide monthly data for minimum temperature and maximum temperature for entire India. The detailed calculation for obtaining annual averaged time series for pre-monsoon temperature is illustrated in Online Resource 2 (refer Online Resource 2).

For NEUS, temperature data at monthly temporal scale is obtained from 'Carbon Dioxide Information Analysis Center' (CDIAC), website. CDIAC is the primary climate-change data and information analysis center of the U.S. Department of Energy (DOE). CDIAC is located at DOE's Oak Ridge National Laboratory (ORNL) and includes the World Data Center for Atmospheric Trace Gases. Monthly station level temperature data for US is available on CDIAC website with ORNL domain (link: http://cdiac.ornl.gov/ftp/ushcn_v2.5_monthly/through_2012/). Monthly temperature data for the months June, July, and August for all the stations in NEUS is extracted from the file. The list of state-wise station IDs is also available on CDIAC website (link: http://cdiac.ornl.gov/ftp/ushcn_daily/ushcn-stations.txt). All the stations, which are located in the study region NEUS are selected for obtaining relatively cold/warmer years. The detailed calculation of spatially averaged annual time series is shown in Online resource 3 (refer Online Resource 3). Missing values in the temperature data (pre-filled with −9999) are less than ~3 % of the overall data and hence, neglected while calculating spatially averaged value. Relatively warmer years are referred to as 'hot' years and relatively colder years are referred to as 'cold' years in this manuscript. Years, which are categorized as hot/cold are different for different study regions (except for some years, which are common).

3.1.4 TR-nonEN-TE-EN and TR-EN-TE-nonEN

ENSO is a leading mode of inter-annual climate variability originating in the tropical Pacific Ocean. The Sea Surface Temperature (SST) anomalies in the equatorial Pacific Ocean can have remote effects on climate globally (Langenbrunner and Neelin 2013). ENSO phases have been linked to Indian summer monsoon. ENSO, the largest known climatic forcing of inter-annual monsoon variability, shows significant influence on ISMR (Krishna Kumar et al. 1999). Even though, in recent years this relationship seems to be weakening (Krishna Kumar et al. 1999), the effect of El-Niño on ISMR cannot be fully denied. With more emission of greenhouse gases, the frequency of occurrences of El Niño may increase (Timmermann et al. 1999). Hence, the third set of complimentary experiments consists of training period selection based on the occurrence of ENSO in a year. The list of El-Niño and non El-Niño years is obtained from Chou et al. (2002). This set consists of a pair of experiments which are complimentary to each other viz. TR-nonEN-TE-EN and TR-EN-TE-nonEN. TR-EN-TE-nonEN represents selection of El Niño years from the time slice 1951–2005 as training period and TR-nonEN-TE-EN represents selection of non-El Niño years from the time slice 1951–2005 as training period. Similar to the experiment TR-C-TE-H, TR-nonEN-TE-EN is an experimental simulation of hypothetical future climate change signal. The list of years that are selected as training period for different experiments is illustrated in Table 2.

Table 2 shows that there are more years (37) pertaining to the training period for the experiment TR-nonEN-TE-EN cases than for the experiment TR-EN-TE-nonEN cases (18). Also, there are instances, where an El Niño event is immediately followed by a La Niña event and vice versa, indicating the El Niño and La Niña years are not widely separated. Significant difference in the number of training years for both experiments is likely to have an impact on the accuracy of the downscaled product. These are the limitations emanating from the design of this experiment.

3.2 Experiment series 2: validations based on expected signature of climate changes

The SD models, showing good skills in capturing historic climatology of observed rainfall, may not necessarily show the same level of proficiency in simulating changed climate. Standard evaluation procedures do not consider this criterion; though the credibility of future projections by a SD model depends on it. Here, we consider this criterion in our DOE, with an indirect 'signature based' approach (Fig. 4). We use two CMIP5 scenarios viz. (1) 'Pre-industrial run' which corresponds to 'no anthropogenic GHG emissions' and (2) RCP8.5 which corresponds to 'Highest GHG emissions'. Pre-industrial run is an unforced run which serves as the baseline for analysis of historical and future runs (Taylor et al. 2012). RCP8.5 corresponds to the strongest radiative forcing, reaching 8.5 Wm^{-2} at the end of 2100 (Taylor et al. 2012). Here, we identify signature (based on synoptic scale circulation) of 'pre-industrial scenario' (PI) and future 'RCP8.5 scenario' predictors in recent past (1981–2005) time period. For individual year, from the recent past (1981–2005), the Euclidean distance is computed for centroid of predictor field between that year and the selected scenario (PI or RCP8.5). Years having predictor field close to these two different scenarios, based on Euclidean distance, form two different subsets, each of 15 years. The subset, close to 'average climatic conditions

Table 2 Different time slices used to calibrate the model for experiment series 1

TR-RAN-TE-RAN	TR-C-TE-H India	TR-H-T-C	TR-C-TE-H NEUS	TR-H-T-C	TR-nonEN TE-EN	TR-EN-TE-nonEN
1951	1953	1982	1992	2005	1951	1957
1952	2002	1965	1982	1955	1952	1958
1953	2004	1971	2000	1973	1953	1965
1954	1999	1983	1958	1995	1954	1966
1955	1958	1990	1965	1952	1955	1968
1956	1973	1957	1985	1999	1956	1969
1957	1954	1997	1972	1988	1959	1972
1958	1985	1967	1962	1959	1960	1973
1959	1969	1963	1986	2002	1961	1982
1960	1980	1951	2004	1991	1962	1983
1961	1956	1968	1964	1983	1963	1986
1962	1964	1987	1956	1994	1964	1987
1963	1959	1976	1954	1993	1967	1991
1964	1984	1979	1963	1975	1970	1992
1965	1996	1993	1960	1987	1971	1994
1966	1961	1992	1979	2001	1974	1995
1967	1974	1995	1951	1966	1975	1997
1968	1994	1986	1997	1953	1976	1998
1969	2003	1989	1977	1984	1977	
1970	1988	1960	1974	1980	1978	
1971	2001	1978	1971	1990	1979	
1972	1998	1981	1976	2003	1980	
1973	1970	1962	1978	1970	1981	
1974	1952	1977	1968	1967	1984	
1975	1975	1955	1996	1981	1985	
1976	2000	1972	1961	1998	1988	
1977	1966	1991	1957	1969	1989	
1978	2005	2005	1989	1989	1990	
1979	1991	1966	1969	1957	1993	
1980	1972	1952	1998	1961	1996	
					1999	
					2000	
					2001	
					2002	
					2003	
					2004	
					2005	

pre-industrial (ACC PI)' is named as 'ACC-PI signatures' and the other subset, close to 'average climatic conditions RCP8.5 (2070–2099) (ACC RCP85)', is named as 'ACC-RCP85 signatures'. The difference in mean observed rainfall for 'ACC-PI signatures' and 'ACC-RCP85 signatures' represents the possible signature of GHG emissions to rainfall. Similar difference is computed for the projected rainfall and results are compared. The criteria of evaluation, here is the skill of downscaling model in simulating this difference. Higher skill of a model in simulating this difference indicates higher credibility in projections. Here we use Canadian (CanESM) GCM for the PI and RCP8.5 circulation patterns.

An important implicit assumption behind the approach is that the years, relatively closer to average climatic conditions of a scenario e.g. RCP8.5, have more signatures of RCP8.5, compared to other years. For measuring the closeness, we have adopted Euclidean distance, and this can be

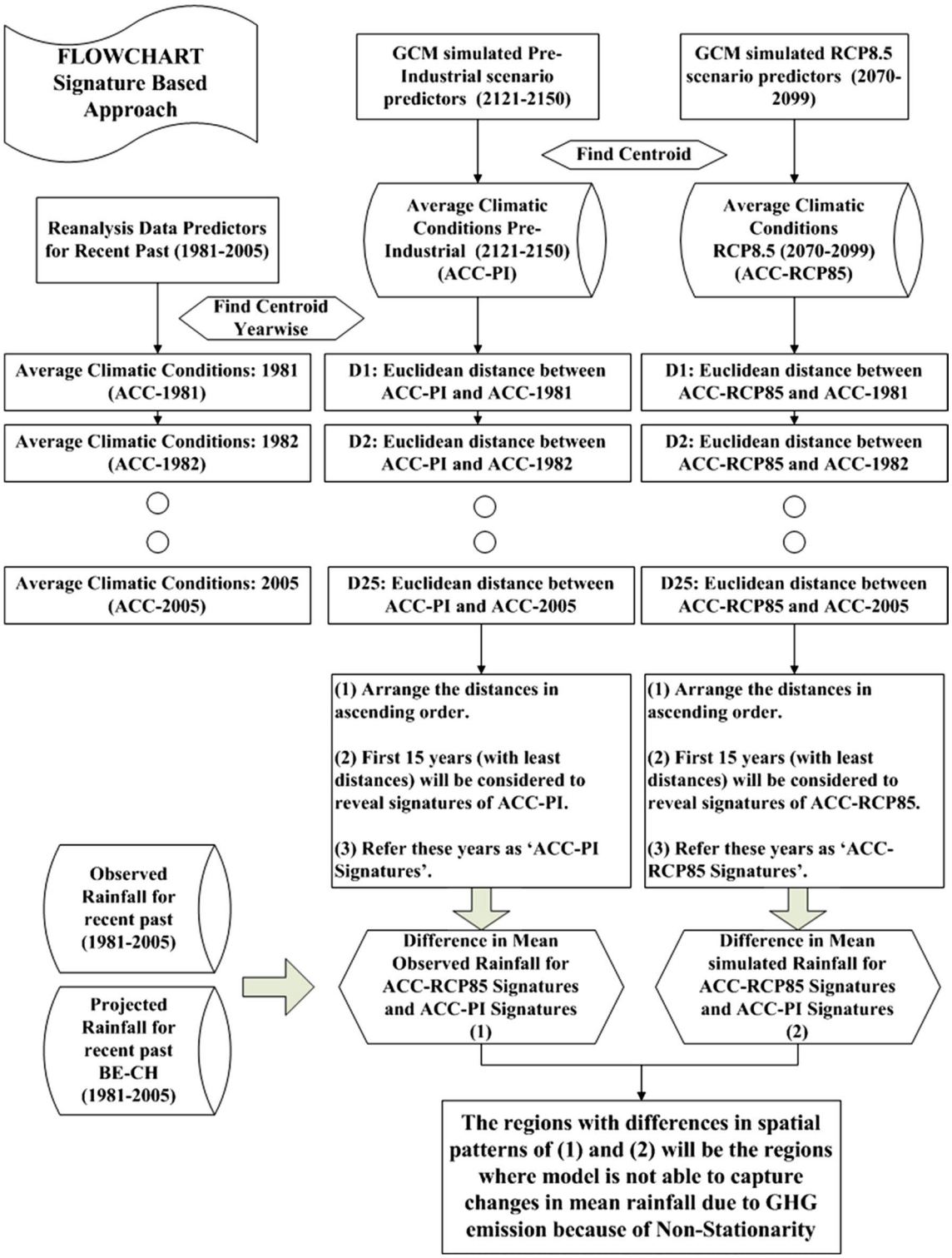


Fig. 4 Flowchart of signature based approach designed for investigating the performance of statistical downscaling methodology under changing climatic conditions (experiment series 2)

improved further with better statistical methods to identify analogues (e.g. Local dynamical analogs, Li and Ding 2011).

The capability of the downscaling model to capture the expected difference in mean rainfall due to GHG emissions is tested with the experiments SB-AP-PI (with

