Growth in mid-monsoon dry phases over Indian region: Prevailing influence of anthropogenic aerosols

Rohit Chakraborty[1], Bijay Kumar Guha[2], Shamitaksha Talukdar[*1], Madineni Venkat Ratnam[1], Animesh Maitra[3]

[1]National Atmospheric Research Laboratory, Gadanki, India,

[2]National Institute of Technology, Rourkela, India

[3]Institute of Radiophysics and Electronics, Kolkata, India

rohitc744@gmail.com, bijayguha74@gmail.com, shamit@narl.gov.in[*]

Abstract

A detailed investigation on the potentially drought prone regions over India has been presented in this study based on the balance between precipitation and potential evapotranspiration (PET) during the South West Asian mid-monsoon season. We methodically introduce a parameter named dry day frequency (DDF) which is found suitable to present the drought index (DI) in mid-monsoon season hence strongly associated with the possibility of drought occurrences. The present study investigates the probable aspects which influence the DDF over these regions revealing that the abundance of anthropogenic aerosols especially over urbanized location have prevailing role on the growth of DDF during last few decades. The prominent increasing trend in DDF over Lucknow (26.84° N, 80.94° E), a densely populated urban location situated in the Indo-Gangetic plain, strongly reflects the dominant association of anthropogenic aerosols with the increasing dry phase occurrences. Increase in DDF (~90%) during the last 60 years is observed over this urban area compared to a broader region in its surroundings. In addition, periodic impacts of large scale phenomena like ENSO (El Niño–Southern Oscillation) or SSN (Sun spot number) become weaker when the study location is downscaled towards an urbanized region. Finally, when long term projections of DDF are drawn using the high urbanization scenario of RCP 8.5 a huge rise in dry days are seen during mid-July to mid-September (reaching up to 50 dry days by the year 2100 over Lucknow) which will be a crucial concern for policy makers in future.

1. Introduction

Drought is a natural and recurrent phenomenon which occurs in all forms of climate. Although similar to aridity in many ways, droughts are mainly temporary in nature thus it should not be confused with the water scarcity due to excess of water demand over available supply. On the other hand these weather extremes are more reasonably linked with the distribution and frequency of rainfall over any region. Although, there are no generally accepted definitions for drought, the American Meteorological Society has categorized it into four types namely: meteorological or climatological, agricultural, hydrological and socioeconomic (Heim, 2002). A prolonged drought lasts several months or even years while the absence or reduction of precipitation creates meteorological droughts. On the other hand, short-term (few weeks) dryness in surface layer could results an agricultural drought (Heim, 2002). However, when prolonged meteorological droughts reduce the ground water level severely then hydrological droughts occur. Finally, all first three droughts with a deficit in water availability are named as socioeconomic drought. Among these four, the agricultural drought might be a serious issue when the farming or crop producing in humid or sub humid zones are concerned. The situation has however become more serious in the present due to rapid population growth across all continents, thereby also producing a hike in their global demand (Sivakumar, 2011).

India is a country where agriculture and its allied activities act as major source of livelihood and hence it is expected to be deeply affected by drought occurrences especially if it occurs in the mid-monsoon period (as it experiences ~80% of the annual rainfall due to the southwest monsoon). Generally drought events originate from the deficiency in precipitation, and water shortage over a particular region and time. As rainfall observation data is available from past two centuries, mostly all the calculations of drought indices includes this variable either single headedly or in combination with other meteorological parameters (WMO, 1975). Some early drought index were simply represented the drought duration or intensity upon satisfying the drought defining criteria, e.g. Munger (1916) defined the drought index as the length of period without 24 hours precipitation with a minimum of 1.27 mm. Marcovitch (1930) used temperature data along with the precipitation while Benton (1942) used the length of drought in days, where the count was terminated upon occurrence of 2.54 mm of rainfall over a span of 48 hours. Likewise, many other drought index can be found in the past literature where precipitation has been used as a primary factor (Palmer, 1965; Lloyd-Hughes and Saunders, 2002). Recently, the multi-scaler drought index like Standardized Precipitation Index

(McKee et al., 1993) is widely used by several researchers in analysing the drought characteristics. However, no single index has the ability to precisely represent the drought duration and intensity and its possible impacts (Wilhite and Glantz, 1985). Again, apart from the rainfall, there are also some other parameters that affects the drought severity, e.g. potential evapotranspiration (PET) and soil water holding capacity (Dai et al., 2004).The Palmer Drought Severity Index (Palmer, 1965) is an effective parameter which uses all these three parameters; however, it has some limitations when applying over climatic zones like India (Kumar et al., 2013). In addition, gathering all these parameters in gridded form and then quantifying the drought index will be very difficult over the Indian region. On the other hand, the standardized precipitation–evapotranspiration index (SPEI) uses only precipitation and temperature, and is considered to be better for analysing drought occurrence (Begueria et al., 2010).

India happens to be one of the most vulnerable drought-prone countries, as severe droughts occur at least once in a three year time span since the past few decades. In addition, there are numerous instances of severe drought conditions during Monsoon as reported in recent past (Pai et al., 2011). Consequently, several studies have been carried out in the recent years in order to understand the drought occurrences during the Indian summer monsoon period (Gore and Sinha Ray, 2002). Bhalme and Mooley (1980) defined the Drought Area Index for drought intensity assessment using monthly rainfall distribution. Raman and Rao (1981) suggested a possible relation between summer droughts and prolonged brake phase of southwest monsoon over the Indian sub-continent. Parthasarathy et al. (1987) identified the extreme drought years by analysing the decade long anomalies in the Indian summer monsoon rainfall. Tyalagadi et al. (2015) analysed more than 100 years of rainfall and identified 21 drought years, half of which were associated with El Niño. Gadgil et al. (2003) explained the excess rainfall or drought in terms of Equatorial Indian Ocean Oscillation (EQUINOO) during 1972 – 2002, especially during monsoon season. Francis and Gadgil (2010) also suggested the role of El Niño Southern Oscillation (ENSO) and EQUINOO behind the 48% deficit of June rainfall over India., Apart from these oscillations like ENSO or IOD (Indian Ocean Dipole) there are also lots of other parameters which may have prominent influences on drought occurrence, e.g. Himalayan ice cover, Eurasian snow cover, the passage of intra-seasonal waves, effects of accumulated pollution etc., e.g. Krishnamurti et al. (2010) reported the intrusion of desert air mass to be responsible towards the drought occurrences over the central Indian region.

98   In general, most of the previous studies on monsoon droughts are discussed on the basis of
99   rainfall accumulation, and there are very few, which quantify its relation with the direct or
100   indirect radiative effects of aerosols (Twomey, 1977) while considering both rainfall and PET.
101   Absorbing aerosols such as black carbon (BC) or dust have the capabilities of atmospheric
102   heating by absorbing solar radiation, while non-absorbing aerosols (e.g. sulphates) scatter the
103   solar radiation have less effect over the same (Lau and Kim, 2006). Additionally, they have the
104   capability of modulating the cloud characteristics by altering cloud radiative properties
105   (Wencai et al., 2015). Previous studies have shown the presence of the aerosols (mainly dust
106   and BC), and their ability to impact the rainfall (depending upon their sizes) during Indian
107   summer monsoon as described by elevated heat pump hypothesis (Solmon et al., 2015). During
108   late pre-monsoon or early monsoon season, the aerosol loading over India is nearly three times
109   higher than the average due to the dust abundance, which is partly dependent upon the winds,
110   precipitation and surface temperature (Dey, 2004). However, the vice versa can also be true
111   (e.g. Moorthy et al., 2007). Very recently some new attempts were also undertaken to study the
112   long and short term implications of both natural and anthropogenic components in producing
113   several atmospheric processes in the boundary layer which thereby produces hindrance to
114   convective rainfall especially over urbanized coastal locations which may also lead to
115   subsequent drought occurrences (Chakraborty et al., 2017 and Talukdar et al., 2017). Keeping
116   all these assertions in mind, the present study has put an effort in establishing a possible
117   relationship between aerosol loading and summer monsoon rainfall, consequently, over
118   drought occurrences during this period in past few decades.

119   Hence a detailed investigation is presented on the evolution of dry phase leading to
120   drought conditions during mid-monsoon over three Indian regions based on the balance
121   between precipitation and PET during the monsoon season. Next, a new parameter called dry
122   day frequency is used to understand the trends of drought potential over the mentioned Indian
123   regions. This is followed by a three pronged investigation to identify the most dominant factor
124   behind these trends after which future projections of DDF is observed and explained for these
125   locations during the mid-monsoon period.

126   **2. Dataset and methodology**

127   Most of the research attempts in recent past have employed SPEI as an indicator of drought
128   occurrence over the Indian region (Beguería et al., 2010). SPEI which is precipitation minus
129   PET mainly represents the climatic monthly water budget. Interestingly, this parameter is
130   found to be the most reliable identifier of drought occurrences as it can be expressed in terms

131 of standardized Gaussian variance with zero mean and one standard deviation. Another

132 advantage of using SPEI over any other multi-scalar drought indicators (e.g. SPI) is that it not

133 only includes the effect of the evaporative demand in its calculation, but also can be calculated

134 for different time scales (Beguería et al., 2010), unlike the PDSI which rely on a water balance

135 of a particular system. In this study the SPEI is calculated using monthly precipitation and PET

136 from the CRU TS3 dataset (http://badc.erc.ac.uk/data/cru/), where the PET is calculated

137 considering the monthly mean temperature and the geographical location of the concerned

138 region as per the method suggested by Thornthwaite (1948). Hence, it provides long-term

139 information about the drought conditions over any location with a high spatial resolution of

140 $0.5°\times0.5°$ on monthly basis. However, the available precipitation (P) data is provided in the

141 form of monthly accumulated value, whereas, the PET represents the monthly mean.

142 Therefore, the difference (D) or SPEI is calculated for each month as follows:

$$D = P - (PET \times \text{number of days in a month}) \qquad (1)$$

143 It may be noted that for this analysis, the value of D is normalized with respect to the

144 climatic mean and 1 sigma standard deviation to obtain comparable values for all regions of the

145 country. These normalized values of D are hereafter referred to as DI. This study considered

146 the length of the dry phase as an indicator of drought occurrence and severity, which is

147 calculated from $0.25°\times0.25°$ daily gridded rainfall datasets as in the National Data Center,

148 India Meteorological Department (IMD) (Guhathakurta and Rajeevan 2008) during the period

149 of 1901-2015. Owing to its better temporal and spatial resolution, the IMD rainfall dataset has

150 been used in several research attempts in the recent past for analysing the morphology of

151 drought occurrences over India (e.g. Gore and Sinha Ray, 2002). In previous literatures there

152 have been various mentions for identifying certain days as dry, based on some predefined daily

153 rainfall accumulation thresholds. Singh et al. (2010) has mentioned that days having rainfall

154 less than 5mm/day can be considered as dry. But this criterion is only valid for ecological

155 droughts and hence it will not be a suitable threshold for many Indian regions experiencing

156 very low rainfall. Recently, another classification scheme has also been attempted by Sushama

157 et al., (2014) where rainfall accumulation lower than 1 or 3 mm/day is considered as a dry day.