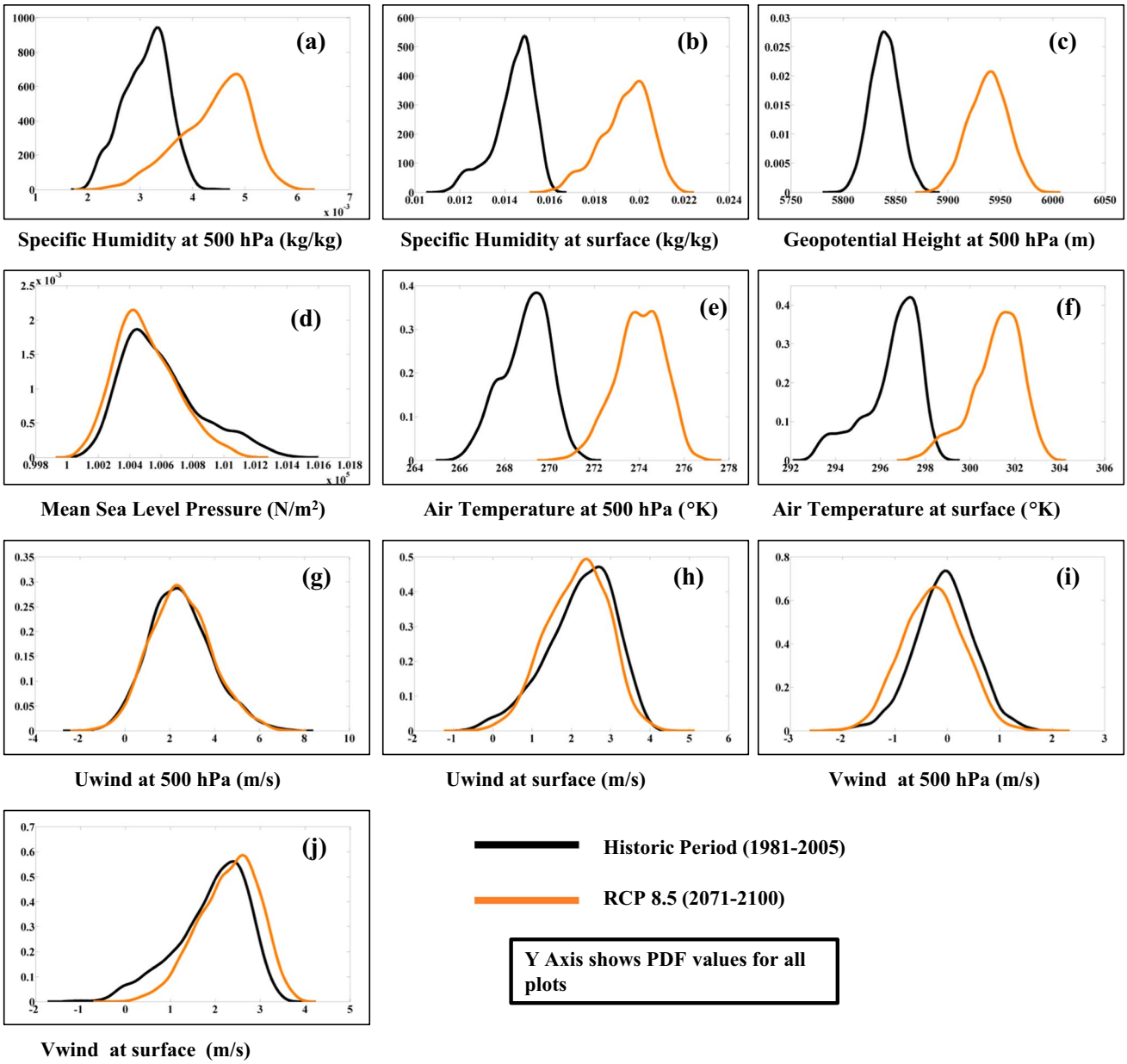


Fig. 5 Probability density function plots for comparison of predictors, specific humidity at 500 hPa (**a**), surface level specific humidity (**b**), Geopotential height at 500 hPa (**c**), mean sea level pressure (**d**), air temperature at 500 hPa (**e**), surface level air temperature (**f**), zonal (U) wind at 500 hPa (**g**), surface level zonal (U) wind (**h**), meridional (V) wind at 500 hPa (**i**), and surface level meridional (V) wind (**j**), simulated for 'Historical' and 'RCP8.5' scenarios over India

'ACC-PI signatures') and SB-AP-RCP85 (with 'ACC-RCP85 signatures'). The signature based approach, developed to achieve this, involves obtaining average climatic conditions for pre-industrial scenario and RCP8.5. The average climatic conditions are obtained with those variables, which show significant climate change from historical period to RCP8.5. In order to identify the predictors showing significant climate change, probability density functions (PDFs) of the predictors are obtained over historical period and RCP8.5 and compared using two sample Kolmogorov–Smirnov test (two sample K–S test) in order to identify whether they belong to same probability distribution. Null hypothesis states that the data in historic period time series and RCP8.5 time series belong to the same distribution (no significant climate change). The predictors for which the null hypothesis is rejected at 5 % significance level are considered as the predictors undergoing significant climate change. The stepwise description of this procedure is as follows (which is the same for India and NEUS).

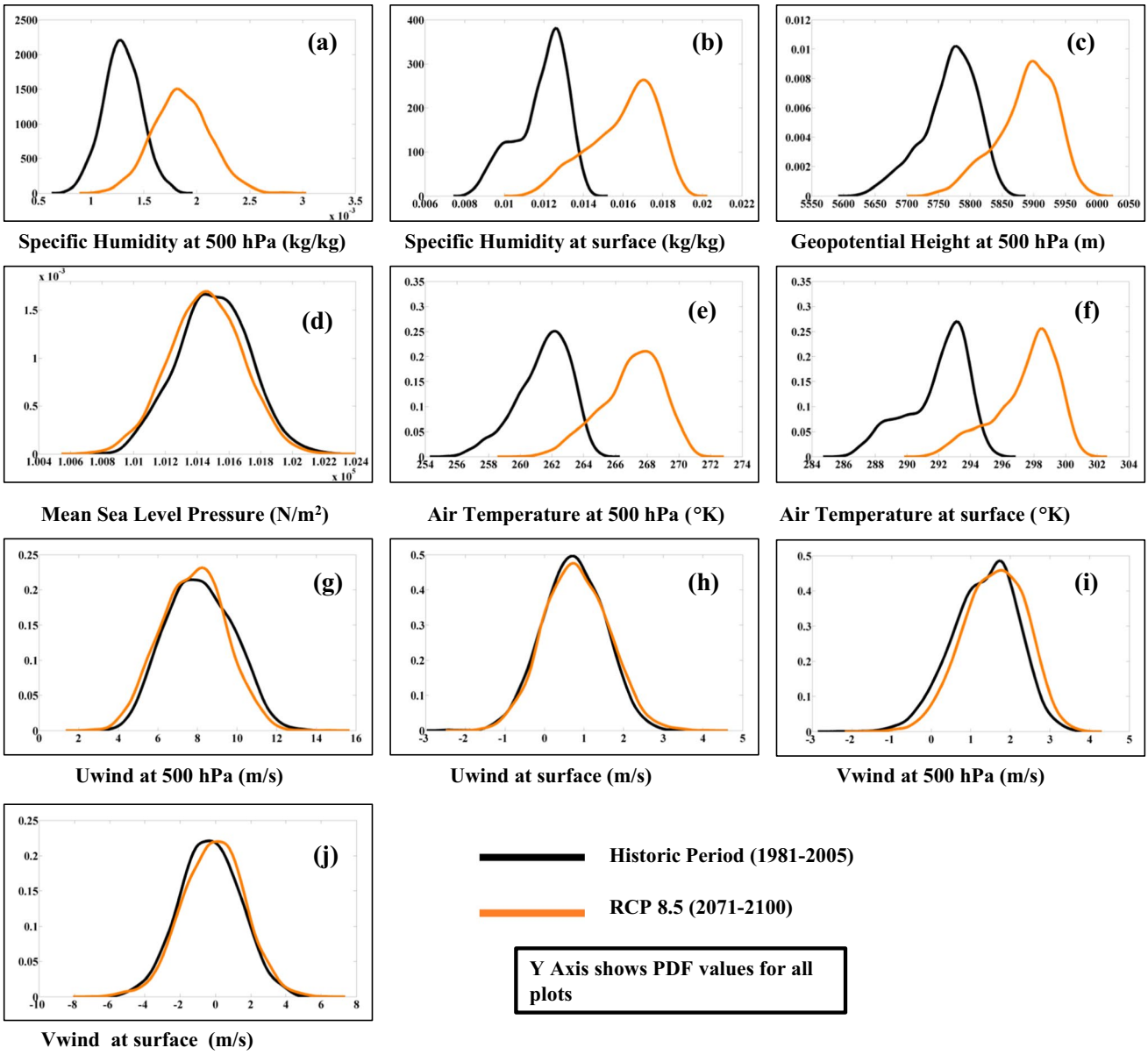


Fig. 6 Probability density function plots for comparison of predictors, specific humidity at 500 hPa (**a**), surface level specific humidity (**b**), Geopotential height at 500 hPa (**c**), mean sea level pressure (**d**), air temperature at 500 hPa (**e**), surface level air temperature (**f**), zonal (U) wind at 500 hPa (**g**), surface level zonal (U) wind (**h**), meridional (V) wind at 500 hPa (**i**), and surface level meridional (V) wind (**j**), simulated for 'Historical' and 'RCP8.5' scenarios over NEUS

1. Select an area around the study region (completely encompassing the study region). For present study, following two areas are selected.

India: latitude 5 N–40 N, longitude 60E–120E; NEUS: latitude 25 N–60 N, longitude 270E–310E.

2. Obtain GCM simulated data for predictors over these areas (for both scenarios historic and RCP8.5).
3. Remove systematic error from GCM simulated data using bias correction methodology at each grid. In this study, quantile based mapping methodology by Li et al. (2010) is used for removing bias from GCM simulations.
4. Obtain spatially averaged time series using bias corrected GCM simulations over historic period and future (RCP8.5). After this step, each predictor will have two spatially averaged time series for historic scenario and RCP8.5 scenario.
5. Fit a probability distribution to both the data and obtain PDF. In the present study, non-parametric kernel probability density function is used.

Table 3 Experiment identification (IDs) with their details

Experiment ID	Details
TR-CH-TE-CH	Training period selection in chronological order (first X years as training period, next Y years as testing period) Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training period: 1951–1980 Testing period: 1981–2005
TR-RAN-TE-RAN	Training period selection based on random number generation (ensuring complete mixing of past and recent past data) Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training/testing period: decided based on generation of uniform random number between 0, 1 Training period: 0 < random number < 0.5430 for India and 0.5515 for US Testing period: random number > 0.5430 for India and 0.5515 for US
TR-C-TE-H	Cold years as training period Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training period: relatively cooler years Testing period: remaining years from the set (1951–2005)
TR-nonEN-TE-EN	Non Enso years as training period (Non- El Nino years) Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training period: non-EL-NINO years Testing period: remaining years from the set (1951–2005)
TR-H-TE-C	Hot years as training period Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training period: relatively warm years pre-monsoon temperature for March, April and May Testing period: remaining years from the set (1951–2005)
TR-EN-TE-nonEN	BE-El: Non Enso years as Training period (Non- El Nino years) Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training period: non-EL-NINO years Testing period: remaining years from the set (1951–2005)
SB-AP-PI, SB-AP-RCP8.5	SB-AP-PI: signature based approach, years close to pre-industrial (PI) run period for Non-overlapping predictors SB-AP-RCP85: signature based approach, years close to RCP8.5 period for Non-overlapping predictors Predictors: NCEP/NCAR reanalysis data Predictand: APHRODITE rainfall data (0.25° resolution) Training period: 1951–1980 PI run period: 2121–2150 Testing period: 1981–2005 RCP8.5: 2070–2099 Details: the years from the testing period (15 out of 25 years), for observed and simulated rainfall, are found out which are close to the centroid of PI run predictors (simulated by GCM) based on Euclidean distance. Similar exercise is carried out with the centroid of RCP8.5 (2070–2099) predictors. The spatial pattern of difference in mean, obtained for observed and simulated data are compared to check whether the methodology is capturing the changes in the scenarios.

6. Apply two sample K–S test and identify whether the predictor undergoes a change from historic runs to RCP8.5, at 5 % significance level.

Figures 5 (for India) and 6 (for NEUS) show comparison of PDFs between different predictors that are simulated over historic period (1981–2005) and future (RCP8.5, 2071–2100). PDF comparison plots for predictors such as specific humidity at 500 hPa (Figs. 5a, 6a), surface level specific humidity (Figs. 5b, 6b), geopotential height at 500 hPa (Figs. 5c, 6c), air temperature at 500 hPa (Figs. 5e, 6e), and surface level air temperature (Figs. 5f, 6f) show visually significant shift, whereas the PDF comparison for the predictors such as mean sea level pressure (Figs. 5d, 6d), U wind at 500 hPa (Figs. 5g, 6g), surface level U wind (Figs. 5h, 6h), V wind at 500 hPa (Figs. 5i, 6i), surface level V wind (Figs. 5j, 6j) show overlapping nature. However, application of two sample K–S test for the predictors over India and NEUS revealed that all ten predictors show significant climate change at 5 % significance level.

Consolidated information about different experiments that are designed to test assumption of stationarity is tabulated in Table 3.

Along with the experiments, designed in the present study, we also apply two methods developed by Duan and McIntyre (2012) and Hertig and Jacobeit (2013) for

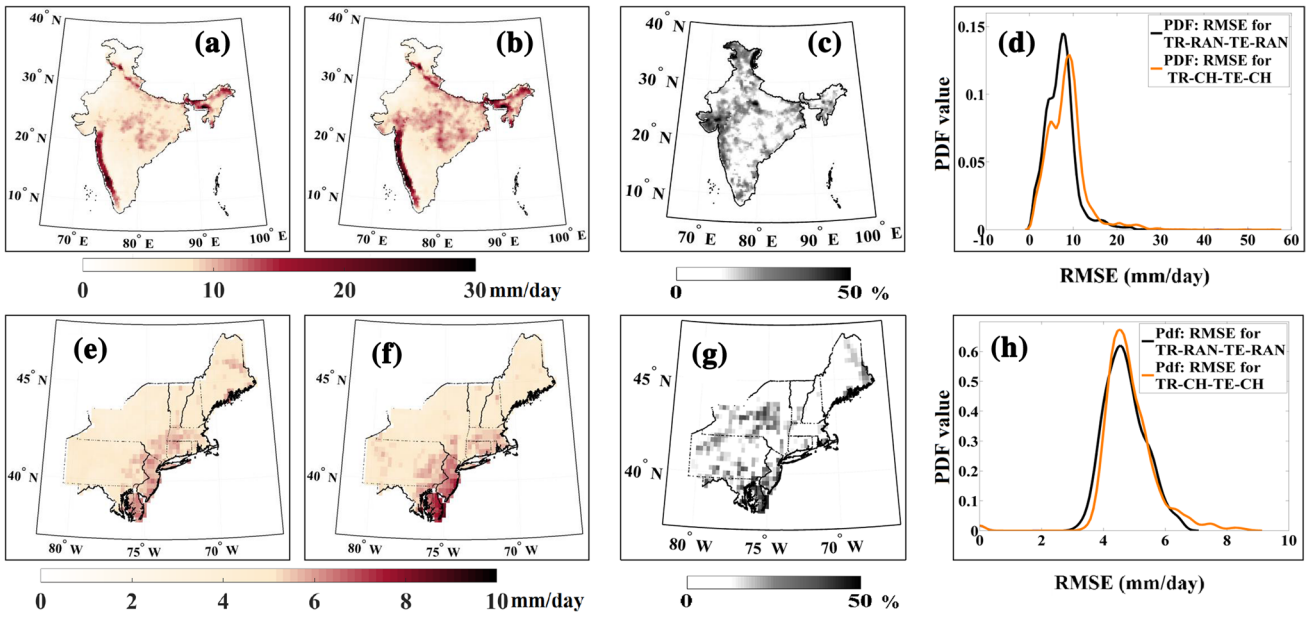


Fig. 7 Comparison between outcomes of experiments TR-CH-TE-CHand base experiment (TR-RAN-TE-RAN) over India and NEUS, in terms of magnitude and spatial distribution of RMSE. For India, **a** RMSE of experiment TR-RAN-TE-RAN, **b** RMSE of experiment TR-CH-TE-CH, **c** grid-wise absolute percentage difference between (**a**) and (**b**) with (**a**) as reference, **d** Comparison between RMSE of TR-RAN-TE-RAN and TR-CH-TE-CH, represented as PDFs. For NEUS, **e** RMSE of experiment TR-RAN-TE-RAN, **f** RMSE of experiment TR-CH-TE-CH, **g** grid-wise absolute percentage difference between (**e**) and (**f**) with (**e**) as reference, **h** comparison between RMSE of TR-RAN-TE-RAN and TR-CH-TE-CH, represented as PDFs

demonstration purpose. In both the methods, training/calibration of the model is performed for multiple overlapping time slices, obtained by applying a moving window to the analysis period. Significant trends in the regression coefficients associated with the predictors (Duan and McIntyre 2012) or significant differences in error metrics (Hertig and Jacobeit 2013); obtained by applying the model to multiple time slices; indicate violation of assumption of stationarity.