158 So, to further check which threshold provides best results, the correlation coefficient of DI

159 verses DDF are plotted in **Table S2**. The correlation coefficients follow some spatial diversity

160 but interestingly, they do not exhibit much change with respect to the rainfall threshold. Hence,

161 to understand its implication, the number of days having rainfall accumulation above 1 and 3

162 mm (during JJAS) is expressed in the form of ratio in **Figure S1**. The ratio indicates that for

all the months and regions, days having rainfall accumulation above 1 mm/day are more in number compared to the days having rainfall accumulation above 3 mm/day. This makes it reasonable to put 1 mm/day as threshold rainfall accumulation for DDF consideration as it will filter out only the intensely dry conditions which will make the drought identification more reliable. Hence, this study is progressed using 1mm/day as the dry day identification threshold. Further, the DI values obtained are normalized with respect to mean and standard deviation for simplicity. Data sets of number of dry days and drought index are passed into three dependence tests: first, using three equal sized grouped box whisker distributions; second by principle component analysis of variances of two main contributors. The third and final approach involves a multi-linear regression in order to see the net contribution of the various components on dry or wet condition,

Datasets of sunspot numbers are considered here as a reliable representative of solar activity, which in turn may modulate the earth's hydrological balance. There have been several scientific mentions in the past underlying the effect of solar intensity on tropical rain and monsoon strengths both over India and abroad (Agnihotri et al., 2001). Monthly averaged sun spot numbers are obtained from the Solar Influences Data analysis Center (SIDC) in the Royal Observatory of Belgium from the year 1749 till present (Cliver et al., 2013). This study also considered ENSO index, obtained from the Oceanic Niño Index (ONI), which is calculated using 3 month running mean of Extended Reconstructed Sea Surface Temperature, Version 5 (ERSST.v5) SST anomalies in Niño 3.4 region (5°N – 5°S, 120° – 170°W) with a 30-year base period (Huang et al., 2017). Conditions resulting in values beyond the threshold of ±0.5°C are considered to be either an El Niño or La Niña. These datasets are obtained from 1950 to present. Present study also uses 0.5°×0.625° gridded datasets of AOT at 550 nm, Black Carbon (BC), dust (pm2.5 only), Organic Carbon (OC), sea salt and sulphate obtained from MERRA-2 (Modern-Era Retrospective analysis for Research and Applications version 2) provided by NASA. MERRA-2 provides global reanalysis product since 1980 to present (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/). The reliability of the aerosol products from MERRA-2 have already been authenticated in studies like Buchard et al. (2017) and in recent years, these datasets have also been utilized for similar climatic investigations over the Indian region (Pandey et al., 2017). Moreover, in this study an additional validation of the aerosol components from MERRA2 have been validated against in-situ datasets from Aethelometer measurements over Kolkata. However, to preserve the parity with monthly averaged Black Carbon Extinction as in MERRA2, the observation datasets are also monthly

196   averaged for a net period of 36 months during 2013, 2015 and 2017. Consequently, a well

197   matching is observed between the two sources as shown in **Figure S2**. To double check, the

198   datasets of BC AOT and concentrations are both normalized and then their probability

199   distributions are plotted. The distributions fitted with Gaussian curves shows almost similar

200   behaviour in both the cases, which shows the suitability of this datasets in subsequent sections.

201   In addition, the ERA Interim reanalysis cloud cover data is utilized

202   (http://www.ecmwf.int/) at 0.75°×0.75° default resolution (Beriford et al., 2011). As DDF is

203   being observed mostly over the month of August, hence monthly averaged data of total, high,

204   medium and low cloud covers are extracted over the required regions and are plotted for the

205   same time period (as in for aerosol parameters) during 1980-2015. The idea behind using this

206   dataset was to identify the association between increased cloudiness and reduced rain

207   accumulation during the mid-monsoon months. Additionally, in order to show its relation with

208   cloud microphysics, dataset of cloud particle radius are utilized, and is obtained from NASA

209   Earth Observation (NEO) portal

210   (https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MODAL2_M_CLD_RD). The dataset

211   provided by Terra/ Aqua Satellite of MODIS on daily, weekly or monthly basis with a good

212   spatial resolution of 1°×1°, and is available only over a relatively shorter span of 2000 – 2018.

213   The monthly averaged values of CER have been utilized during the month of July and August

214   for the present study.

215   This study uses gridded population density (as a proxy of urbanization), obtained from

216   Gridded Population of the World (GPWv4), and provided by the CIESIN-SEDAC database

217   from Columbia University for the year 2000, 2005, 2010 and 2015. This data set is constructed

218   by extrapolating the population data from national or sub-national administrative units all

219   around the world. The resolution of the product is 30 arc-seconds, or approximately 1 km at the

220   equator, further details about the data can be obtained from

221   http://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-rev10.

222

223   **3. Results and Discussion**

224   **3.1. Identification of potentially drought prone regions over India**

225   Considerable conditions for drought occurrences are identified on the basis of the

226   balance between monthly PET and rainfall accumulation during June-September as depicted in

227   **Figure 1**. It is seen that due to arid climates, north western India experiences higher values of

228   PET particularly up to July which may happen due to late arrival of monsoons at that location

7

229 and hence this region may be considered for the analysis. On the other hand the south eastern
230 peninsula of India experiences higher PET values, hence it has been considered for further
231 analysis. However, the rest of the country experiences much lesser values of PET. In contrast,
232 precipitation values are consistently lesser both in the north western India as well as the south
233 eastern peninsula, so both these regions may face more probability to experience negative DI,
234 hence are selected for analysis. Another highlight from the figure is that, the mid-section of
235 IGP depicts a sharp gradient of precipitation. This diversity becomes more prominent during
236 the months of July-August as during this period, the entire IGP experiences very heavy rain
237 accumulation (>300 mm on average) but the mid-IGP experiences much lesser rainfall ~200
238 mm. Consequently, this mid IGP region is also selected for analysis. Accordingly, the grid
239 points with 0.5 degree resolution in these three regions are identified and accumulated to form
240 three main study regions which are numbered 1, 2 and 3 corresponding to IGP, South Eastern
241 peninsula and North West India, respectively as shown in **Figure S3**.

242 **3.2. Importance of dry day frequency (DDF) in analysing the drought conditions**

243 After the identification of the drought prone regions, the main objective is to determine
244 a suitable parameter which best represents the probability of droughts and which also can be
245 related to other natural and anthropogenic factors in all regions. Hence an assumption is taken,
246 that if the temporal distribution of rainfall is considered constant month wide, then a drought is
247 only possible when both PET is high and precipitation is low. Now low precipitation and high
248 PET mainly arises from multiple dry day occurrences in a month leading to droughts. So, for
249 simplicity, during each of four months in three seasons, the difference between precipitation
250 and PET is calculated over 115 years and the obtained data is normalized with respect to mean
251 and 1 sigma std for simplicity, Next, the DDF time series is calculated from daily precipitation
252 values as already described previously after which the correlation coefficients between drought
253 index DI and this dry days frequency are calculated and shown in **Figure 2(a)**. The correlation
254 analysis is done for two overlapping periods of 115 and 60 years namely: 1901-2015 and 1956-
255 2015. The reason for this two part analysis is that during the second part, more technological
256 advancement may lead to more reliable daily rainfall data, this is because during recent years
257 the advent of more accurate rain gauges have led more reliability in deciding whether daily
258 accumulation < 1mm and thus more reliable dry day frequencies are calculated. Another reason
259 is that, second part witnessed more station and satellite data sources, so possibility of
260 relationship is expected to be stronger in last 60 years. However, to bear better evidence to the

above stated hypothesis **Figure S4** shows scatter plots of DI and dry day frequencies for all regions and months.

The correlation analysis depicts a set of reasonable correlation coefficients in both regions 1 and 3 over 115 years. Importantly, better correlation values are observed typically over July in region 3 and August in region 1 while it is lesser in all other cases. It may be noted that regions situated in the western and north western parts of the country experiences delayed monsoon (supported from many independent sources) which have led to high correlation values during June and July over region 3. However, region 1 specially shows good correlation in August which is mid monsoon month which need more attention in coming sections. Here, it may be noted from Figure S4 that the agreement between DDF and DI is not strong in most of the cases other than 3-4 instances only. The reason is that firstly, DI is dependent on the monthly accumulated difference between precipitations and PET while DDF depends upon the erratic distribution of daily rainfall accumulation, hence the temporal scales of these two parameters are different from each other. The second reason is the presence of an independent factor called PET which depends upon a set of various components (location, season, vegetation and soil type, temperature, moisture content, wind speed, surface pressure and net radiation flux) but not on precipitation. Hence this explains the disagreement between the two parameters on a climatic scale.

Considering the last 60 years, correlation coefficients are improved in all regions and months as expected. Region 3 shows high correlations in July followed by August, while region 1 depicts comparatively much higher values during July and August. Thus the consideration of delayed monsoon onset may bring out more dry days in region 1 and 3. But on the other hand region 1 shows a high association between DI and dry days in August which needs to be studied. Region 2 is mainly influenced by precipitation occurring during the late monsoon months i.e. September and not by the mature monsoon stage which is evident from the higher correlation values at that time. Hence this region may not fit with the scope of the present study. Additionally, as more significant correlation values are obtained during 60 years span compared to 115 year scale, hence DDF trends will be studied over the last 60 years span in the coming sections.

3.3 Determining the time spans for the analysis of DDF trend over mentioned regions

3.3.1. The importance of partitioning Region 1 for further analysis

9

It can be seen from the preceding sections that the correlation between DI and dry days for region 1 is noticeable but it is not highly prominent due to the presence of many outliers in the scatter plots (Figure S4). This is because region 1 encompasses a total spatial coverage of 5^oX8^o which has a lot of topographical and climatic diversities between them. A better example of this has already been depicted from the precipitation diversity in Figure 1 where the precipitation gradient was found to change abruptly even within region 1. So these spatial diversities can interrupt the association between droughts and dry days. Hence to have more realistic investigation, the region is now partitioned horizontally along 81.25^oE which lies in the middle of IGP. This gives rise to two different regions in the east and west of region 1 which will hereafter be referred as region 1a and 1b, respectively. Next the total distribution of data of DDF for 1a, 1b, 2 and 3 are again investigated for two overlapping periods 1901-1960 and 1956-2015 for parity in **Figure 2(b)**. Region 2 and 3 show almost no change in the distribution before or after 1955, hence it is not given importance. Region 1b shows sight increase in mean and median but with no prominent change in the distribution while the same thing is very prominent over region 1a. In this case, since last 60 years the mean and median values changed by more than 4 days which is a very alarming fact. Most importantly, the upper quartiles and whisker have ascended to a maximum value of 30 days, which indicates severe drought occurrences. Thus it can be inferred that region 1a emerges as a prominent drought prone region showing an abrupt rise in DDF especially over the month of August and hence it will be investigated in detail in the coming sections.

3.3.2. Analysing the climatic trends of DDF using a 15 day window

It has already been discussed that the drought intensity has significant correlation with DDF on a monthly basis. However, it is also necessary to investigate whether the intra-monthly distribution of rainfall may also have its own impact in modulating the dry day frequencies especially during the mid-monsoon months which experience maximum precipitation variability. Hence, the monsoon months (JJAS) are now divided into 8 equal slots of 15 days each and the 60 year time series for all these regions are obtained. Next the robust-fit trend analysis at 95% confidence level is done to find the mean yearly trends, which is multiplied by 60 years and then normalized respect to mean to generate a percentage wise change in DDF.

The percentage changes are shown in **Figure 2(c)** which depicts an overall increase of DDF for all regions with a few exceptions. Region 2 shows very weak trends (< 5%) all throughout monsoon, however, by the end of September, a reasonable trend of ~20% is seen

which may link to dry phase developments in the later months. However, this period falls at the declining phase of monsoon which is beyond main scope of this study; hence neglected.

Region 3 shows quite weak but alternating dry day trends over June followed by the month-long increase in July. This indicates a probable change in the timing of monsoon rainfall over region 3. However, this cannot be firmly confirmed as there is no particular time slot having a prominent trend value (all cases showing trends < 5%). Rainfall in June and early August lead to dry region conditions over July, but the cumulative monthly growth in July is ~10% which is not very strong enough and hence it will be discussed later in the study.

Finally, Region 1 shows very strong increasing trends in dry day frequencies with similar pattern over 1a and 1b. Both these sub-regions experience relative wetting at late June, followed by a prolonged dry phase up to September. But the main difference between the two sub-regions is that the trends are consistently high all throughout in 1a with as much as 60% and 20% increase over August which also continues onto September; while in region 1b the trend values are comparatively lesser (40% and 5%) during August. Thus, it can be inferred that though a clear increase in DDF is obtained all throughout region 1 during July-September, yet the trends are relatively stronger in region 1a especially during August which demand primary importance throughout the study.

3.4 Investigating the probable influence of natural and anthropogenic components on DDF for region 1a

In light of the previous sections, the probable influences behind the increasing trends in dry day occurrences are investigated over region 1. Number of natural or anthropogenic factors may be responsible for this phenomenon. While natural factors mainly include the effect of solar activity, ENSO variability or moisture tendencies, the anthropogenic constituents mainly include aerosols which again encompass a lot of organic and inorganic pollutants. Now, to quantify the effect of aerosols, the aerosol extinction coefficient values can be utilized from either satellite observations (MISR) or from dedicated model simulations (MERRA2). Since observational datasets from MISR satellites are very sparse during monsoon season and also the total measurement period is only 16 years, hence MERRA 2 datasets are used for further analysis. Keeping the availability of AOD datasets in mind further analysis has to be concentrated on 36 years span between 1980 and 2015. Owing to the prominence of DDF trends during the month of August, further studies are concentrated on this period only. As already mentioned, natural factors like solar activity and ENSO oscillations (hereafter referred

as SSN and ENSO) may have some impact on precipitation variability which is also supported

from previous attempts taken, hence they are considered. Additionally, moisture content also

directly controls precipitation and so their monthly means at 850 hpa (corresponding to

maximum moisture content during monsoon) are also utilized from MERRA 2 reanalysis

database. To understand the dependence of these factors on DDF, first, the monthly DDF

values during August 1980-2015 are arranged in descending order and then the sorted dataset is

divided into three equal groups as Short Dry Phase (SDP) corresponding to normal conditions

(8-10 days with average of 9), then Medium Dry Phase (MDP) signifying near drought (10-14

days, with average of 12.5 days) and Long Drought Period (LDP) which represents a full

drought conditions (14-18 days, average ~ 16) as depicted in **Table 1** and they are also

hereafter mentioned as SDP, MDP and LDP. Next, for all these three groups, the distribution of

total aerosol extinction (AOT), SSN, ENSO index and SHUM at 850 hpa are shown in the

form of box plots in **Figure 3**. It is seen that as DDF increases, the distribution of total aerosols

start increasing, as evident from the rise in median and upper whisker values. The variation of

SSN is almost random in all cases hence neglected. Additionally, ENSO intensity changes

fairly with droughts. The upper whiskers and median rises slightly, but its effect is doused due

to a dominant overlapping between the groups which fails to indicate a clear relationship.

Specific humidity shows a minor decrease in all groups, though the median and quartiles do not

show any prominent change (from 15 to 13 g/kg). Hence the importance of this factor cannot

be ascertained.

As the dry phase length distribution fails to identify the dominant factor behind the rise

of dry days in region 1a, hence all these four factors are passed through principal component

analysis test (PCA) and the results are shown in **Figure S5(a)**. The analysis produced a set of

three orthogonal components out of which pc1 and pc2 account for 50 and 25% of variances so

we can neglect the contribution of the 3rd component. Next, the corresponding variance scores

of these components are plotted in Figure S5 which shows that SSN and humidity have very

less variance according to the pc1 axis hence considered as less important, but aerosols and

ENSO have comparatively higher values so they are considered important for further analysis.

Further, multi linear regression analysis is done to see the independent contribution of

these four parameters to DDF. All datasets are normalized so as to get uniform variability to

enable easy identification of the dominating factors. The MLR concludes that the coefficients

for aerosol, SSN, ENSO and SHUM are 0.393, 0.008, 0.161 and -0.207 as shown in **Table 2**,

SSN does not show any effect hence finally rejected. ENSO and specific humidity have

significant contributions but in opposite manner and also their distribution analysis showed significant overlapping; hence they should not be considered in order to remove ambiguity. Finally aerosols have a coefficient of 0.393 which is much higher than the others as also observed in the PCA test and distribution analysis. Hence one has to consider aerosols as more dominating factor compared to other natural components in modulating dry day occurrences.

In view of the dominance of total aerosol AOT over DDF, now the analysis is concentrated on the datasets of various aerosols components over region 1a..Total columnar extinction values of 5 aerosol components namely: black carbon (BC), Dust PM2.5, organic carbon (OC), Sea Salt and Sulphates are obtained from MERRA 2. BC and OC mostly comes from anthropogenic sources and significantly contribute in warming up the atmosphere. It has been reported in earlier studies that the presence of BC aerosol in rain cloud may have "burn off" effect on the cloud due to heating [Ackerman et al, 2000]. On the other hand aerosols like PM2.5which may have both natural and anthropogenic sources can also influence the cloud life time by increasing cloud droplet number (Sato et al, 2018). Thus, the cloud cover is modulated and precipitation process is affected. Now the change in concentration of these parameters during last 36 years over region 1a has been discussed in the next section.

Though it has been discussed in the previous sections that aerosols have a dominating influence over dry day occurrences, however, it is yet to be specified which type of aerosols (natural or anthropogenic, organic or inorganic) are becoming major influencing factor for this phenomenon over region 1a. Hence time series datasets of these five components are again taken for 36 years and are grouped with respect to the corresponding dry day ranges as already explained in previous section. After that the corresponding distributions are plotted in box plots in **Figure 3**. The distribution analysis depicts that the sea salts show some overlapping which reduces the impact on DDF. Sulphates have quite high values all throughout but their medians or distribution does not exhibit any deterministic sequence (first decrease then increase); so they also cannot be used here. Dust AOT values are less but its median shows weak contribution towards drying, but the overlapping in distribution makes the overlap association very weak. But compared to the others BC and OC have shown a better association with DDF along with reasonably increasing tendencies in medians and quartiles. But this phenomenon also hints towards a dominant component of pollution coming from certain highly urbanized sectors of region 1a such as Lucknow, Allahabad (25.43° N, 81.84° E) and Varanasi (25.31° N, 82.97° E). Again out of these two, BC has relatively better variation as it has the least

421 overlapping nature so it may be considered the most dominant factor. But still to have better
422 evidence, the PCA and regression analysis are attempted.

423   The PCA analysis results are depicted in **Figure S5(b)** which shows the contribution of
424 pc1 alone is 60% followed by pc2 of 25% to be more prominent hence there may not be a need
425 to study pc3 here. From the scores it is found that sulphate and dust behave similarly in their
426 variances with high pc1 and low pc 2 values, but OC and BC have both high pc1 and pc2
427 components, so they may be found responsible for the variability in dry day changes. However,
428 sea salt also may have some influence but it is not much clearly understood from the figure.

429   To clarify any remaining misconceptions, the MLR coefficients are computed which
430 gives the values as 0.542, 0.129, 0.263, 0.326 and 0.124 (shown in **Table 3)**. It is expectedly
431 obtained that the dust and sulphate have very less contributions so should be neglected. BC,
432 OC and sea salt have higher values, of which OC and sea salt have comparable magnitudes,
433 but, sea salt has much less AOT values with lesser pc1 variance score and also reasonable
434 distribution overlapping, so the effect of OC may be considered better. BC has very high MLR
435 coefficient with high pc1 score and also a clear variability of distributions. Hence, it may be
436 concluded that owing to urbanization, the effect of BC followed by OC has much stronger
437 association with drought intensity and dry day occurrence.