4 Results

The experiments, detailed in the previous sections, are executed over the study regions. The approach by Duan and McIntyre (2012) involves regression based SD model with 30 year moving window time slices e.g. 1951–1980, 1952–1981…, where Box–Cox transformed rainfall data at each grid is used as predictand. Statistically significant changes in regression coefficient values with time indicate non-stationarity. Supplementary figure 3 (refer Online Resource 1, SF3) shows the results after applying the methodology, developed by (2012) to the Indian landmass. It is observed that the assumption of stationarity is violated over different regions such as West Central India (WCI), the northern part of India, the Western Ghats, south east coastal region, some part of Gujarat and Jammu and Kashmir (JAK). Temperature, V wind and geopotential height (all at 500 hPa) are the predictors for which the assumption of stationarity gets violated over a relatively larger region as compared to other predictors.

Similarly, we also apply the methodology developed by Hertig and Jacobeit (2013) to Indian landmass, with RMSE as the performance metric. The premise behind demonstration of these methodologies is to provide glimpse of the recent literature on some of the ways in which the violation of assumption of stationarity is assessed and not the comparison. Hence, we have not demonstrated the application of bootstrap approach while implementing the methodology, developed by Hertig and Jacobait (2013). Rainfall simulations are obtained with different calibration periods e.g. 1951–1981, 1952–1982… with 31 year moving window and RMSE for different simulations are compared. The RMSE patterns look similar for all the 30 windows obtained from the analysis period (refer Online Resource 1, Supplementary Figure 4) except for the northern India viz. JAK and the western part of Rajasthan. These are the regions, where stationarity assumption is getting violated. The spatial pattern of RMSE for other regions is similar, indicating persistence of stationarity.

The results obtained by executing statistical experiments (designed in the present study), are discussed in the following subsections.

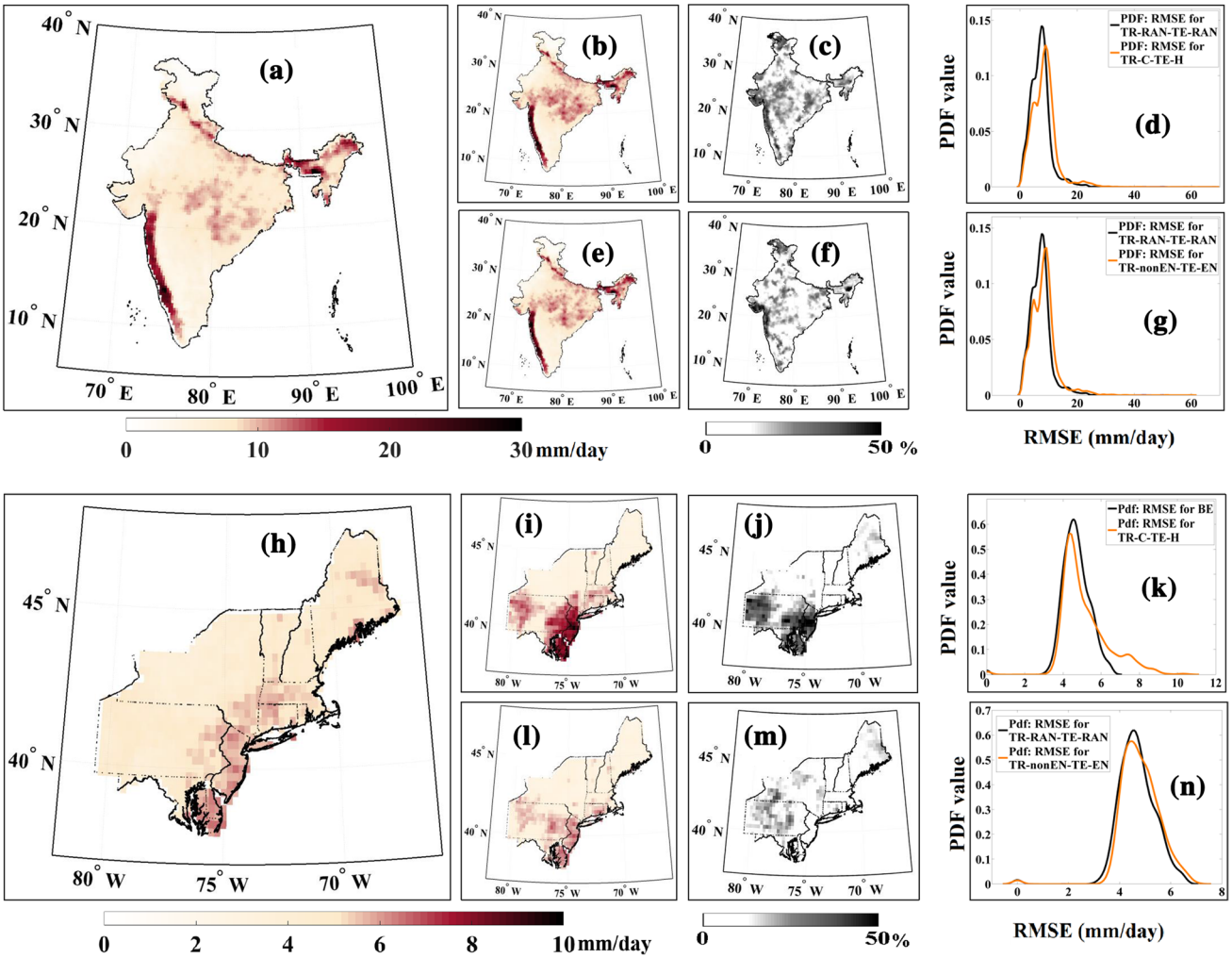


Fig. 8 Comparison between outcomes of 'Hypothetical climate change scenario' experiments (TR-C-TE-H, and TR-nonEN-TE-EN) and base experiment (TR-RAN-TE-RAN) over India and NEUS, in terms of magnitude and spatial distribution of RMSE. For India, **a** RMSE of experiment TR-RAN-TE-RAN, **b** RMSE of experiment TR-C-TE-H, **c** grid-wise absolute percentage difference between (**a**) and (**b**) with (**a**) as reference, **d** comparison between RMSE of TR-RAN-TE-RAN and TR-C-TE-H, represented as PDFs, **e** RMSE of experiment TR-nonEN-TE-EN, **f** grid-wise absolute percentage difference between (**a**) and (**e**) with (**a**) as reference, **g** comparison between RMSE of TR-RAN-TE-RAN and TR-nonEN-TE-EN, represented as PDFs. For NEUS, **h** RMSE of experiment TR-RAN-TE-RAN, **i** RMSE of experiment TR-C-TE-H, **j** grid-wise absolute percentage difference between (**h**) and (**i**) with (**h**) as reference, **k** comparison between RMSE of TR-RAN-TE-RAN and TR-C-TE-H, represented as PDFs, **l** RMSE of experiment TR-nonEN-TE-EN, **m** grid-wise absolute percentage difference between (**h**) and (**l**) with (**h**) as reference, **n** comparison between RMSE of TR-RAN-TE-RAN and TR-nonEN-TE-EN, represented as PDFs

4.1 Experiments TR-RAN-TE-RAN and TR-CH-TE-CH

Figure 7 shows the comparison of two experiments TR-RAN-TE-RAN and TR-CH-TE-CH in terms of root mean square error (RMSE) over India and NEUS. The RMSE for the experiment TR-RAN-TE-RAN for India (Fig. 7a) shows higher magnitude over the windward side of the Western Ghats region, Central Northeast India (CNI), north-east Indian region and the Himalayan foothills. It should be noted that, these are the regions receiving higher rainfall as compared to other parts of India. The RMSE for the experiment TR-CH-TE-CH (Fig. 7b) shows similar spatial pattern as that of TR-RAN-TE-RAN with slightly elevated magnitude. Absolute percent difference in RMSE (Fig. 7c) points out the locations such as coastal regions on West, East, and South East side, northern India and JAK region, where the RMSE for the experiments TR-RAN-TE-RAN and TR-CH-TE-CH differ with comparatively larger magnitude. These regions indicate lack of stationarity for the given set of experiments over India. The differences in RMSE magnitudes (Fig. 7d) are also presented in terms of

their PDFs and the PDF of TR-CH-TE-CH shows a shift with respect to that of TR-RAN-TE-RAN. For NEUS, the RMSE for experiment TR-RAN-TE-RAN (Fig. 7e) and experiment TR-CH-TE-CH (Fig. 7f) do not show similarity in magnitude and spatial pattern, except over the southern coastal part. This is evident from the absolute percentage differences in RMSE plot (Fig. 7g, h), showing comparison of PDFs. Figure 7g indicates that the absolute percentage difference is <10 % for almost entire study region, except for some of the locations in western and southern side. As compared to India, stationarity is observed to prevail in NEUS over most of the regions for the set of experiments, executed here.

4.2 Experiments TR-RAN-TE-RAN, TR-C-TE-H, and TR-nonEN-TE-EN

(1) Training with cold years and testing for hot years (TR-C-TE-H) and (2) training with non-El-Niño years and testing for El-Niño years (TR-nonEN-TE-EN) are the experiments, where diametrically opposite climatic conditions are considered for testing the validity of assumption of stationarity. The RMSE of the experiments TR-C-TE-H and TR-nonEN-TE-EN are compared with RMSE of TR-RAN-TE-RAN for India and NEUS and illustrated in Fig. 8. The RMSE for TR-RAN-TE-RAN (Fig. 8a) is observed to be less in magnitude, as compared to that of TR-C-TE-H (Fig. 8b) and TR-nonEN-TE-EN (Fig. 8e), mainly in the Western Ghats and Central India; however for both the cases, the spatial pattern remains similar. Absolute percentage difference in RMSE for TR-C-TE-H (Fig. 8c) shows certain regions of higher magnitudes, mainly in the northern part of India, central and the west coast of India, JAK, and the southeast coast. This indicates that the assumption of stationarity in predictor-predictand relationship is getting violated, in majority of the regions, in warming environment. Absolute percentage difference in RMSE between TR-nonEN-TE-EN and TR-RAN-TE-RAN (Fig. 8f) shows high magnitude locations in the northern part, northeast region, central India, and mainly in the west coast. These findings are also observed in the form of shifts, when PDFs for TR-C-TE-H (Fig. 8d) and for TR-nonEN-TE-EN (Fig. 8g) are compared with TR-RAN-TE-RAN. The outcomes of the same experiments, executed over NEUS, are displayed in Fig. 8h–n. The southern and the western parts of NEUS show violation of the assumption of stationarity for these experiments. The RMSE for TR-RAN-TE-RAN (Fig. 8h) and TR-C-TE-h (Fig. 8i) show higher differences in the southern part. The same finding is revealed in absolute percent difference plot (Fig. 8j), which shows higher magnitude of differences and Fig. 8k, showing shifts in PDFs. On the other hand, the RMSE for TR-RAN-TE-RAN (Fig. 8h) and TR-nonEN-TE-EN (Fig. 8l) show similarities in spatial pattern and magnitude. The findings are consistent with those, shown by absolute percent difference plot (Fig. 8m, n), showing comparison of PDFs. Relative comparison of results over two study regions shows that the assumption of stationarity seems to hold good in NEUS as compared to India for the experiments TR-C-TE-H and TR-nonEN-TE-EN, that are derived based on hypothetical expected climate change. For India, absolute percentage differences are in higher magnitudes at different locations, spreading across coastal regions, central India, and northeast region. For NEUS, the southern part seems to be affected with the problem of non-stationarity for experiment TR-C-TE-H.

4.3 Experiments TR-RAN-TE-RAN, TR-H-TE-C, and TR-EN-TE-nonEN

The experiments (1) Training with hot years and testing for cold years (TR-H-TE-C) and (2) training with El-Niño years and testing for nonEl-Niño years (TR-EN-TE-nonEN) are hypothetical scenarios posing reverse climate change. The RMSE of the experiments TR-H-TE-C and TR-EN-TE-nonEN are compared with RMSE of TR-RAN-TE-RAN for India and NEUS (refer Online Resource 1, Supplementary Figure 5). Comparison of RMSE over India revealed that high differences exist in absolute percentage RMSE over the northern India, the central India, South east coast, upper part of the Western Ghats, Jammu and Kashmir region, and northeast region indicating lack of stationarity. For NEUS, regions 1 and 4 show violation in the assumption of stationarity.