438 **3.5. Investigating the probable influence of natural and anthropogenic components on**
439 **DDF over Lucknow**

440   From the previous section, it has surfaced that anthrpogenic emissions (of BC and OC)
441 as a result of urbanization may have a dominant association with the increase in DDF during
442 August. To be definite about this, a re-investigation has been done over Lucknow (26.8^{o}N,
443 80.9^{o}E) which is the state capital of the state Uttar Pradesh, and is a more urbanized point
444 location belongs to region 1a. However the relationship of DDF with SSN, ENSO and SHUM
445 is not shown as Lucknow already falls in region 1a whose synoptic effect would not change
446 within the region. Here, the effect of individual aerosol components is also depicted in the
447 distribution analysis as shown in **Figure 4**. Now, in case of Lucknow the variability in dry day
448 values are much stronger as shown by SDP (4-12 dry days average 9.5) MDP (13-17 days with
449 average 15) and LDP (18-30 days with average at 22 days) mentioned in Table 1. The
450 distribution analysis on total aerosol AOT shows much larger values over Lucknow than in
451 region 1a and also the variability of the median values with the quartiles and whiskers are also
452 far more deterministic here which may have influenced the entire distribution towards more

14

453 dry conditions. Next, coming to sea salts and sulphates, they have much less values than in
454 region 1a due to its significant distance from the seas. Sulphates show no meaningful variation,
455 hence are rejected straightaway, sea salt values are less but the variation of median and upper
456 whisker shows a prominent increase which may be important. However, the lower quartile is
457 very small and overlapping in all three cases which serve as a setback to its variability.
458 However, Dust does not such variations due a considerable overlapping in it. On the other
459 hand, BC and OC do not have much overlapping and they also have clear increase in medians
460 and both quartiles thus supporting the more sensitivity of this region towards dry days.

461   **Figure S6** shows the distribution analysis of these components with PCA tests. The
462 analysis reveals the presence of three strong principal components where pc1 is 60% and pc2
463 of 30%; hence pc3 is not considered further. Next, when the variance scores for these
464 parameters are plotted, then all factors show almost similar values of pc1 score, so pc2
465 becomes important. While judging the pc2 scores, we see that BC followed by OC has the best
466 variability in this set hence they may be considered for the dry day variation. To confirm this,
467 multi linear regression is done on the components and the results yield values of 0.864, 0.218,
468 0.556, 0.0106 and 0.155 (Table 3). According to previous results, the contribution of BC and
469 OC is much higher than the others, with BC shows a higher correlation in all cases compared to
470 OC, hence the dependence of dry days can be primarily associated with urbanization. Dust
471 follows this parameters but its dependence is comparatively much smaller than both BC and
472 OC which further supports these findings.

473   In the previous sub-section, the effect of aerosols with BC in particular is found to be
474 strongly associated with low rainfall occurrence. However, the effect of all meteorological
475 parameters was not isolated in the previous analysis. Also a time series analysis showing the
476 impact of present AOD on impending rainfall accumulation was not demonstrated earlier.
477 Hence, an attempt has been progressed over Lucknow as it is an urbanized location in Region
478 1a. Now, to isolate the effect of various meteorological parameters such as temperature,
479 pressure, winds, moisture content and rainfall accumulation, these datasets have been collected
480 and then plotted in Figure S7 for 16-30 July of 1980-2015. The long term mean and 2 sigma
481 standard deviations are also shown to exclude years having abnormal weather conditions. The
482 screening process revealed that three years: 1980, 1987 and 2002, have exhibited
483 meteorological variations beyond the general range, hence they are obliterated.

484   In the next attempt, it was required to see the effect of low rainfall periods and AOD on
485 impending DDF for the next few days during on these years. Hence a set of years having

15

comparatively lower rainfall accumulation during 16-31 July were identified. A total of 16 years were recorded which had rainfall values between the 50th and 25th percentile of the population. It may be noted that certain years experienced rainfall below the 1st quartile and hence they were neglected to preserve the data uniformity. Next the average AOD values were accumulated for those years and interestingly, two well separated clusters having a set of non-adjacent 8 years in each were observed: one with AOD below 0.3 and other above 0.4. To study the effect of these two AOD clusters on rainfall, their corresponding DDF values are observed for the next 15 days (1-15 August). This time shift was employed in order to investigate the net effect changing AOD on impending rainfall distributions. It was observed that DDF values are distinctly higher for high AOD compared to the lower AOD case. Hence this supports that higher AOD necessarily leads to more DDF in next few days.

3.6. Comparative analysis on the DDF trend of last 60 years and cloud properties among Region 1, 1a and Lucknow

The preceding sections have given an idea of how urbanization is influencing the evolution of dry day occurrences. But to understand quantitatively its climatic impact now the averaged DDF of last 60 years are plotted for regions 1, 1a, Lucknow. In order to examine the change in DDF patterns as one downscales from a broad synoptic scale (IGP) to a small localised urban location. **Figure 5** reveals that region 1 has a weak but discernible increase from 5 to 15 days in last 60 years. When robust-fit analysis was performed, it was inferred that the net change in dry day frequencies over region 1 is ~35% with respect to the 60 year average. However, the existence of some periodicities in the data was observed while no evident extremes were observed in the time frame. The value of the slope is found to be less (0.074) which leads to a poor r of 0.384. For region 1a the total variability is from 5 to 18 days; so the slope is expected to improve a bit (with a robust-fit net trend of ~44% with respect to the average) while the periodicity seems to be apparently disturbed due to presence of more data extremes. Finally, in case of Lucknow, huge change is observed from 4 to 25 days which indicates a complete shift in rain climatology with trend values as high as 61% with respect to 60 year average during August when normally, the maximum rainfall occurs over India. Huge number of outliers and extremes are seen some of which are close to 30 days (indicating no rain over August at all). The periodicity also seems to be disturbed due to outliers resulting in a very sharp slope of 0.139 per year. Thus, the severity in drought climatology is well explained with respect to urbanization as already hypothesized earlier. But it may be noted that the increasing trends and correlations are mainly caused by more occurrence of high dry days in

present rather than a gradual rise in the mean values; additionally there are also some periodicities in the signal which results in the correlation being less than 0.5.

It is reported earlier that increase in anthropogenic aerosols may lead to more number of CCN causing reduction in cloud particle radius which may result into less-occurrence of rain in spite of the increase in cloud cover. From previous section it is clear that dry day frequency exhibits a definite increase in magnitude over region 1a and Lucknow. Since anthropogenic components have shown highest possible dominance on dry day occurrences, so an attempt is made to identify how cloud parameters has changed with time over region 1, 1a and Lucknow having different urbanization growth and so on the anthropogenic components. Region 1 which is covering a broad area does not show prominent change in DDF and it is also observed that that the change in cloud cover over region 1 (~ 2%) and reduction in cloud particle size are very feeble. But interestingly as the region of concern is downscaled to Region 1a followed by a further downscaling to a region the urbanization impact becomes prominent and that is also reflected in the observed cloud parameters. A significant increase in cloud lifetime is observed which has resulted in a more definite growth of mid and low level clouds. The situation however, becomes more prominently worse in case of Lucknow where the cloud cover increased consistently (~18%) reflecting the impact of urbanisation. As a consequence, the dry day frequency ascends at a rapid rate over Lucknow in spite of increasing cloud cover which definitely needs to be studied in more detail in future approaches.

The long term trends of dry day occurrences have exhibited a prominent growth in dry days but the effect of this trends were found to be subdued to some extent by several periodicities over the last 60 years in both region 1 and 1a. To understand their role to a quantitative scale, periodicity analysis is done on last 60 years using autocorrelation functions and the results are depicted in **Figure S8**. The ACF values show highest value of 1 for a time lag 0, hence it is removed. Also there is no use in understanding periodicities greater than half of the period hence the maximum period is fixed to 30 years. 1 sigma bars are provided to understand which periodicity may be significant enough to impact the long term trends. Figure shows that the ACFs are reducing with time for all regions just as expected. However, only two points are found considerable, one is at the local maxima of 4 years corresponding to ENSO, where as expected the synoptic influences will be stronger in larger spatial scales. Another periodicity is expected to lie at ~1-2 years which represents the year-year varying component of urbanization. However, this effect is found to be much lesser in region 1 as it has a much higher spatial scale. But in case of region 1a the 1 year periodicity is expected to more

17

prominent than in region 1 which is also supported with the comparatively lesser contribution of ENSO in region 1a as also shown. Again, because of the same reason, the year to year variability (shown by periodicity 1) should be most dominant in Lucknow followed by 1a and then 1. Same thing follows in the figure and interestingly, the effect of urbanization overshadows the other factors like ENSO in the periodicity analysis for Lucknow (due to presence of many outliers) as shown previously. The contribution of both outliers extremes with periodicities are seen almost comparable in region 1a. But in region 1 the effect of periodicities is more than the outliers as clearly seen with higher ACF in ENSO for region 1 compared to 1 year periodicity case. This clearly infers about the effect of urbanization which suppresses the effect of ENSO periodicity and thereby results in the drastic increase in DDF over Lucknow.

3.7. Probable influence of natural and anthropogenic components on DDF for region 3

In most of the preceding sections, the variability of DDF has been studied over Region 1 falling in the IGP. However, the north-western part of the country also comes under high drought severity zone as already discussed; hence this region is studied in detail now. Figure 2 has showed that the DDF trend is comparatively higher during the month of July; hence DDF during that month will be considered hereafter for further analysis over region 3. But it may be noted that the change is not so much prominent here as in region 1 (with a cumulative average of ~8% rise) and also the yearly fluctuations are too large which has subdued the trends resulting in a feeble rise of two days in the last 60 years (23-25 days) over this region shown in **Figure 6**. To start with the distribution analysis, three classes are made as SDP (14-20 dry days average 19) MDP (21-24 dry days average 22.6) and LDP (24.5-27.5 days with average 26 days) as depicted in Table 1. It may be noted that the values themselves have high magnitudes for all classes and the variability is also quite less (19-26 days) here compared to 9-22 in Lucknow; so the observed variation also should not be much prominent which is also evident from **Figure 6**. Further, as this region generally experiences arid climate, hence specific humidity can be an important factor here. Accordingly a decreasing trend is seen as supported by the median and lower bounds. But there is more overlapping among the classes and the total variance of humidity at 850 hpa is only between 12-10 which may not be strong enough to modulate drought intensities all by itself. SSN shows no definite variation hence not considered further. Aerosols and ENSO seem to have a weak increasing trend in their medians which again is diffused by more overlapping in these distributions. So this weaker variability is in

584 good agreement with the feeble trend in dry days, but simultaneously makes it difficult to
585 determine the potential driving factor behind the increasing DDF in region 3.