4.4 Possible reasons behind non-stationarity

Here, we attempt to identify the possible reasons behind non-stationarity as revealed from series 1 experiments. Figure 9 brings out the results of pilot analysis, providing a hint towards one of the "possible" drivers behind violation of assumption of stationarity. The regions in India and NEUS, showing lack of stationarity for experiment TR-C-TE-H (Fig. 9(a) and (d)) and TR-nonEN-TE-EN (Fig. 9(b) and (e)), can be identified as the locations with high magnitude of absolute percentage different in RMSE (darker shades of gray). Preliminary analysis, based on visual inspection revealed that these (darker shades of gray) locations are high population dwelling areas in both the study regions. Considering reasonable chances in favor of the proposition that high population locations are likely to be urbanized, an analysis is carried out by identifying such high population locations and urbanized areas in India and NEUS. Figures (9c) and (9f) illustrate the high population locations (as black circles) and the positions of urbanized regions (as shaded regions), superimposed over each other.

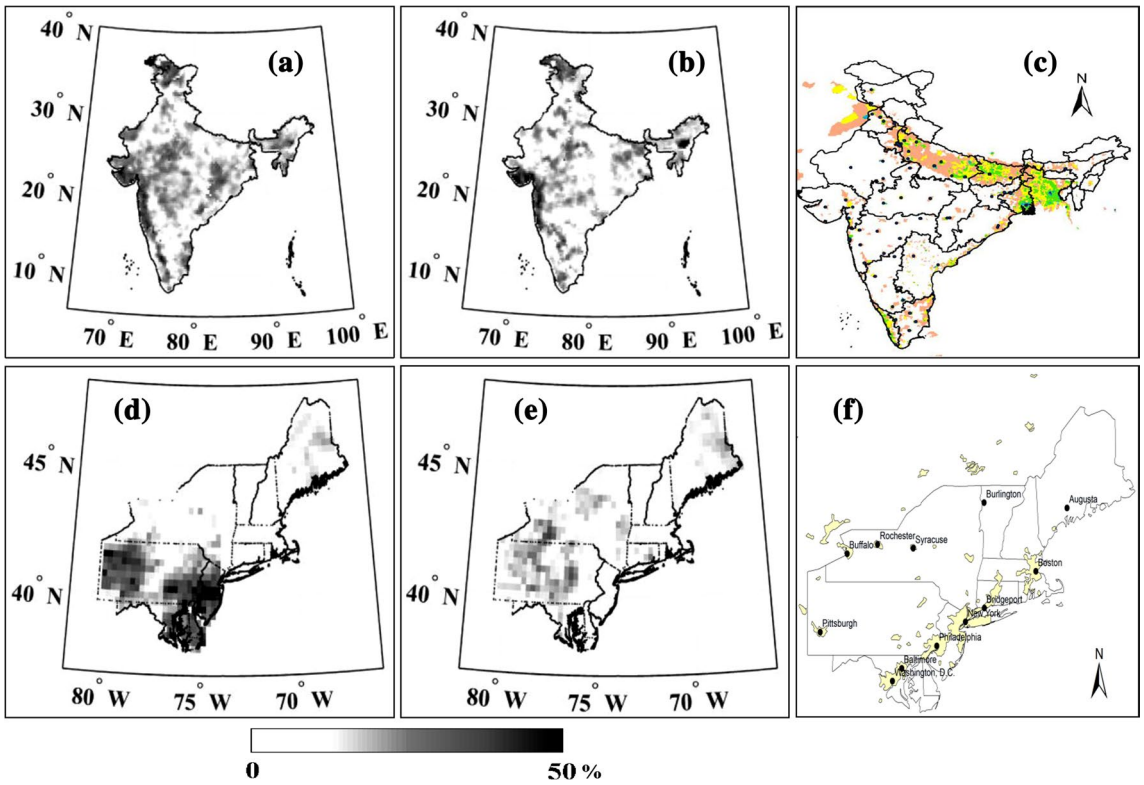
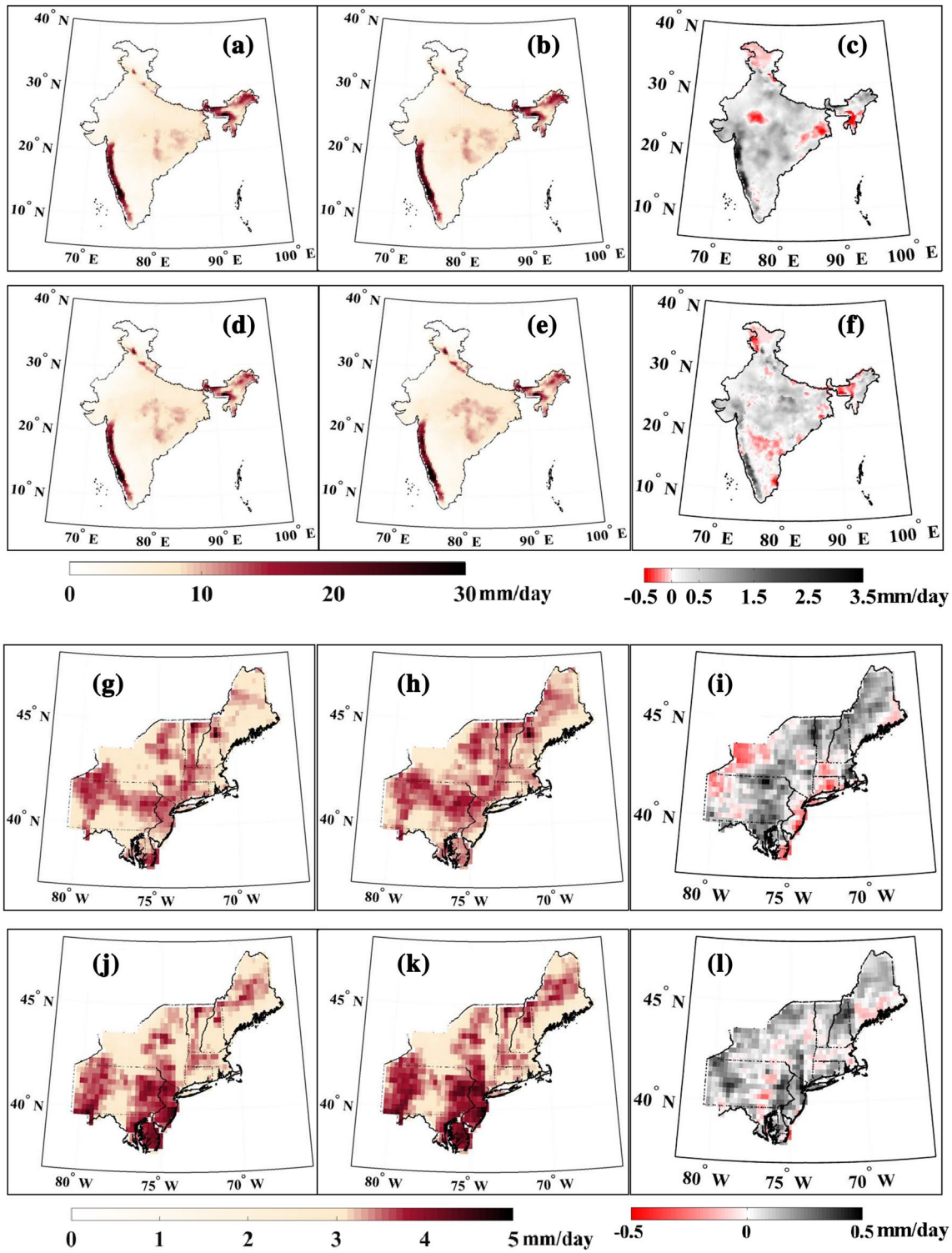


Fig. 9 Possible influence of urbanization over violation in assumption of stationarity. For India, **a** grid-wise absolute percentage difference between RMSE of experiments TR-RAN-TE-RAN and TR-C-TE-H, **b** grid-wise absolute percentage difference between RMSE of experiments TR-RAN-TE-RAN and TR-nonEN-TE-EN (both showing *shaded regions* where assumption of stationarity is violated), **c** location of high population cities (*black dots*) and urbanized areas (*shades regions*). For NEUS, **d** grid-wise absolute percentage difference between RMSE of experiments TR-RAN-TE-RAN and TR-C-TE-H, **e** grid-wise absolute percentage difference between RMSE of experiments TR-RAN-TE-RAN and TR-nonEN-TE-EN (both showing *shaded regions* where assumption of stationarity is violated) **f** location of high population cities (*red dots*) and urbanized areas (*shades regions*)

The population data and the data for urbanized regions are obtained from http://www.naturalearthdata.com/downloads/50m-cultural-vectors/ (Junk et al. 2014; Lin et al. 2014, Jenkinsa et al. 2015) and http://www.diva-gis.org/gdata (Wisz et al. 2008; Jimenez-Valverde et al. 2008; Masello et al. 2015) collectively. It is clearly visible from the Fig. 9c, f that most of the high population locations are urbanized. Also, it is clear from Fig. 9 that the locations, where assumption of stationarity fails to hold good (gray shades) and locations of high population urbanized areas show fair match. Possible reasons behind this match can be explained as follows. Urban areas are observed to have different climatology (Kishtawal et al. 2010; Mishra and Lettenmaier 2011; Shastri et al. 2015). The downscaling model, which is deployed in the present study, does not incorporate the effect of urbanization and hence it is not able to capture the changes in the rainfall patterns that are because of local level modifications. This may be one of reasons behind the violation of stationarity assumption. However, this is a possible hypothesis, which may be tested with follow on model based research activities.

4.5 Experiments SB-AP-PI and SB-AP-RCP85

These are the second series of experiments, where the existence of stationarity in the system is checked based on the capability of the model to simulate changes in mean rainfall under the dominance of GHG emissions. It is important to note that the two set of years, which are close to preindustrial scenario and to RCP8.5, show some overlap, with 7 years for India and 9 years for NEUS. This is one of the shortcomings of this approach, which can be improved with longer validation period and improved analogue identification technique. Figure 10 shows the results for these experiments. Figure 10a is the mean of observed rainfall for ACC-PI signatures years and Fig. 10b is the mean of observed rainfall for ACC-RCP85 signatures years. Figure 10c is the difference between two central tendencies, which illustrates changes in mean observed rainfall, expected due to GHG emissions. Figure 10d–f are exact counterparts of Fig. 10a–c, obtained for simulated rainfall. Again, similarities in spatial pattern and magnitude are visible for Fig. 10d, e. Figure 10f shows different spatial patterns of changes in mean



as compared to Fig. 10c over majority of locations in the central India, the northeast India, and in the Western Ghats; however, over some of the locations in north, western, and Peninsular India, the spatial patterns show a good match. These are the only regions where model is able to capture the changes in mean rainfall due to strongest radiative forcing. Similar plots for NEUS (Fig. 10g–l) show that, SD model fails to simulate the expected changes in majority of the areas except northern part of NEUS (region 4).

In order to have more confidence in analysis, experiments SB-AP-PI and SB-AP-RCP85 are carried out using another GCM MIROC-ESM (resolution ~2.8°). This

◀**Fig. 10** Results of experiment series 2: Signature based approach SB-AP-PI and SB-AP-RCP85. For India, **a** mean of observed rainfall, obtained for the years from recent past (1981–2005) that showed signatures of Pre-industrial (PI) climatic conditions, **b** mean of observed rainfall, obtained for the years from recent past (1981–2005) that showed signatures of RCP8.5 climatic conditions, **c** changes in mean of observed rainfall because of GHG emissions [difference between (**b**) and (**a**)], **d** mean of projected rainfall, obtained for the years from recent past (1981–2005) that showed signatures of PI climatic conditions, **e** mean of projected rainfall, obtained for the years from recent past (1981–2005) that showed signatures of RCP8.5 climatic conditions, **f** changes in mean of projected rainfall because of GHG emissions [difference between (**e**) and (**d**)]. For NEUS, **g** mean of observed rainfall, obtained for the years from recent past (1981–2005) that showed signatures of PI climatic conditions, **h** mean of observed rainfall, obtained for the years from recent past (1981–2005) that showed signatures of RCP8.5 climatic conditions, **i** changes in mean of observed rainfall because of GHG emissions [difference between (**h**) and (**g**)], **j** mean of projected rainfall, obtained for the years from recent past (1981–2005) that showed signatures of PI climatic conditions, **k** mean of projected rainfall, obtained for the years from recent past (1981–2005) that showed signatures of RCP8.5 climatic conditions, **l** changes in mean of projected rainfall because of GHG emissions [difference between (**k**) and (**j**)]

analysis is carried out only for India. The results, obtained using this GCM are illustrated in Supplementary Figure 6 (refer Online Resource 1, SF6). Comparison between PDFs of climate variables simulated by MIROC-ESM for PI and RCP8.5 shows statistically significant changes. SD model is not able to capture changes in mean rainfall over the Western Ghats, the central India, and the northeast India.