586   A better insight into the inter-dependence of all these components are investigated by
587 the PCA test in **Figure S9 (a).** The analysis reveals four PCA components out of which three
588 PCs are considered to explain the complete range of variances in dry days. The scores signify
589 no definite pattern with the total aerosol AOT assuming high pc1 and low pc2 pc3 while ENSO
590 has high pc2 and pc3 with lesser pc1 and SHUM falls in completely different quadrant. Now
591 since aerosols have higher pc1 component which is comparatively stronger than other pcs so it
592 may be a deciding factor. To clarify this confusion, MLR coefficients are calculated which
593 come around 0.107, 0.078, 0.056 and -0.267 also shown in Table 2. It is clear from the MLR
594 outputs that specific humidity has a strong negative influence on dry days so it will have good
595 effect on drought occurrences. But apart from this, the second dominant factor behind droughts
596 is still found to be the aerosols. However, this fact needs to be supported with more detailed
597 analysis as shown in the later sections.

598   In view of the previous sub-section, now the analysis is concentrated on the aerosols
599 components over region 3. The distribution analysis of aerosol components are now shown in
600 **Figure 6** which depicts that as usual, sea salt aerosols and sulphates have no role in
601 modulating the DDF. But it may be noted that here the magnitude of sea salts and sulphates are
602 higher than in region 1 or 1a may be due to its transport from the nearby seas which has not
603 been washed away by rain in its path owing to the arid climate. However, experience a very
604 prominent overlapping between the components which reduces the overall trend. The variation
605 of OC is not clear and hence is obliterated. BC as usual has a deterministic variance with some
606 overlapping; but still the whiskers and median values indicate its impact on dry days. Another
607 important aspect here is that, the range of values for these parameters are much lesser here due
608 to lesser urbanization which still affects the DDF. But the contribution of dust aerosols
609 emerges as the dominant component here as it not only shows higher values compared to all
610 other regions but it also signifies a clear trend in the medians and distribution values. Thus it
611 can be inferred that both dust and BC may contributed to this phenomena.

612   To investigate which parameter has more dominance in dry day formation, PCA
613 analysis is done on the individual components and the results are depicted in **Figure S9(b)**.
614 Here four PCAs are obtained, but the first two PCAs contribute 80% of variability so the 2D
615 variance is seen. Also the contribution of pc1 is comparable to pc2 so here both will be
616 important. While analysing the scores it is observed that only dust and BC have both high pc1

617 and pc2 so should be considered while most of the others have lower pc2 scores so they can be

618 neglected. Further investigation is done on MLR analysis towards the trend contribution which

619 also gives similar outputs as 0.464, 0.431, 0.120, 0.182, and 0.033 (Table 3). Again here both

620 BC and dust emerge potentially significant for the region 3 to be considered in association with

621 the slant rise in dry days. Both of these two components may have local sources but owing to

622 its location, there are possibilities of having added amount of dust aerosols being transported

623 from adjoining deserts or from dust storms and fumigation of dust from the ground during

624 intense dryness which are not found prominent over the region 1a (where BC and OC was high

625 due to high urbanization). Further for more meticulous observation cloud cover values have

626 also been checked (**Figure 7**) which shows that all the types of cloud cover have remained

627 almost unchanged over the years unlike region 1a and Lucknow. This is again in good

628 agreement with less prominent increase of anthropogenic emissions or in short less increase in

629 urbanization over region 3 compared to region 1a or Lucknow. This is further discussed in

630 coming sections. But few things are important to mention here: the trend of dry days in region

631 3 though it is weaker compared to region 1a may have serious impact in future as the region

632 already experiences high number of dry days itself so a slight increasing trend is also alarming.

633 Thus the effect of urbanization will be still an important parameter contributing towards the

634 hike of BC and (some of) dust aerosols growth and in turn leading to more strong trends in

635 DDF over this region.

636 **3.8. Impact of urbanization on DDF trends**

637 From the previous section, a strong association has been observed between dry day

638 frequency and anthropogenic emissions such as BC and OC which in turn is closely related

639 with the urbanization growth On the other hand, high population density is also generally

640 associated with the growth of urbanization and hence it may be taken as a suitable proxy for

641 the latter in this study. The population density values have been taken from the gridded 1°

642 population densities during 2000 -2015 (from SEDAC website) while the BC AOT extinction

643 datasets are utilized from MERRA-2. The BC AOT values have been averaged during the

644 month of August over a moving window of 5 years for Region 1, 1a Lucknow and Region 3 to

645 be synchronous with the population density measurements and their variations are shown in

646 Figure 8. The figure suggests that both population and BC AOT show a similar increasing

647 trend and this supports the utility of using the population data for further analysis

648 The primary distribution of population for year 2000 is shown in **Figure S10** which

649 depicts, more values at region 1a compared to region 1b, and another thing is that, Lucknow is

still found as a patch of very high population even at 2000. On the other hand, region 3 had much lesser populations at the same time. Next, the long term variation of population density are again observed over region 1, 1a and Lucknow from **Figure 8**. It may be noted from the figure that all throughout region 1, population density rises from 650 to 800 persons per sq kilometre which is quite a high value. Next, region 1a shows even higher values than region 1 with a steep rise from 760 to 1000 persons per square km. Thus it follows that region 1a has consistently higher population average and trends leading to higher OC and BC. However, the situation worsens in Lucknow where population density changes drastically from 850 to 1100 persons with most of the change happening in last 10 years; hence this phenomenon strongly supports the amplified DDF trend over Lucknow compared to 1a. But region 3 shows very less variations in last few years (100 to 140) leading to comparatively lesser BC and OC emissions. However, it may also be noted that the relative change over region 3 is higher (40%) compared to Lucknow (30%). Hence in future, if urbanization and population persists to grow at the same rate over region 3 then BC , OC and dust will also expectedly grow to alarming limits which can cause a drastic change in DDF over North-Western Indian regions.

3.9. Future trends of DDF over Region 1 and 3 using RCP 8.5 scenario

The next concern of this study is to investigate the projected change of dry phase lengths over the foreseeable future. Many attempts in the recent years have employed CMIP5 GCM simulations to provide future projections for any urbanization scenario. In accordance with the present study, RCP 8.5 projections of rainfall (and DDF) corresponding to maximum urbanization levels has been considered over the mentioned regions. It may be noted that in the last sixty years itself, DDF values have reached ~ 30 days in August, hence it is useful to study DDF in a two months span of mid-July to mid-September (having a reasonable increasing trend in dry days). The future projections of DDF over this time span is now obtained from 1950-2100. But the reliability and accuracy of these datasets first need to be validated from in-situ measurements. Hence, historical daily precipitation datasets of r1i1p1 realization from 11 well known GCM simulations are taken during 1955-2005 for all grid points in region 1 and 3 after which the DDF is calculated and recorded. Finally the averaged DDFs from each model was compared with the IMD data and the correlation coefficient with the normalized standard deviation values in **Table S3** indicate that three models namely: CAN ESM2, CNRM CM5, NORESM 1M show better agreement; hence they can be utilized to generate future projections for region 1 and 3 up to year 2100. For simplicity the yearly means of DDF historical data from

682   the models are also shown in **Figure S11** which again are found to follow the expected trends

683   of DDF in all three regions

684   Next the total variation in dry days are investigated over region 1 and 3 including both

685   historical and CMIP5 RCP 8.5 projections data to get a 150 year trend of dry day frequencies

686   in **Figure 9(a)**. The DDF for all 29 grid points in region 1 and 20 grid points over region 3 are

687   averaged yearly and then depicted in Figure 5 and 7. The multi model mean data shows that

688   even when averaged spatially, dry days show clear increase from ~ 8 days in 1950 to ~40 days

689   near 2100. Thus, Region 1 will experience a rise in DDF from 10% to 70% during mid

690   monsoon phase which is highly alarming and is attributed to the rapid pace of urbanization

691   over those regions in the future. Again, this trend looks less discretely increasing compared to

692   the historical trends over Lucknow. Again, in certain cases the projected DDF is expected to

693   increase up to ~50 days (80%) during the 2100 monsoon which should lead to severe drought

694   conditions. Again, the trends look comparatively weaker in first fifty years (8-12), then it gets

695   stronger (12-24) and finally shoots up to very high values (24-42 days) after 2050 which is

696   primarily caused due to high urbanization rate over this region in the future. However, when

697   the same analysis was done for region 3 DDF was found to increase steadily from 20 to 40

698   days over 150 years. The trends of DDF are clearly much weaker in region 3 compared to

699   region 1 while the standard error bars are also less here. Both of this factors can be attributed to

700   the fact that region 3 has much less urbanization components than region 1. But it may be

701   noted that if region 3 continues to face urbanization at the present rate, then in future it will

702   experience more number of dry days. Additionally, it has been observed that, the trends have

703   increased almost steadily in region 3 with no abrupt change in DDF in the last 50 years like

704   region 1. This is attributed to the low urbanization levels at region 3 at present.

705   Hence region 1 creates a more alarming situation with dry days increasing by around 5

706   times compared to the other regions. So to further investigate this abrupt change spatially, the

707   model averaged data of DDF for 50 years span are shown for region 1 in the bottom panel of

708   **Figure 9(b)**. The figure shows an expected high value around Lucknow for the 50 year

709   periods; but its effect diffuses as one goes towards the outskirts of Lucknow facing lesser

710   urbanization. Another thing is that the places adjoining Lucknow show a very drastic change

711   only after 2010. Thus, most of the places adjoining Lucknow shows very high number of dry

712   days (>45 days) near the end of this century which will grossly affect the monsoonal rainfall

713   leading to severe droughts and so it needs to be addressed by policy makers.

714

4. Conclusions

It is an essential aspect to study the probability of drought occurrences over India during monsoon as agricultural and economical issues are directly related with it. Here, a detailed study on the occurrence of dry days during monsoon over the Indian region is presented. The study investigates three potentially drought prone regions in India based on the dearth of precipitation and abundance of PET. Region 1 mostly belongs to the State of Uttar Pradesh (UP), Region 2 covers major parts of the states of Andhra Pradesh and Tamil Nadu and small portion of Karnataka while Region 3 encompass the arid part of Rajasthan. A series of investigations are progressed which infer that over the eastern part of region 1 which is referred as region 1a urbanization plays significant role in increasing DDF. Prevailing impact of anthropogenic emission like BC or OC aerosols becomes more prominent as the study goes in depth with a downscaling approach from a broad region 1 to a specific urbanized location like Lucknow which is one of the urbanized sectors of IGP. The increase in cloud cover and non-occurrence of rain events indicate rain suppression phenomena over region 1 which is yet to be investigated in detail. This also indicates the scope of the study over several other point locations having drought occurrence record but could not be included in the present study approach. Finally, the long term projections of DDF are drawn over region 1a and 3 using intense urbanization scenario of RCP 8.5 and an average of 70% rise in dry days are seen which may be a very crucial concern by the year 2100 and hence it needs to be considered by policy makers in future aspects. However, this study is mainly done from modelled components of aerosols, so a far more accurate analysis can later be done over IGP subject to more availability of aerosol in-situ data in the other major urban locations over India. The main findings of the study are shown in a schematic presentation in **Figure S12** and are highlighted as follows:

➢ The DDF (based on the frequency of days having local precipitation accumulation less than 1mm) has a significant level of correlation with the universally accepted monthly SPEI Drought Index (DI) especially in the last sixty years. Further, the correlation levels between DI and DDF are more prominent during August in Region 1a and during July in region 3.

➢ The trends of DDF (within 15 days window) are more prominent during August for region 1a. However, region 3 shows a descent trend during July while region 2 shows the same during late September, (corresponding to the monsoon retreating phase) hence it has been neglected as it may not completely reflect a monsoonal drought.

23

748   ➢   Results from region 1a indicate prevailing contribution of aerosols compared to ENSO,
749   Humidity or SSN. Further studies show that BC and OC aerosols over urbanized region
750   are more active in increasing the DDF, which is also supported from distribution, PCA and
751   MLR analysis

752   ➢   The trend analysis on DDF reveals that the increasing trends become stronger as the
753   spatial coverage is downscaled from region1 to 1a and followed by a local urbanized
754   location of Lucknow. About 50% increase in DDF is found in Lucknow compared to 17%
755   all through region 1. Further, a periodicity of 4 and 8 years is found stronger in region 1
756   which gets overpowered by the randomly varying urbanization component over Lucknow.

757   ➢   Population density maps have been taken as a proxy of the urbanization component
758   owing to its significant agreement with anthropogenic carbonaceous emissions (BC). A
759   higher population density is observed over Lucknow (average of 850 persons/km^2 and
760   trends of~35%) compared to the rest of region 1 and 1a. Further the population density
761   values are very less in region 3 (100 persons/km^2) which is in good agreement with lesser
762   impact of urbanization on DDF over this region.

763   ➢   In depth investigation revealed that the increase of urbanization components like BC or
764   OC exibits a significant association with increased cloud lifetime (~ 18% rise in LCC)
765   over Lucknow which results in a stronger gradient of dry day occurrences (from 9 days in
766   1956 to ~17 days at present).

767   ➢   Though in region 3 aridity plays a major role to experience a high number of dry days
768   (~23) still dust aerosols show an increasing trend and hence it probably influences a
769   further increase in DDF (an increase from 23 days in 1956 to 25 days at present) which is
770   alarming for region 3.

771   ➢   The climatic projections of dry day frequency from CMIP5 simulations of 3 GCM
772   model (CNRM CM5, CAN ESM and NOR ESM 1M) show a sharp increase in dry days
773   during July 15 to September 15 with DDF reaching up to 50 dry days over region 1 and 45
774   days over region 3 by 2100.

775

Author contributions

777 RC performed the main analysis, ST and MVR provided the initial concept, main guidance,
778 writing and editing of the manuscript, BKG and AM provided the needed data and contributed
779 to the analysis, discussion, and editing.

780

References

1. Ackerman, A. S., Toon, O. B., Taylor, J. P., Johnson, D. W., Hobbs, P. V and Ferek, R. J.: Effects of aerosols on cloud albedo: Evaluation of Twomey�s parameterization of cloud susceptibility using measurements of ship tracks, J. Atmos. Sci., 57(16), 2684–2695, 2000.

2. Agnihotri, R., Dutta, K., Bhushan, R. and Somayajulu, B. L. K., Evidence for solar forcing on the Indian monsoon during the last millennium, Earth and Planetary Science Letters 198, 521 – 527, 2002.

3. T. Alaka, Gaddam, G. and others: Monsoonal Droughts In India--A Recent Assessment, Pap. Glob. Chang. IGBP, 2015.

4. BEGUER'\iA, S., Vicente-Serrano, S. M. and Angulo-Mart'\inez, M.: A multiscalar global drought dataset: the SPEIbase: a new gridded product for the analysis of drought variability and impacts, Bull. Am. Meteorol. Soc., 91(10), 1351–1356, 2010.

5. Benton, G. S.: Drought in the United States analyzed by means of the theory of probability, United States Department of agriculture., 1942.

6. Bhalme, H. N. and Mooley, D. A.: Large-scale droughts/floods and monsoon circulation, Mon. Weather Rev., 108(8), 1197–1211, 1980.

7. Buchard, V., Randles, C.A., Da Silva, A.M., Darmenov, A., Colarco, P.R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A.J., Ziemba, L.D. and Yu, H., 2017. The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies. Journal of Climate, 30(17), 6851-6872, 2017, DOI:10.1175/JCLI-D-16-0613.1

8. Chakraborty, R., Talukdar, S., Saha, U., Jana, S. and Maitra, A.: Anomalies in relative humidity profile in boundary layer during convective rain, Atmos. Res., 191, 74–83, 2017.

9. Cliver, E. W., Clette, F. and Svalgaard, L.: Recalibrating the sunspot number (SSN): the SSN workshops, Cent. Eur. Astrophys. Bull, 37(2), 401–416, 2013.

812   10. Dai, A., Trenberth, K. E. and Qian, T.: A global dataset of Palmer Drought Severity Index
813   for 1870--2002: Relationship with soil moisture and effects of surface warming, J.
814   Hydrometeorol., 5(6), 1117–1130, 2004.

815   11. Dey, S., Tripathi, S. N., Singh, R. P. and Holben, B. N.: Influence of dust storms on the
816   aerosol optical properties over the Indo-Gangetic basin, J. Geophys. Res. Atmos.,
817   109(D20), 2004.

818   12. For International Earth Science Information Network (CIESIN)�Columbia University, C.:
819   Gridded Population of the World, Version 4 (GPWv4): Population Density, 2016.

820   13. Francis, P. A. and Gadgil, S.: Towards understanding the unusual Indian monsoon in 2009,
821   J. Earth Syst. Sci., 119(4), 397–415, 2010.

822   14. Gadgil, S., Vinayachandran, P. N. and Francis, P. A.: Droughts of the Indian summer
823   monsoon: Role of clouds over the Indian Ocean, Curr. Sci., 1713–1719, 2003.

824   15. Gore, P. G. and Ray, K. C. S.: Variability in drought incidence over districts of
825   Maharashtra, Mausam, 53(4), 533–538, 2002.

826   16. Guhathakurta, P. and Rajeevan, M.: Trends in the rainfall pattern over India, Int. J.
827   Climatol., 28(11), 1453–1469, 2008.

828   17. Heim Jr, R. R.: A review of twentieth-century drought indices used in the United States,
829   Bull. Am. Meteorol. Soc., 83(8), 1149–1165, 2002.

830   18. Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H.,
831   Menne, M. J., Smith, T. M., Vose, R. S. and Zhang, H.-M.: Extended reconstructed sea
832   surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons, J.
833   Clim., 30(20), 8179–8205, 2017.

834   19. Krishnamurti, T. N., Thomas, A., Simon, A. and Kumar, V.: Desert air incursions, an
835   overlooked aspect, for the dry spells of the Indian summer monsoon, J. Atmos. Sci., 67(10),
836   3423–3441, 2010.

837   20. Kumar, K. N., Rajeevan, M., Pai, D. S., Srivastava, A. K. and Preethi, B.: On the observed
838   variability of monsoon droughts over India, Weather Clim. Extrem., 1, 42–50, 2013.

839   21. Lloyd-Hughes, B. and Saunders, M. A.: A drought climatology for Europe, Int. J.
840   Climatol., 22(13), 1571–1592, 2002.

841   22. Marcovitch, S.: The measure of droughtiness, Mon. Weather Rev., 58(3), 113, 1930.

842   23. McKee, T. B., Doesken, N. J., Kleist, J. and others: The relationship of drought frequency
843   and duration to time scales, in Proceedings of the 8th Conference on Applied Climatology,
844   vol. 17, pp. 179–183., 1993.

845   24. Moorthy, K. K., Babu, S. S., Satheesh, S. K., Srinivasan, J. and Dutt, C. B. S.: Dust
846   absorption over the ◆Great Indian Desert◆ inferred using ground-based and satellite
847   remote sensing, J. Geophys. Res. Atmos., 112(D9), 2007.

848   25. Munger, T. T.: Graphic method of representing and comparing drought intensities., Mon.
849   Weather Rev., 44(11), 642–643, 1916.

850   26. Pai, D. S., Sridhar, L., Guhathakurta, P. and Hatwar, H. R.: District-wide drought
851   climatology of the southwest monsoon season over India based on standardized
852   precipitation index (SPI), Nat. hazards, 59(3), 1797–1813, 2011.

853   27. Palmer, W. C.: Meteorological drought. Research Paper No. 45. Washington, DC: US
854   Department of Commerce, Weather Bur., 59, 1965.

855   28. Pandey, S.K., Vinoj, V., Landu, K. and Suresh Babu, S., Declining pre-monsoon dust
856   loading over South Asia: Signature of a changing regional climate Scientific Reports, 7:
857   16062, 2017,

858   29. Parthasarathy, B., Sontakke, N. A., Monot, A. A. and Kothawale, D. R.: Droughts/floods in
859   the summer monsoon season over different meteorological subdivisions of India for the
860   period 1871--1984, J. Climatol., 7(1), 57–70, 1987.

861   30. Raman, C. R. V and Rao, Y. P.: Blocking highs over Asia and monsoon droughts over
862   India, Nature, 289(5795), 271–273, 1981.

863   31. Sato, Y., Goto, D., Michibata, T., Suzuki, K., Takemura, T., Tomita, H. and Nakajima, T.:
864   Aerosol effects on cloud water amounts were successfully simulated by a global cloud-
865   system resolving model, Nat. Commun., 9(1), 985, 2018.

866   32. Singh, N. and Ranade, A.: The wet and dry spells across India during 1951--2007, J.
867   Hydrometeorol., 11(1), 26–45, 2010.

868   33. Sivakumar, M., Stone, R., Sentelhas, P. C., Svoboda, M., Omondi, P., Sarkar, J. and
869   Wardlow, B.: Agricultural drought indices: summary and recommendations, in Agricultural
870   drought indices Proceedings of an expert meeting, pp. 2–4., 2010.

871   34. Solmon, F., Nair, V. S. and Mallet, M.: Increasing Arabian dust activity and the Indian
872   summer monsoon, Atmos. Chem. Phys., 15(14), 8051, 2015.