To understand the reasons behind such failure in the regions of India and NEUS, we plot the partial correlation between predictor and predictand for both data subsets ACC-PI and ACC-RCP85. Partial correlation involves calculating the correlation between two variables, holding constant the external influences of third. In a regression model, the regression coefficients are obtained using the partial correlation values between predictor and predictand. Transfer function based SD model, being a regression model, relies on partial correlation. If the partial correlation coefficients between a specific predictor and predictand are different for two time periods with different climatic condition, the predictor-predictand relationship will also be different. In such cases the regression equation developed for the first period will not be valid for the second, which is the reason behind the violation of assumption of stationarity. Hence, partial correlation coefficient plots are obtained between observed rainfall and the set of predictors for years showing signatures of (1) PI scenario and (2) RCP8.5 scenario. For India, this analysis is carried out for two zones viz. WCI and CNI, whereas for NEUS, this analysis is carried out for region 1, and region 4. The partial correlation coefficients are obtained for all ten predictors (listed in Table 1). Figure 11 illustrates the results of this analysis carried over WCI and CNI zone of India. Partial correlation for each predictor is illustrated in a pair of plots e.g. Figure 11a1, a2 correspond to the partial correlation between principal component of specific humidity at 500 hPa with observed rainfall (for WCI region) for years showing signatures of PI scenario (Fig. 11a1) and that of RCP8.5 scenario (Fig. 11a2). Out of the ten pairs of plots obtained for WCI zone, the difference in spatial pattern of the two corresponding plots are prominent for specific humidity [surface (Fig. 11b1, b2) and pressure level (Fig. 11a1, a2)], mean sea level pressure (Fig. 11c1, c2) and subtle for temperature [surface (Fig. 11e1, e2) and pressure level (Fig. 11d1, d2)] and geopotential height (Fig. 11j1, j2). For wind variables (surface and pressure levels, Fig. 11f1–i1, f2–i2) the patterns match well. Exactly similar behavior is visible for partial correlation plots, obtained for CNI zone. The difference in spatial pattern of the two corresponding plots are prominent for specific humidity [surface (Fig. 11l1, l2) and pressure level (Fig. 11k1, k2)], mean sea level pressure (Fig. 11m1, m2) and subtle for temperature [surface (Fig. 11o1, o2) and pressure level (Fig. 11n1, n2)] and geopotential height (Fig. 11t1, t2). For wind variables (surface and pressure levels, Fig. 11p1–s1, p2–s2) the patterns match well.

Similar analysis is carried out for two regions in NEUS that is illustrated in Fig. 12. For region 1, mean sea level pressure (Fig. 12c1, c2), air temperature at 500 hPa (Fig. 12d1, d2), and geopotential height (Fig. 12j1, j2) show dissimilarities. However, for region 4, the same predictors are accompanied by surface level specific humidity (Fig. 12l1, l2) and surface air temperature (Fig. 12o1, o2). In both the figures (Fig. 12l1, l2), the predictors, showing difference in spatial pattern of partial correlation plots are responsible for bringing down performance of SD model in capturing changes in mean rainfall.

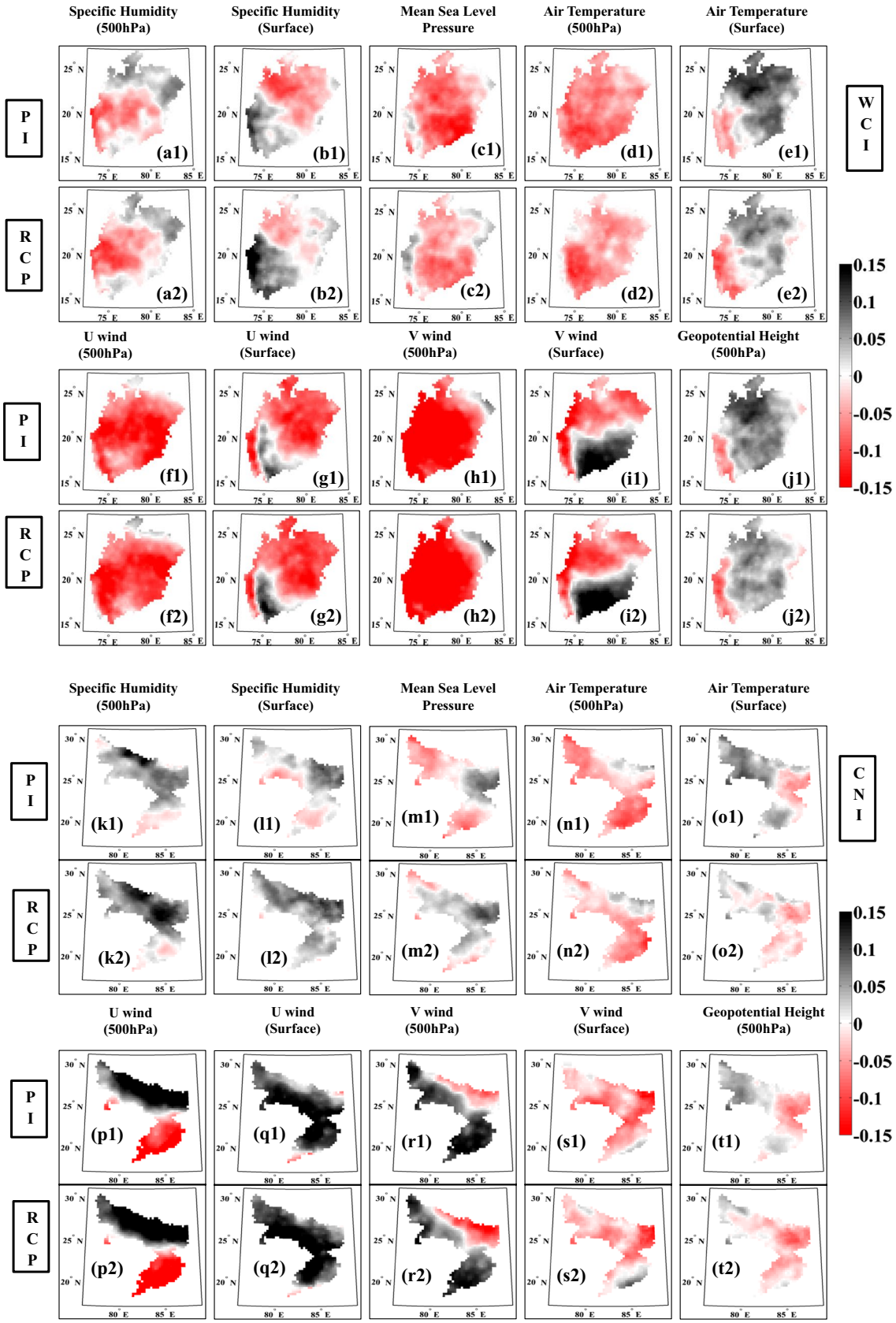
4.6 Improvement in SD model under nonstationary conditions

Both set of experiments viz. series 1 and series 2, which are executed over two study regions India and NEUS revealed (1) different locations in the study regions where assumption of stationarity is getting violated (outcome of series 1 experiments) and (2) inability of downscaling model to capture the changes in mean rainfall on account of GHG emissions (outcome of series 2 experiment). As discussed before, violation of assumption of stationarity may be because of the changes, occurring at a large scale or possible interventions by some local factors. Identifying the exact root cause of such non-stationarity and incorporating its solution in SD model might help to overcome the limitation of data driven models. However, it is a nontrivial task and through analysis would be necessary to pinpoint the exact reason behind non-stationarity, which may vary

from region to region. Partial correlation analysis (Fig. 11) brings out an interesting fact that the ability of SD model to capture changes in mean rainfall gets sabotaged because of inclusion of certain predictors. Hence, there is a possibility that SD model might performs better under nonstationary conditions by inclusion of relevant predictors. This finding is consistent with the proposition by Charles et al. (1999) and Wilby et al. (2004), which say that the assumption of stationarity may be robust if the choice of the predictor is judicious. Different predictor selection criteria such as partial correlation analysis, step-wise regression etc. already advocated in literature (Charles et al. 1999; Wilby et al. 2003). For the present study, we refer to different indices that have been defined to measure and to predict the yearly variations and future developments of the monsoon strength. Different indices, such as precipitation based indices (Parthasarathy et al. 1992; Goswami et al. 1999), vertical wind shear based indices (Webster and Yang 1992; Chen et al. 2007; Goswami et al. 1999; Dobler and Ahrens 2011) and longwave radiation based indices (as a measure of convection) (Wang and Fan 1999) are discussed in literature. Even though, there is no single best index in estimating ISMR strength (Wang and Fan 1999; Goswami 2000; Wang 2000), here, we consider meridional and zonal wind shear indices (Dobler and Ahrens 2011) and include meridional and zonal winds at 850 and 250 hPa as additional predictors. Along with these wind variables, we use other climate variables such as air temperature, geopotential height, and specific humidity at the mentioned pressure levels (As specific humidity at 250 hPa is not available with NCEP/NCAR data, we used specific humidity at 300 hPa). In reality, the atmospheric levels 850 hPa (lower) and 200 hPa (upper) are standard levels for analyzing the strength and direction of the large-scale circulation important for the Indian summer monsoon. However, instead of 200 hPa, we use data at 250 hPa because GCM simulations are commonly available at this pressure level. The list of the predictors that are used for the analysis is mentioned in Table 1 (Sect. 3). Using these predictors, the designed experiments are executed for India and Fig. 13 shows consolidated results of series 1 and series 2 experiments.

For series 1 experiments, spatial distribution of RMSE for TR-RAN-TE-RAN (Fig. 13X) is compared with the RSME for TR-C-TE-H (Fig. 13a), TR-nonEN-TE-EN (Fig. 13c), TR-CH-TE-CH (Fig. 13e), TR-H-TE-C (Fig. 13g), and TR-EN-TE-nonEN (Fig. 13i). The corresponding absolute percentage differences in RMSE are illustrated in Fig. 13b, d, f, h, j. Comparing all the plots, showing absolute percentage differences together, with the previously obtained plots brings out two important findings. First, there is a reduction in spatial extent where absolute percentage difference shows high magnitude, indicating assumption of stationarity is valid for a larger extent in study region. Secondly, we find that all the experiments show consistent locations, where the assumption of stationarity does not hold good. All the plots show violation in the assumption of stationarity near the southern west coast, a small region near the east coast, and in the central India, unlike the previous version of results (obtained without including predictors at 850 and 250 hPa), which showed different regions, where violation of assumption of stationarity is encountered. The results, obtained for experiment series 2 (Fig. 13k–p) show the capability of SD model to simulate changes in mean rainfall because of GHG emissions. Figure 13m shows the changes in mean rainfall for observed data and Fig. 13p shows changes in mean rainfall for downscaled data. These demonstrate a good match in spatial pattern as compared to their counterparts in previous run (Fig. 10c, f). The performance of SD model is improved by incorporating relevant climate predictors. We further try to analyze these improvements with the help of partial correlation plots. These plots are obtained for two zone viz. WCI and CNI zones in exactly the same manner as discussed before. Supplementary Figures 8 and 9 (refer Online Resource 1, SF8 and SF9) show the partial correlation for WCI and CNI zones respectively. For WCI, geopotential height (at all pressure levels) and air temperature at 500 hPa show difference in spatial pattern of partial correlation coefficients, whereas for CNI, surface level specific humidity shows difference in spatial pattern of partial correlation coefficients. The partial correlation plots for remaining predictors (other than those mentioned above) show similar magnitudes and spatial patterns. In comparison with the partial correlation plots as shown in Fig. 11, we encounter better matching of spatial pattern of partial correlation plots for different predictors, resulting in improved results.

Fig. 11 Partial correlation analysis over WCI and CNI regions in India. Comparison between (1) partial correlation, obtained with observed rainfall and different predictors for years (from recent past, 1981–2005) showing signatures of PI run [specific humidity at 500 hPa (**a1**), surface level specific humidity (**b1**), mean sea level pressure (**c1**), air temperature at 500 hPa (**d1**), surface level air temperature (**e1**), zonal 'U' wind at 500 hPa (**f1**), surface level zonal wind (**g1**), meridional 'V' wind at 500 hPa (**h1**), surface level meridional wind (**i1**), geopotential height at 500 hPa (**j1**)] and (2) partial correlation, obtained with observed rainfall and different predictors for years showing signatures of RCP8.5 [specific humidity at 500 hPa (**a2**), surface level specific humidity (**b2**), mean sea level pressure (**c2**), air temperature at 500 hPa (**d2**), surface level air temperature (**e2**), zonal 'U' wind at 500 hPa (**f2**), surface level zonal wind (**g2**), meridional 'V' wind at 500 hPa (**h2**), surface level meridional wind (**i2**), geopotential height at 500 hPa (**j2**)] over WCI is illustrated. Similar analysis carried out over CNI (**k1–t1** and **k2–t2** are exact counterparts of **a1–j1** and **a2–j2** respectively, obtained for CNI)


Specific Humidity (500hPa)
Specific Humidity (Surface)
Mean Sea Level Pressure
Air Temperature (500hPa)
Air Temperature (Surface)
PI
RCP
WCI
(a1)
(b1)
(c1)
(d1)
(e1)
(a2)
(b2)
(c2)
(d2)
(e2)
U wind (500hPa)
U wind (Surface)
V wind (500hPa)
V wind (Surface)
Geopotential Height (500hPa)
(f1)
(g1)
(h1)
(i1)
(j1)
(f2)
(g2)
(h2)
(i2)
(j2)
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
Specific Humidity (500hPa)
Specific Humidity (Surface)
Mean Sea Level Pressure
Air Temperature (500hPa)
Air Temperature (Surface)
CNI
(k1)
(l1)
(m1)
(n1)
(o1)
(k2)
(l2)
(m2)
(n2)
(o2)
U wind (500hPa)
U wind (Surface)
V wind (500hPa)
V wind (Surface)
Geopotential Height (500hPa)
(p1)
(q1)
(r1)
(s1)
(t1)
(p2)
(q2)
(r2)
(s2)
(t2)

5 Summary and discussion

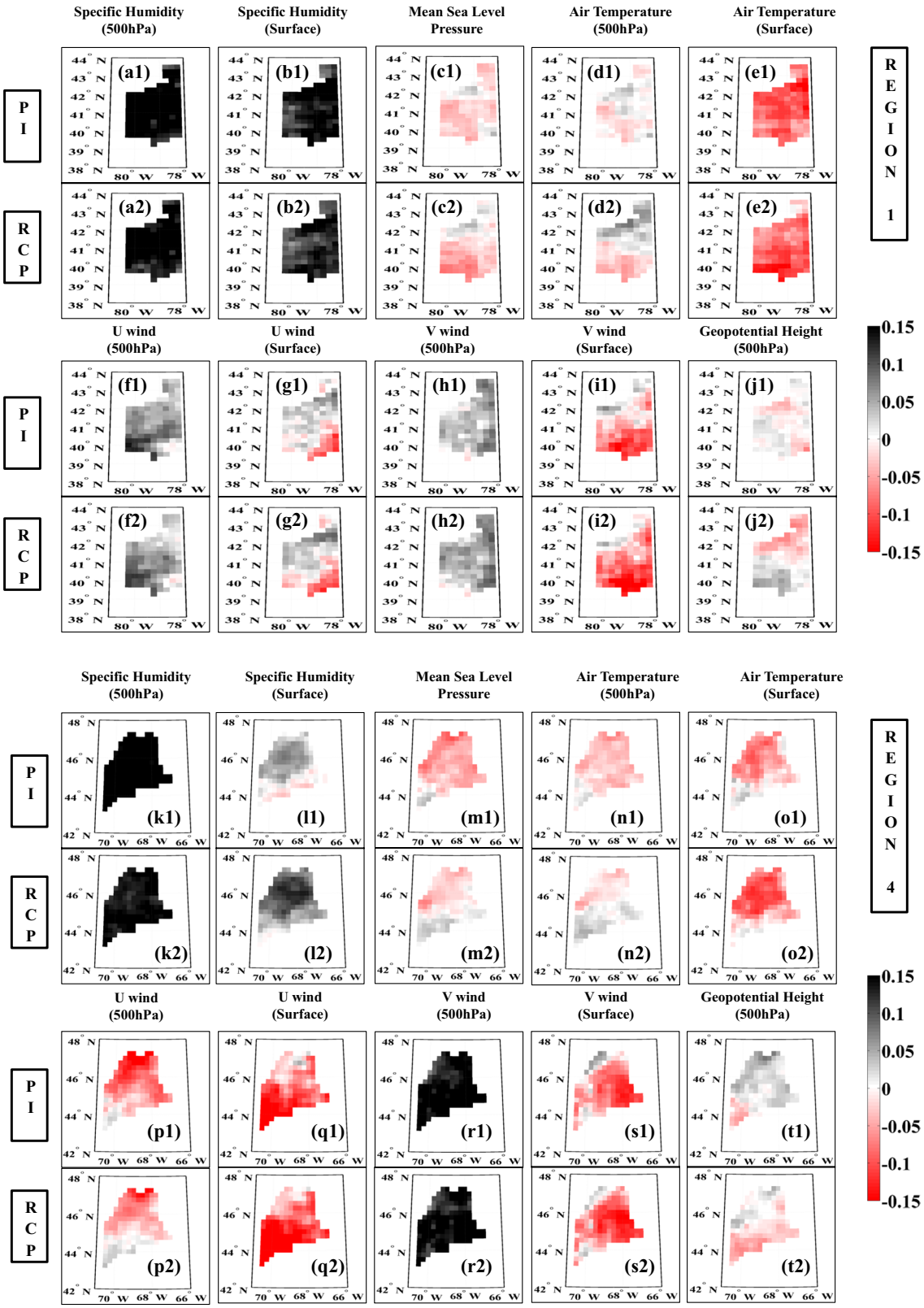
SD methodologies show good skills in capturing different statistical properties of evidentiary target data over historic period simulations. However, the violation of assumption of stationarity under changing climatic conditions may reduce the credibility of the model to produce realistic future projections. The present study is undertaken with the rationale of developing statistical experiments to (1) identify the regions where the assumption of stationarity gets violated and (2) to judge the ability of SD model to perform under changing climatic conditions. First series of experiments is based on different criteria of selecting training period such as (1) statistical based (random, chronological), (2) hypothetical climate change based (cold years as training, nonEl-Niño years as training), and (3) reverse climate change (hot year as training, El-Niño years as training). The second series of experiment involves, validating the credibility of the downscaling model to perform under changing climatic conditions, by comparing the effect of GHG emissions on observed and simulated mean rainfall. This is realized by identifying two sets of years (from recent past), which show very close resemblance to the average climatic conditions of preindustrial run and extreme GHG emission run (RCP8.5). The difference in the mean rainfall between the two sets will show the effects of GHG emission on mean rainfall.

Experiments are performed over two study regions with completely different climatic conditions, specifically, India, and North Eastern United States, with precipitation as the variable of interest and reanalysis data as the predictors. Over India, challenges in SD owing a possible lack of stationarity, is primarily observed over the north, in the Western Ghats, and long the southeast and east coasts. Over Northeast US, the corresponding regions are the southern and western parts. The regions, where non-stationarity results in degradation of performance are found to be high population locations. However, this observation is based on the visual inspection and such hypothesis must be validated with follow on research activities including model runs with urban canopy coupling. The second series of experiments shows that the model fails to capture the possible changes in mean rainfall, due to GHG emissions, mainly in WCI, the Western Ghats, some part of PI, Gujarat indicating lack of stationarity over these regions. Some of the regions such as Rajasthan, CNI show that the model is able to capture the changes in mean rainfall. The same experiment, carried out for NEUS, shows lack of stationarity in all regions except region 1. Partial correlation analysis is carried out over two regions (WCI and CNI for India and regions 1 and 4 for NEUS) in each study area. It is observed that, for India, all the predictors (except wind variables) showing different pattern of partial correlation for the time periods having synoptic circulation patterns similar to PI and RCP 8.5. For NEUS, temperature at 500 hPa, surface level specific humidity, mean sea level pressure, and geopotential height at 500 hPa are the only predictors showing different spatial pattern of partial correlation. Selection of appropriate predictors might help in reducing the nonstationarity in relationship, however, the selected predictors for SD, in that case should carry the climate change signal. Taking this into considerations, all the experiments (series 1 and series 2) are executed over India including pressure level predictors at 850 and 250 hPa. The results showed improvements in terms of reduction in spatial extent of regions with nonstationarity and elevated capability of SD model to capture changes in mean rainfall because of GHG emissions.

Fig. 12 Partial correlation analysis over region1 and region4 in NEUS. Comparison between (1) partial correlation, obtained with observed rainfall and different predictors for years (from recent past, 1981–2005) showing signatures of PI run [specific humidity at 500 hPa (**a1**), surface level specific humidity (**b1**), mean sea level pressure (**c1**), air temperature at 500 hPa (**d1**), surface level air temperature (**e1**), zonal 'U' wind at 500 hPa (**f1**), surface level zonal wind (**g1**), meridional 'V' wind at 500 hPa (**h1**), surface level meridional wind (**i1**), geopotential height at 500 hPa (**j1**)] and (2) partial correlation, obtained with observed rainfall and different predictors for years showing signatures of RCP8.5 [specific humidity at 500 hPa (**a2**), surface level specific humidity (**b2**), mean sea level pressure (**c2**), air temperature at 500 hPa (**d2**), surface level air temperature (**e2**), zonal 'U' wind at 500 hPa (**f2**), surface level zonal wind (**g2**), meridional 'V' wind at 500 hPa (**h2**), surface level meridional wind (**i2**), geopotential height at 500 hPa (**j2**)] over region 1 are illustrated. Similar analysis carried out over region 4 (**k1–t1** and **k2–t2** are exact counterparts of **a1–j1** and **a2–j2** respectively, obtained for region 4

The results, obtained with methodologies that are proposed in the present manuscript can be improved by taking care of some of the limitations.

1. The experiments, discussed in the manuscript are executed using only one downscaling methodology (developed by Kannan and Ghosh (2013)). It is necessary to check how other downscaling methodologies perform in this context. As the basis behind every data driven downscaling technique is the same (model is calibrated over past data), it is unlikely that other methods such as weather generators or weather typing would show exceptional skills as compared to the methodology used in the present manuscript. However, it will be interesting to see if other methods show capability to generalize under changing climatic conditions.
2. Selection of predictors is a very important step for the reliable performance of downscaling model. In the present study, even though, the predictors are selected based on the criteria mentioned in the literature (Wilby et al. 2004; Kannan and Ghosh 2013), some extra



analysis may be performed for the same. The analysis based on partial correlation, presented in the current manuscript maybe useful, in this regard. Both the partial correlation plots obtained for India and NEUS, show that the temperature at 500 hPa is one of the variable showing difference in the spatial pattern. It

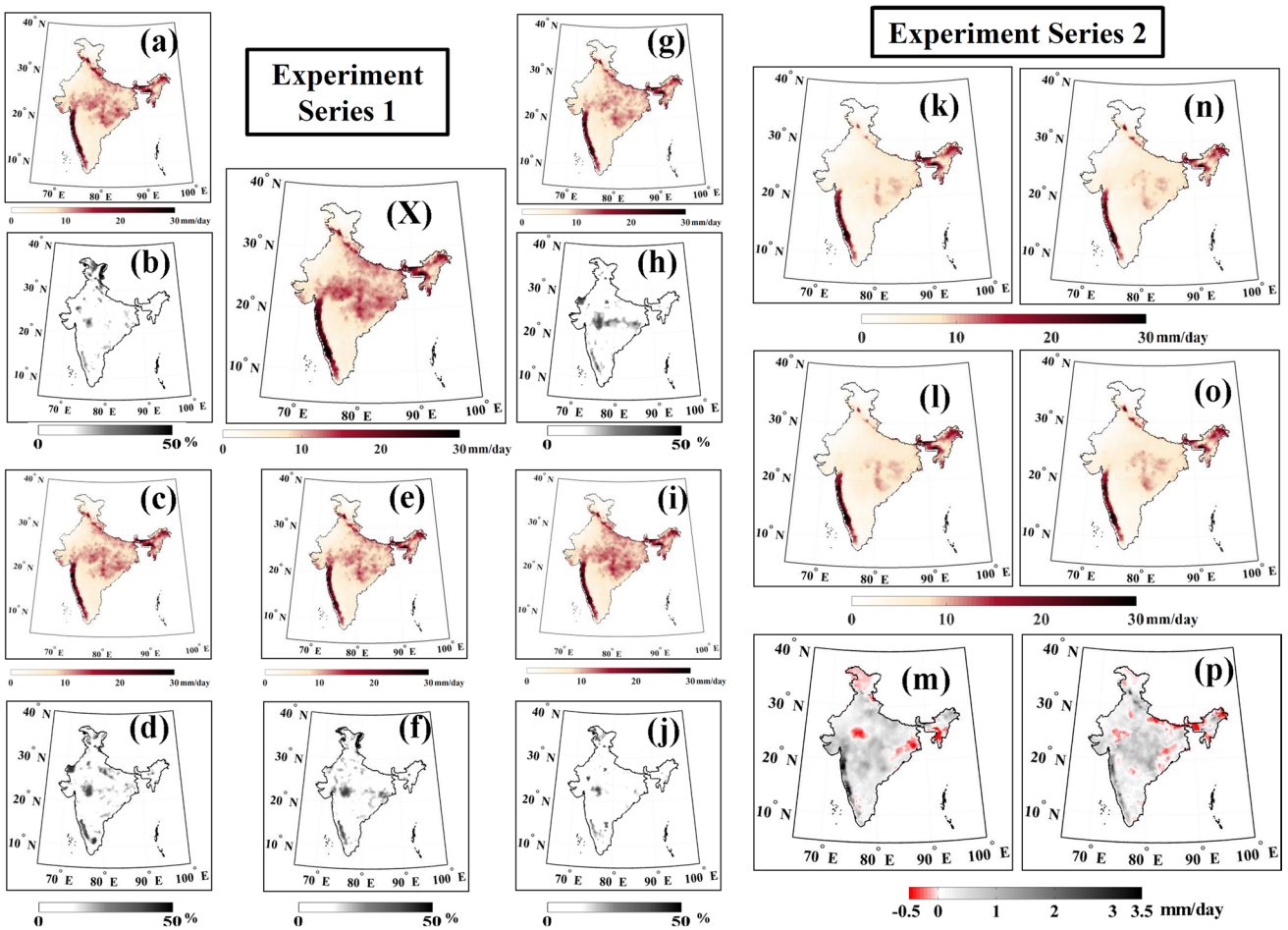


Fig. 13 Improvement in competence of SD model under nonstationary conditions on account of inclusion of pressure level predictors. *Experiment series 1* panel shows comparison between RMSE for Base experiment (TR-RAN-TE-RAN) (**X**) with RMSE of other experiments (TR-C-TE-H, TR-nonEN-TE-EN, TR-CH-TE-CH, TR-H-TE-C, TR-EN-TE-nonEN) on the basis of (1) spatial distribution, illustrated in (**a**, **c**, **e**, **g**, **i**) and (2) difference in magnitude (represented as absolute percentage difference), illustrated in (**b**, **d**, **f**, **h**, **j**) respectively. *Experiment series 2* panel shows results of signature based approach. **k** mean of observed rainfall, obtained for the years from recent past (1981–2005) that showed signatures of Pre-industrial (PI) climatic conditions, **l** mean of observed rainfall, obtained for the years from recent past (1981–2005) that showed signatures of RCP8.5 climatic conditions, **m** changes in mean of observed rainfall because of GHG emissions [difference between (**l**) and (**k**)], **n** mean of projected rainfall, obtained for the years from recent past (1981–2005) that showed signatures of PI climatic conditions, **o** mean of projected rainfall, obtained for the years from recent past (1981–2005) that showed signatures of RCP8.5 climatic conditions, **p** changes in mean of projected rainfall because of GHG emissions [difference between (**o**) and (**n**)]

is quite possible that inclusion of this predictor might bring down the performance of downscaling methodology under the influence of GHG emissions.