873   35. Sushama, L., Said, S. Ben, Khaliq, M. N., Kumar, D. N. and Laprise, R.: Dry spell
874   characteristics over India based on IMD and APHRODITE datasets, Clim. Dyn., 43(12),
875   3419–3437, 2014.

876   36. Talukdar, S., Jana, S. and Maitra, A.: Dominance of pollutant aerosols over an urban region

877   and its impact on boundary layer temperature profile, J. Geophys. Res. Atmos., 122(2),

878   1001–1014, 2017.

879   37. Thornthwaite, C. W.: An approach toward a rational classification of climate, Geogr. Rev.,

880   38(1), 55–94, 1948.

881   38. Twomey, S.: The influence of pollution on the shortwave albedo of clouds, J. Atmos. Sci.,

882   34(7), 1149–1152, 1977.

883   39. Wencai, W., Sheng, L., Jin, H. and Han, Y.: Dust aerosol effects on cirrus and altocumulus

884   clouds in Northwest China, J. Meteorol. Res., 29(5), 793–805, 2015.

885   40. Wilhite, D. A. and Glantz, M. H.: Understanding: the drought phenomenon: the role of

886   definitions, Water Int., 10(3), 111–120, 1985.

887

888

889   **Tables**

890

Region	Case 1(SDP)		Case 2(MDP)		Case 3(LDP)	
	Range	Average	Range	Average	Range	Average
Region 1a	8-10	9	10-14	12.5	14-18	16
Lucknow	4-12	9.5	13-17	15	18-30	22
Region 3	14-20	19	21-24	22.6	24.5-27.5	26

891   **Table 1.** Classification of dry day phase according its length.

Region	Components			
	Aerosol	SSN	ENSO	SHUM
Region 1a	0.393	0.008,	0.161	-0.207
Region 3	0.107	0.078	0.056	-0.267

892   **Table 2.** MLR coefficients for all general factors.

Region	Components				
	BC	Dust	OC	Sea Salt	Sulphate
Region 1a	0.542	0.129	0.263	0.326	0.124
Lucknow	0.864	0.218	0.556	0.011	0.155
Region 3	0.464	0.431	0.120	0.182	0.033

893   **Table 3.** MLR coefficients for aerosol components.

894

895
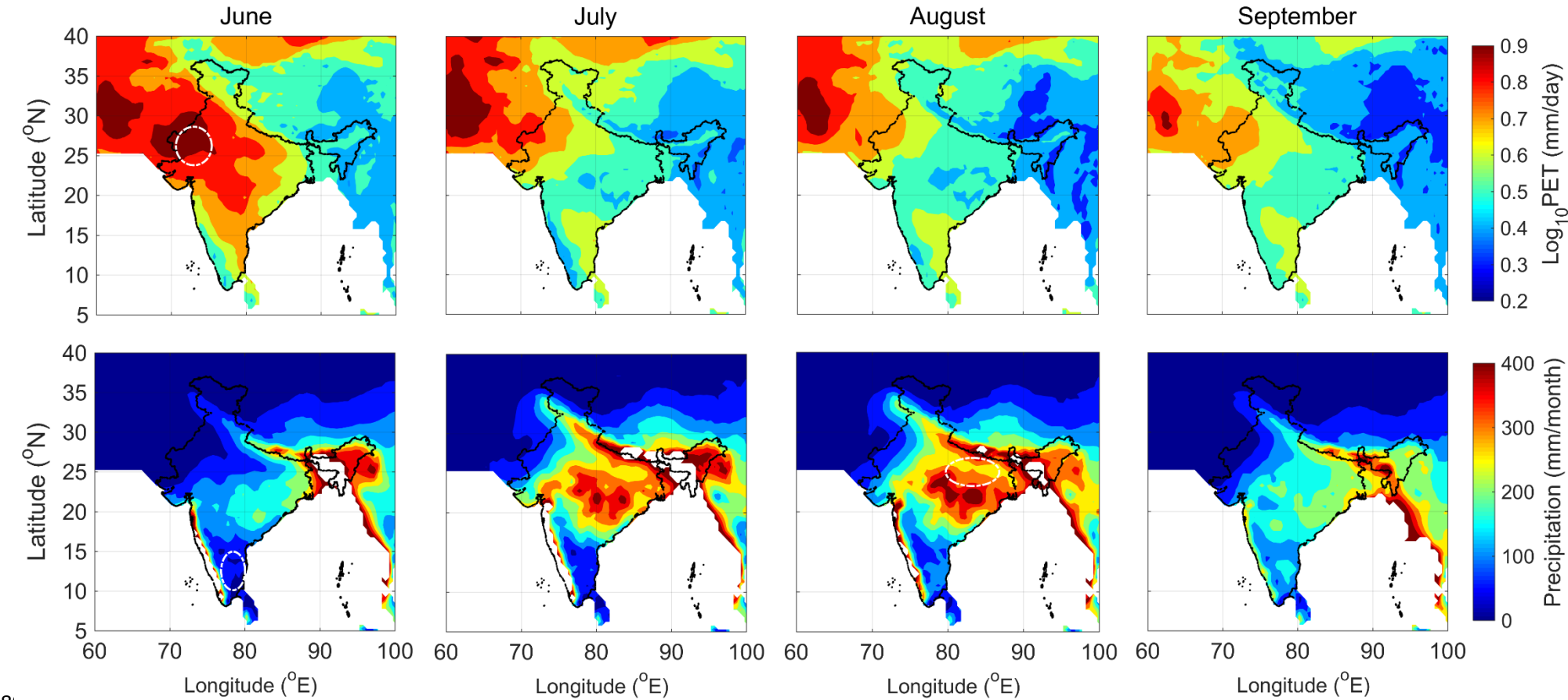
Figures
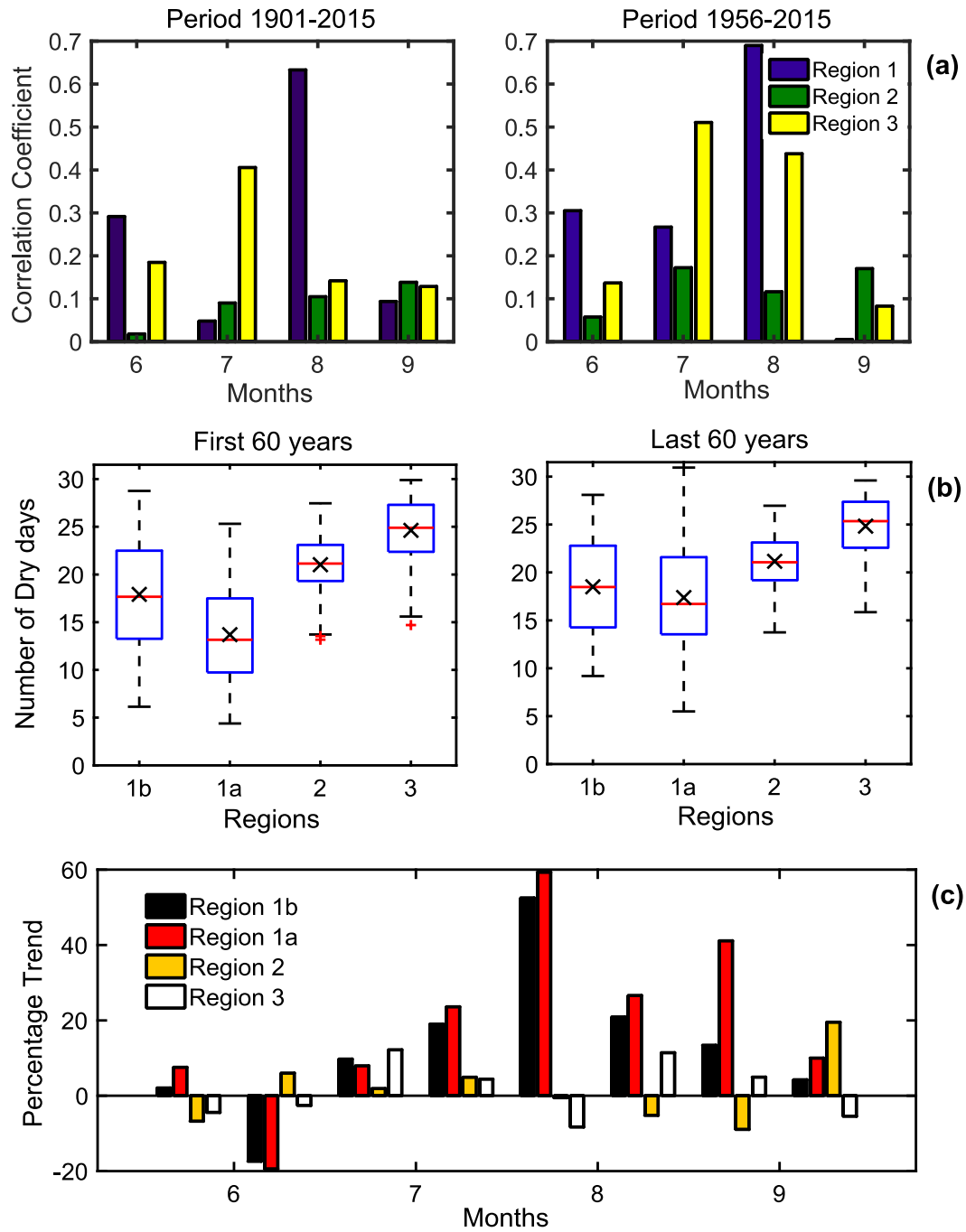
Figure 1. Monthly averaged maps of potential evapo-transpiration rate and precipitation during June-September.