3. Present study is carried out with NCEP/NCAR reanalysis data as the source of gridded climate variables. However, there are other reanalysis products such as the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data (ERA40) (Uppala et al. 2005), NCEP–DOE Atmospheric Model Intercomparison Project (AMIP-II) reanalysis (Kanamitsu et al. 2002), and Japanese 25-year ReAnalysis (JRA-25) (Onogi et al. 2007) from the Japan Meteorological Agency etc. which are at finer spatial resolution as compared to NCEP/NCAR data. Application of different reanalysis products (which are at different spatial resolution) are likely to produce different results, leading to uncertainty (Kannan et al. 2014). At the same time, limitations of NCEP/NCAR reanalysis data in terms of, sensitivity to the shortage of observations in high southern latitudes, presence of artificial trends in the mean sea level pressure fields near Antarctica, not being able to adequately represent the atmospheric state for certain regions and time, limited skills in representation of hydrologic cycle etc. should be taken into consideration before its usage.

4. In order to obtain signatures of average climatic conditions for scenarios PI and RCP8.5 in the recent past years (1981–2005), Euclidean distance approach is used. Deployment of better approach may help to produce more realistic results.

6 Conclusion

This is the first study to the best of our knowledge which attempts to systematically evaluate a SD approach under plausible nonstationarities expected due to climate change. The DOE in the present study is guided by two important principles. First principle is that the ability of SD approaches to generalize under climate nonstationarity depends on the extent to which the extracted relationships are physically meaningful, or at least, consistent with both current and future expected climatology. Thus, a test of generalization in climate, and especially in the context of downscaling, needs to include but go beyond the standard metrics (e.g. Akaike or Bayesian information criteria) or optimization approaches (e.g. ridge regression, LASSO or the Elastic Net) which at some level attempt to balance error or skill metrics with complexity of the statistical model. The second principle is that even under global warming scenarios, historical signatures can be found for climatologies expected in the future at a given location or region. While a general assumption is that the signatures may need to be discovered in both space and time, the present study makes a further assumption that these can be found just in time. While the latter assumption may be relaxed in subsequent studies, the physical implications of these assumptions is just that at any given region or location, natural variability in the observed past may have led to situations that were similar to what may be expected in the future. Given that regional or local climate changes significantly and can never really be considered stationary, this assumption may not be too strong. From a statistical perspective, the approaches are conceptually similar to nearest neighbor approaches, where the neighborhoods in this case are defined based on the similarity of the historically signatures (climatologies) to expected future conditions. Anticipatory assessment of future performance, when the radiative forcing differs from current condition, is based on extraction of historical signatures. Based on the premise laid out here, a partial falsification of the SD-based projections may be possible, and the situations where SD approaches do not work well can be examined in depth. The present study attempt to accomplish this, and based on two distinct case studies in the NEUS and India with precipitation as the variable of interest, different possible reasons for violations of assumption of stationarity emerge out. Outcomes of statistical experiments, belonging to 'series 1' category (designed to identify regions where stationarity fails) hint towards a possible association between regions, where assumption of stationarity is violated and locations of dense populations. Hypothetically, it can be argued that the local changes such as land use land cover change (barren land to urban) are possible outcomes of increase in population density to cater their basic needs. In case of established predictor-predictand relationship, if predictand (rainfall) is highly sensitive to such local changes as compared to predictor, rapid changes in local drivers will influence predictand and the old established relationship might fail to capture these changes. However, such arguments might prove to be a mirage unless supported by thorough hypothetical testing in details. As far as present study is concerned, we limit our scope to put forth the hypothesis that possible implications of high population density have some role to play in disturbing stationary relationship. Secondly, the outcome of 'series 2' experiments (signature based approach) shows that the capability of SD model to capture changes in mean rainfall under high radiative forcing scenarios can be improved by inclusion of correct predictors. This is evident from the partial correlation plots pointing out to different combinations of predictors such as specific humidity, temperature, mean sea level pressure, geopotential height for different regions. Brining out a scientific explanation behind why these predictors bring down performance of SD is a nontrivial task. This may lead to a potent research problem worth undertaking, however, out of scope for present study. Future studies may need to examine the sensitivities when historical signatures are considered in both space and time.

Any scientific endeavor or assertion needs to be ultimately subjected to falsifiability, but this is especially hard in climate change where the community is interested in long lead time projections of a system that is ultimately nonlinear and dynamical. Climate simulations, whether from GCMs or RCMs used for DD, and SD, are often evaluated by examining historical skills and past and future multimodel agreement as well as consistency with known physical processes or established data associations. A contribution of the present work is the development of a design of experiments strategy where the falsification may be accomplished for the kind of nonstationarity expected under both natural climate variability and change at local to regional scales. Follow-on studies may be able to take this further by relaxing certain implicit assumptions by considering, for example, more disparate case studies based on natural climate variability (e.g. in addition to El Niño versus non El Niño years), or more signatures by including spatial climate signatures in addition to temporal (historical) ones. The design of experiments strategy for evaluation under nonstationary climate is not restricted to SD approaches alone

and may be valid for dynamic downscaling and GCMs. Current evaluation of GCMs includes examination of historical skills, multimodel agreement, and adherence to known physical processes. An adaptation of the design of experiments proposed here may examine anticipated performance at local to regional scales under non-stationarity conditions expected owing to global climate change. In the context of downscaling, there is a performance trade-off between SD and DD, since SD relations are purely statistical and hence subject to larger biases under changed climate conditions, while DD parameterizations, despite being physically motivated, are more complex and hence subject to higher variability. In fact, the distinctions between statistical versus parameterized physics may get blurred when the parameterizations get further away from fundamental physics, and statistical methods may have advantages over parameterized physics in situations where the parameters of the latter cannot be estimated from observed data. Nevertheless, SD provides a computationally affordable means of achieving high resolution locally for several applications that would otherwise be out of reach for most research groups (Laprise 2008). Current strategy for modeling anticipated climate changes resulting from human activities consists in making long, multi-decadal simulations, with multiple realizations of a model in an ensemble mode and, when possible, with many models (Laprise 2014). Recently, the World Climate Research Program sponsored the Coordinated Regional Climate Downscaling Experiment project (CORDEX) (Jones et al. 2011), recommended RCM simulations spanning the period 1950–2100. The purpose of long and ensemble simulations is to maximize the signal-to-noise ratio, since the "noise" in individual simulations tends to cancel one another in an ensemble (Laprise 2014). Working Group I have extended coverage of future climate change compared to earlier reports by assessing near-term projections and predictability as well as long-term projections and irreversibility (Stocker et al. 2013). However, the current study focuses on evaluation of SD, since there is a perception that SD performance is somehow more sensitive to climate non-stationarity simply because they are statistical in nature. Our results appear to suggest that the ability of SD to generalize may be subject to the same constraints that GCMs or RCMs may be subject, specifically, processes at local and regional scales such as land use or urbanization changes that are less well understood in terms of climate impacts. Further studies may need to conduct similar or extended versions of, hypotheses-driven design of experiments to test the performance of not just SD, but DD and even GCMs, under nonstationary climate. One word of caution is that a fail-safe test for complete non-stationarity is by definition, impossible to design. However, this manuscript takes a step towards an evaluation strategy that considers the kind of non-stationarity expected in local and regional climatologies under global climate change scenarios.

Acknowledgments The authors thank the anonymous reviewers for valuable comments. KS and SG acknowledge funding from the Space Technology Cell of Indian Institute of Technology Bombay and Indian Space Research Organization while ARG acknowledges support from the United States National Science Foundation (NSF) Expeditions in Computing Grant Award No. 1029166 titled "Understanding climate change: a data driven approach". The collaborative activity between IIT Bombay and Northeastern University was supported through Interdisciplinary Program in Climate Studies, IIT Bombay, funded by Department of Science and Technology, Government of India. The authors thank Amey Pathak of the Indian Institute of Technology, Bombay, as well as Udit Bhatia, Thomas Vandal and Evan Kodra, all of Northeastern University, for helpful suggestions.

References

Anandhi A, Srinivas VV, Ravi S, Nanjundiah Nageshkumar D (2008) Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine. Int J Climatol 28(3):401–420

Barker HW, Cole JNS, Morcrette JJ, Pincus R, Raisanen P, Salzen KV, Vaillancourt PA (2008) The Monte Carlo independent column approximation: an assessment using several global atmospheric models. Q J R Meteorol Soc 134:1463–1478

Bernett TP, Adam JC, Lettenmaier DP (2005) Potential impacts of a warming climate on water availability in snow-dominated regions. Nature 438:303–309. doi:10.1038/nature04141

Bogardi I, Matyasovszky I, Bardossy A, Duckstein L (1993) Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns. J Geophys Res 98(D9):1653–1667

Bromwich DH, Fogt RL (2004) Strong trends in the skill of the ERA-40 and NCEP–NCAR reanalyses in the high and midlatitudes of the Southern Hemisphere, 1958–2001. J Clim 17:4603–4619. doi:10.1175/3241.1

Busuioc A, von Storch H, Schnur R (1999) Verification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. J Clim 12:258–272

Chaboureau JP, Bechtold P (2005) Statistical representation of clouds in a regional model and the impact on the diurnal cycle of convection during Tropical Convection, Cirrus and Nitrogen Oxides (TROCCINOX). J Geophys Res 110:D17103. doi:10.1029/2004JD005645

Charles SP, Bates BC, Hughes JP (1999) A spatio-temporal model for downscaling precipitation occurrence and amounts. J Geophys Res 104(D24):657–669

Chen H, Ding YH, He JH (2007) Reappraisal of Asian summer monsoon indices and the long-term variation of monsoon. Acta Meteorol Sin 21(2):168–178

Chou C, Tu JY, Yu JY (2002) Interannual Variability of the western north Pacific summer monsoon: differences between ENSO and non-ENSO years. J Clim 16:2275–2287. doi:10.1175/2761.1

Crane RG, Hewitson BC (1998) Doubled CO_2 climate change scenarios for the Susquehanna basin: precipitation. Int J Climatol 18:65–76

DelSole T, Chang P (2003) Predictable component analysis, canonical correlation analysis, and autoregressive models. J Atmos Sci

60:409–416. doi:10.1175/1520-0469(2003)060<0409:PCACCA>2.0.CO;2

Dobler A, Ahrens B (2011) Four climate change scenarios for the Indian summer monsoon by the regional climate model COSMO-CLM. J Geophys Res 116(D24):1–13. doi:10.1029/2011JD016329

Duan J, McIntyre N, Onof C (2013) Resolving non-stationarity in statistical downscaling of precipitation under climate change scenarios. In: BHS Eleventh National Symposium, Hydrology for a changing world, Dundee 2012. ISBN: 1903741181, doi:10.7558/bhs.2012.ns16

Gautam R, Hsu NC, Lau KM, Tsay SC, Kafatos M (2009) Enhanced pre-monsoon warming over the Himalayan–Gangetic region from 1979 to 2007. Geophys Res Lett 36:L07704. doi:10.1029/2009GL037641

Ghosh S (2010) SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output. J Geophys Res 115:D22102. doi:10.1029/2009JD013548

Ghosh S, Mujumdar PP (2006) Future rainfall scenario over Orissa with GCM projections by statistical downscaling. Current 90(3):396–404

Gillies RR, Wang SY, Huang WR (2012) Observational and supportive modeling analyses of winter precipitation change in China over the last half century. Int J Climatol 32:747–758. doi:10.1002/joc/2303

Glotter M, Elliott J, McInerney D, Best N, Foster I, Moyer EJ (2014) Evaluating the utility of dynamical downscaling in agricultural impacts projections. Proc Natl Acad Sci USA 111(24):8776–8781. doi:10.1073/pnas.1314787111

Goswami BN (2000) Comments on "Choice of south Asian summer monsoon indices". Bull Am Meteorol Soc 81(4):821–822

Goswami BN, Krishnamurthy V, Annamalai H (1999) Abroad-scale circulation index for the interannual variability of the Indian summer monsoon. Q J R Meteorol Soc 125(554):611–633

Groppelli B, Bocchiola D, Rosso R (2011) Spatial downscaling of precipitation from GCMs for climate change projections using random cascades: a case study in Italy. Water Resour Res 47(3):1–18. doi:10.1029/2010WR009437

Han Z, Zhou T (2012) Assessing the quality of APHRODITE high-resolution daily precipitation dataset over contiguous China. Chin J Atmos Sci 36(2):361–373. doi:10.3878/j.issn.1006-9895.2011.11043 **(in Chinese)**

Hay LE, McCabe GJ, Wolock DM, Ayers MA (1991) Simulation of precipitation by weather type analysis. Water Resour Res 27(4):493–501

Hertig E, Jacobeit J (2008) Assessments of Mediterranean precipitation changes for the 21st century using statistical downscaling techniques. Int J Climatol 28:1025–1045

Hertig E, Jacobeit J (2013) A novel approach to statistical downscaling considering nonstationarities: application to daily precipitation in the Mediterranean area. J Geophys Res Atmos 118(2):520–533. doi:10.1002/jgrd.50112

Hewitson BC, Crane RG (1992) Large-scale atmospheric controls on local precipitation in tropical Mexico. Geophys Res Lett 19(18):1835–1838. doi:10.1029/92GL01423

Hines KM, Bromwich DH, Marshall GJ (2000) Artificial surface pressure trends in the NCEP-NCAR reanalysis over the Southern Ocean and Antarctica. J Clim 13:3940–3952

Hughes JP, Guttorp P (1994a) A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. Water Resour Res 30(5):1535–1546

Hughes JP, Guttorp P (1994b) Incorporating spatial dependence and atmospheric data in a model of precipitation. J Appl Meteorol 33:1503–1515

Hughes JP, Lettenmair DP, Guttorp P (1993) A stochastic approach for assessing the effect of changes in synoptic circulation patterns on gauge precipitation. Water Resour Res 29(10):3303–3315

Huth R (1997) Potential of continental-scale circulation for the determination of local daily surface variables. Theor Appl Climatol 56:165–186

IPCC (2007) Climate change 2007: the physical science base. In: Solomon S, Qin D, Manning M (eds.) Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change

IPCC (2014) Climate change 2014: impacts, adaptation, and vulnerability. Part A: global and sectoral aspects. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds) Contribution of Working Group II to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Jenkinsa CN, Van Houtan KS, Pimm SL, Sexton JO (2015) US protected lands mismatch biodiversity priorities. www.pnas.org/cgi/doi/10.1073/pnas.1418034112

Jimenez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. Divers Distrib. doi:10.1111/j.1472-4642.2008.00496.x

Jones C, Giorgi F, Asrar G (2011) the coordinated regional downscaling experiment: CORDEX—an international downscaling link to CMIP5, CLIVAR Exch 56(16–2): 34–40. http://www.clivar.org/sites/default/files/imported/publications/exchanges/Exchanges_56.pdf

Junk J, Matzarakis A, Ferrone A, Krein A (2014) Evidence of past and future changes in health-related meteorological variables across Luxembourg. Air Qual Atmos Health 7:71–81. doi:10.1007/s11869-013-0229-4

Kalnay E, Kanamitsu M, Kistler R et al (1996) The NCEP/NCAR 40-years reanalysis project. Bull Am Meteorol Soc 77(3):437–471

Kanamitsu M, Ebisuzaki W, Woollen J, Yang S, Hnilo JJ, Fiorino M, Potter GL (2002) NCEP–DOE AMIP-II reanalysis (R-2). Bull Am Meteorol Soc 83:1631–1643. doi:10.1175/BAMS-83-11-1631

Kannan S, Ghosh S (2013) A nonparametric Kernel regression model for downscaling multisite daily precipitation in the Mahanadi basin. Water Resour Res. doi:10.1002/wrcr.20118

Kannan S, Ghosh S, Mishra V, Salvi K (2014) Uncertainty resulting from multiple data usage in statistical downscaling. Geophys Res Lett 41:4013–4019. doi:10.1002/2014GL060089

Kerr RA (2013) Forecasting regional climate change flunks its first test. Science 339:638. doi:10.1126/science.339.6120.638

Khalili MF, Brissette Leconte R (2009) Stochastic multi-site generation of daily weather data. Stoch Environ Res Risk Assess 23:837–849. doi:10.1007/s00477-008-0275-x

Kishtawal CM, Niyogi D, Tewari M, PielkeSr RA, Shepherd JM (2010) Urbanization signature in the observed heavy rainfall climatology over India. Int J Climatol 30:1908–1916

Knutti R, Sedláček J (2012) Robustness and uncertainties in the new CMIP5 climate. Nat Clim Change 2:587–595. doi:10.1038/nclimate1495

KrishnaKumar K, Rajagopalan B, Cane MA (1999) On the weakening relationship between the Indian monsoon and ENSO. Science 284(5423):2156–2159. doi:10.1126/science.284.5423.2156

Kumar D, Kodra E, Ganguly AR (2014) Regional and seasonal intercomparison of CMIP3 and CMIP5 climate model ensembles for temperature and precipitation. Clim Dyn 43(9–10):2491–2518

Langenbrunner B, Neelin JD (2013) Analyzing ENSO Teleconnections in CMIP Models as a measure of model fidelity in

simulating precipitation. J Clim 26:4431–4446. doi:10.1175/JCLI-D-12-00542.1
Laprise R (2008) Regional climate modelling. J Comput Phys 227:3641–3666. doi:10.1016/j.jcp.2006.10.024
Laprise R (2014) Comment on "The added value to global model projections of climate change by dynamical downscaling: a case study over the continental U.S. using the GISS-ModelE2 and WRF models" by Racherla et al. J Geophys Res Atmos 119(3877–3881):2013J. doi:10.1002/D019945
Li J (2002) Accounting for unresolved clouds in a 1D infrared radiative transfer code. Part I: solution for radiative transfer, cloud scattering and overlap. J Atmos Sci 59:3302–3320
Li J, Barker HW (2002) Accounting for unresolved clouds in a 1D infrared radiative transfer code. Part II: cloud subgrid-scale variability. J Atmos Sci 59:3321–3339
Li J, Barker HW (2005) A radiation algorithm with correlated k-distribution. Part I: local thermal equilibrium. J Atmos Sci 62:286–309
Li J, Ding R (2011) Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogs. Mon Weather Rev 139:3265–3283. doi:10.1175/MWR-D-10-05020.1
Li H, Sheffield J, Wood EF (2010) Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. J Geophys Res. doi:10.1029/2009JD012882
Lin W, Frere CH, Karczmarski L, Xia J, Gui D, Wu Y (2014) Phylogeography of the finless porpoise (genus Neophocaena): testing the stepwise divergence hypothesis in the northwestern Pacific. Sci Rep. doi:10.1038/srep06572
Litta AJ, Idicula SM, Mohanty UC (2011) A comparative study of convective parameterization schemes in WRF-NMM model. Int J Comput Appl 33(6):32–40
Masello JF, Montano V, Quillfeldt P, Nuhlickova S, Wikelski M, Moodley Y (2015) The interplay of spatial and climaticlandscapes in the genetic distribution ofa South American parrot. J Biogeogr. doi:10.1111/jbi.12487
Maurer EP, O'Donnell GM, Lettenmaier DP, Roads JO (2001) Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model. J Geophys Res 106(D16):17841–17862
Mehrotra R, Sharma A (2005) A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurrences. J Geophys Res 110:D16108. doi:10.1029/2004JD005677
Mehrotra R, Sharma A (2006) A nonparametric stochastic downscaling framework for daily rainfall at multiple locations. J Geophys Res 111:D15101. doi:10.1029/2005JD006637
Mishra V, Lettenmaier DP (2011) Climatic trends in major U.S. urban areas, 1950–2009. Geophys Res Lett 38:L16401. doi:10.1029/2011GL048255
Murphy J (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. J Clim 12:2256–2284
Nadaraya EA (1964) On estimating regression. Theory Probab Appl 10:186–190
Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S, Wada K, Kato K, Oyama R, Ose T, Mannoji N, Taira R (2007) The JRA-25 reanalysis. J Meteorol Soc Jpn 85:369–432
Özelkan EC, Duckstein L (1996) Relationship between monthly atmospheric circulation patterns and precipitation: fuzzy logic and regression approaches. Water Resour Res 32(7):2097–2103
Parthasarathy B, Kumar KR, Kothawale DR (1992) Indian summer monsoon rainfall indexes—1871–1990. Meteorol Mag 121(1441):174–186
Parthasarathy B, Rupakumar K, Munot AA (1996) Homogeneous regional summer monsoon rainfall over India: interannual variability and teleconnections, Research report no. RR-070
Pielke RA, Adegoke J, Beltran-Przekurat A, Hiemstra A, Lin J, Nair US, Niyogi D, Nobis TE (2007) Tellus 59B:587–601. doi:10.1111/j.1600-0889.2007.00251.x
Pincus R, Barker HW, Morcrette JJ (2003) A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. J Geophys Res 108:4376
Pinto JO, Monaghan AJ, Monache LD, Vanvyve E, Rife DL (2014) Regional assessment of sampling techniques for more efficient dynamical climate downscaling. J Clim 27:1524–1538. doi:10.1175/JCLI-D-13-00291.1
Racherla PN, Shindell DT, Faluvegi GS (2012) The added value to global model projections of climate change by dynamical downscaling: a case study over the continental U.S. using the GISS-ModelE2 and WRF models. J Geophys. doi:10.1029/2012JD018091
Raje D, Mujumdar PP (2009) A conditional random field–based downscaling method for assessment of climate change impact on multisite daily precipitation in the Mahanadi basin. Water Resour Res 45:W10404. doi:10.1029/2008WR007487
Ramesh KV, Goswami P (2014) Assessing reliability of regional climate projections: the case of Indian monsoon. Sci Rep. doi:10.1038/srep04071
Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour Res 17:182–190
Road J, Betts A (1999) NCEP–NCAR and ECMWF reanalysis surface water and energy budgets for the Mississippi river basin. J Hydrometeorol 1:88–94
Ryu J, Hayhoe K (2013) Understanding the sources of Caribbean precipitation biases in CMIP3 and CMIP5 simulations. Clim Dyn. doi:10.1007/s00382-013-1801-1
Salvi K, Kannan S, Ghosh S (2013) High resolution multi-site daily projections in India with statistical downscaling for climate change impacts assessment. J Geophys Res Atmos. doi:10.1002/jgrd50280
Schmith T (2008) Stationarity of regression relationships: application to empirical downscaling. Notes Corresp. doi:10.1175/2008JCLI1910.1
Shashikanth K, Salvi K, Ghosh S, Rajendran K (2013) Do CMIP5 simulations of Indian summer Monsoon rainfall differ from those of CMIP3? Atmos Sci Lett. doi:10.1002/asl2.466
Shastri H, Paul S, Ghosh S, Karmakar S (2015) Impacts of urbanization on Indian summer Monsoon rainfall. Extremes. doi:10.1002/2014JD022061
Slonosky VC, Jones PD, Davies TD (2001) Atmospheric circulation and surface temperature in Europe from the 18th century to 1995. Int J Climatol 21:63–75
Sperber KR, Annamalai H, Kang IS, Kitoh A, Moise A, Turner A, Wang B, Zhou T (2012) The Asian summer monsoon: an intercomparison of CMIP5 vs CMIP3 simulations of the late 20th century. Clim Dyn. doi:10.1007/s00382-012-1607-6
Stickler A, Brönnimann S (2011) Significant bias of the NCEP/NCAR and twentieth century reanalyses relative to pilot balloon observations over the West African monsoon region (1940–1957). Q J R Meteorol Soc. doi:10.1002/qj.854 **(published online)**
Stocker TF, Qin D, Plattner GK, Alexander LV et al. (2013) Technical summary. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Doschung J, Nauels A, Xia Y, Bex V, Midgley PM (Eds) Climate Change 2013: the physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, pp. 33–115, doi:10.1017/CBO9781107415324.005
Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498
Timmermann A, Oberhuber J, Bacher A, Esch M, Latif M, Roeckner E (1999) Increased El Niño frequency in a climate model forced by future greenhouse warming. Nature 398:694–697

Tojo B, Kotera A, Nakai K, Nagano T, Kobayashi S, Moji K (2011) Evaluation of recent forest cover change in Savannakhet Province, Lao PDR, using AVNIR-2 and MODIS satellite images. Global Environ Res V15N2:119–130

Underwood FM (2009) Describing long-term trends in precipitation using generalized additive models. J Hydrol 364(3–4):285–297

Uppala SM, Kållberg PW, Simmons AJ et al (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131:2961–3012. doi:10.1256/qj.04.176

Wang B (2000) Comments on "Choice of South Asian summer monsoon indices"—reply. Bull Am Meteorol Soc 81(4):822–824

Wang B, Fan Z (1999) Choice of South Asian summer monsoon indices. Bull Am Meteorol Soc 80(4):629–638

Wang SY, Gillies RR (2013) Influence of the Pacific quasi-decadal oscillation on the monsoon precipitation in Nepal. Clim Dyn 40:95–107. doi:10.1007/s00382-012-1376-2

Wang QJ, Robertson DE, Chiew FHS (2009) A Bayesian joint probability modelling approach for seasonal forecasting of streamflows at multiple sites. Water Resour Res 45(5):1–18. doi:10.1029/2008WR007355

Watson GS (1964) Smooth regression analysis. Sankhya Ser A 26:359–372

Webster PJ, Yang S (1992) Monsoon and ENSO—selectively interactive systems. Q J R Meteorol Soc 118(507):877–926

Wilby RL (1994) Stochastic weather type simulation for regional climate change impact assessment. Water Resour Res 30:3395–3403

Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr 21:530–548

Wilby RL, Hassan H, Hanaki K (1998) Statistical downscaling of hydro-meteorological variables using general circulation model output. J Hydrol 205:1–19

Wilby RL, Hey LE, Leavesly GH (1999) A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River Basin, Colorado. J Hydrol 225:67–91

Wilby RL, Tomlinson OJ, Dawson CW (2003) Multi-site simulation of precipitation by conditional resampling. Clim Res 23:183–194

Wilby RL, Charles SP, Zorita E et al (2004) The guidelines for use of climate scenarios developed from statistical downscaling methods. In: Supporting material of the Intergovernmental Panel on Climate Change (IPCC), Prepared on behalf of Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA). (http://ipccddc.cru.uea.ac.uk/guidelines/StatDownGuide.pdf)

Wilks DS (1999) Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain. Agric For Meteorol 96:85–101

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. Divers Distrib 14:763–773

Xue Y, Janjic Z, Dudhia J, Vasic R, Sale FD (2014) A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. Atmos Res 147–148(2014):68–85. doi:10.1016/j.atmosres.2014.05.001

Yatagai A, Arakawa O, Kamiguchi K, Kawamoto H, Nodzu MI, Hamada A (2009) A 44 year daily gridded precipitation data set for Asia based on a dense network of rain gauges. SOLA 5:137–140. doi:10.2151/sola.2009-035