901

902

Figure 2(a). Correlation coefficients between DI and DDF values for all monsoon months for two different climatic periods 1901-2015 and 1956-2015, (b)Monthly mean values of DDF for regions 1a, 1b, 2 and 3 during 1901-1960 and 1956-2015, (c) 15 day trends of DDF trends during 1956-2015.
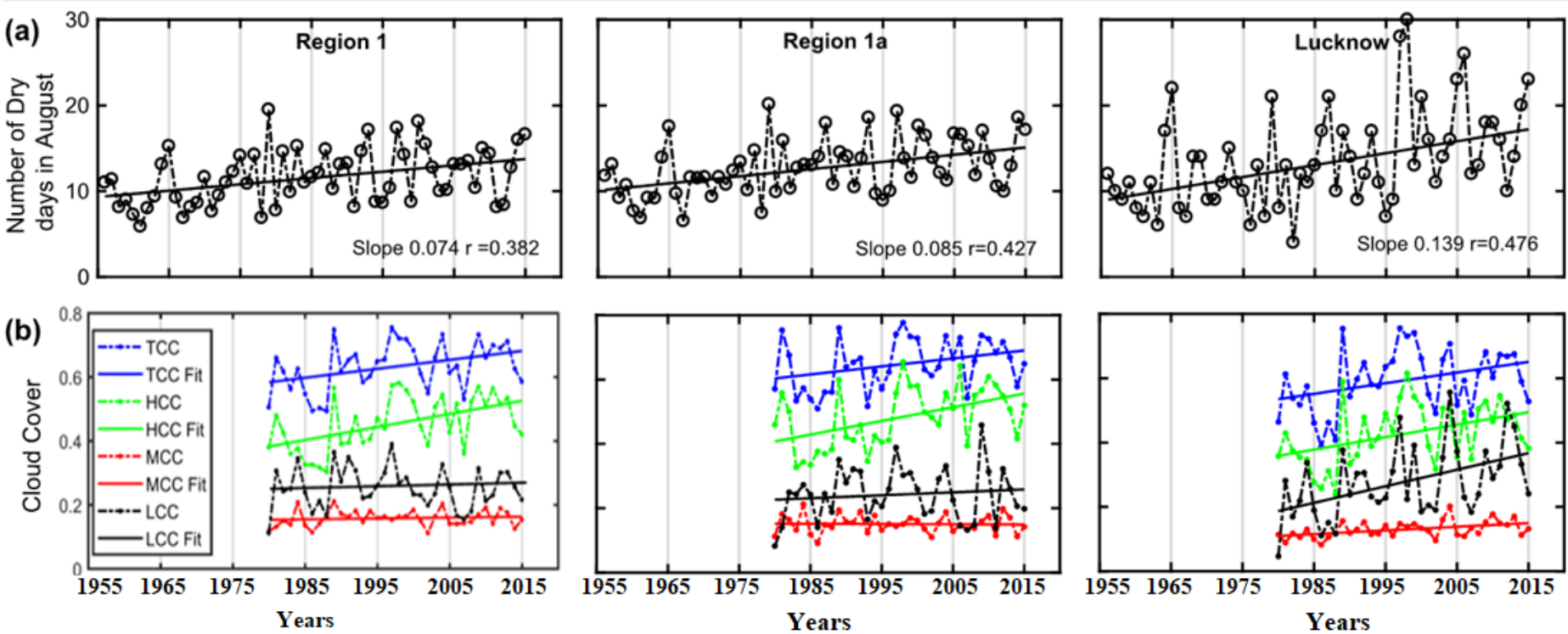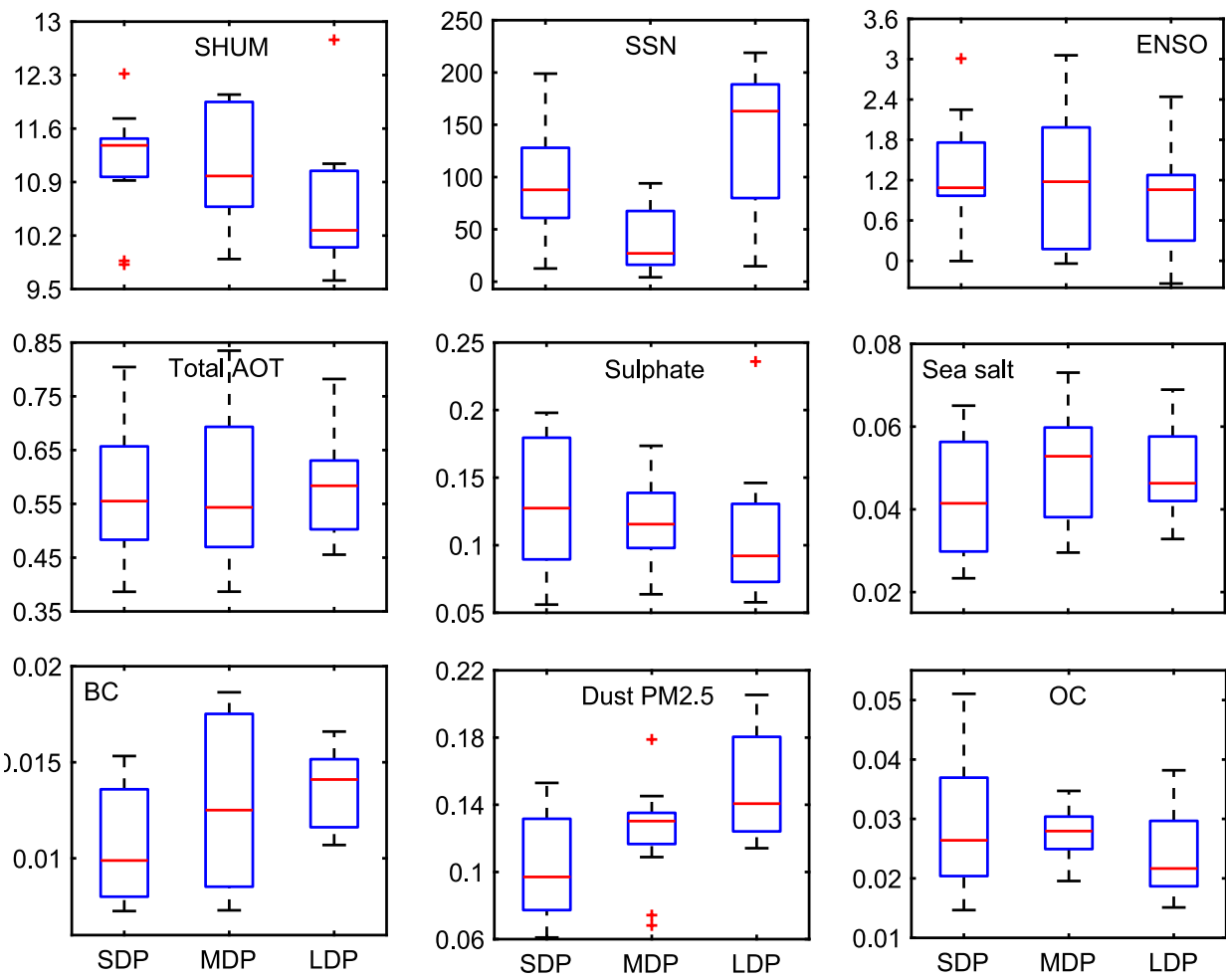
906

907

908
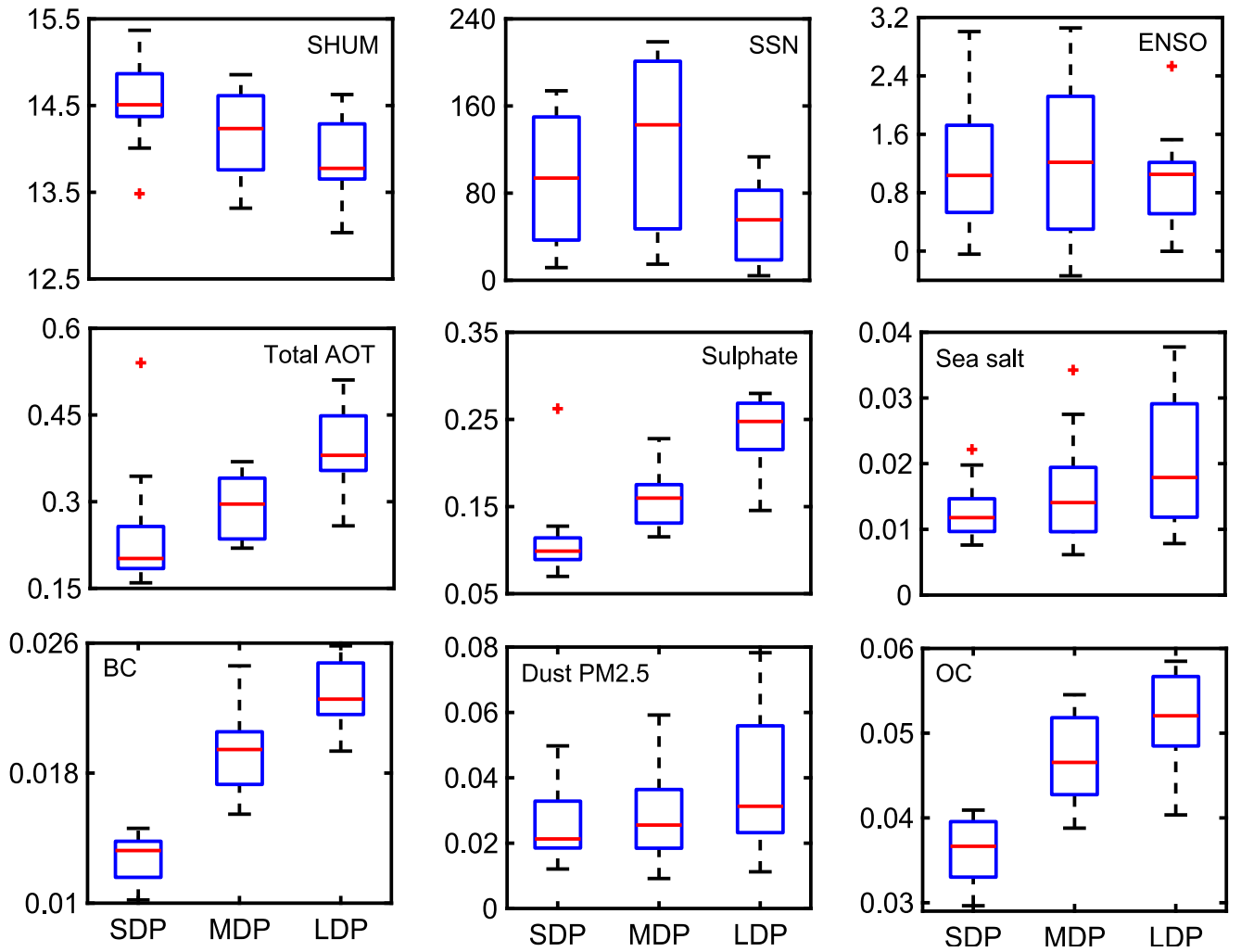
909 **Figure 3.** Frequency distribution analysis results of various controlling factors behind DDF evolution
910 for various types of dry phase lengths over region 1a, namely: Humidity, SSN, ENSO Total aerosols,
911 Sulphates, Sea Salt, BC, Dust PM 2.5 and OC.

912

913

914

915
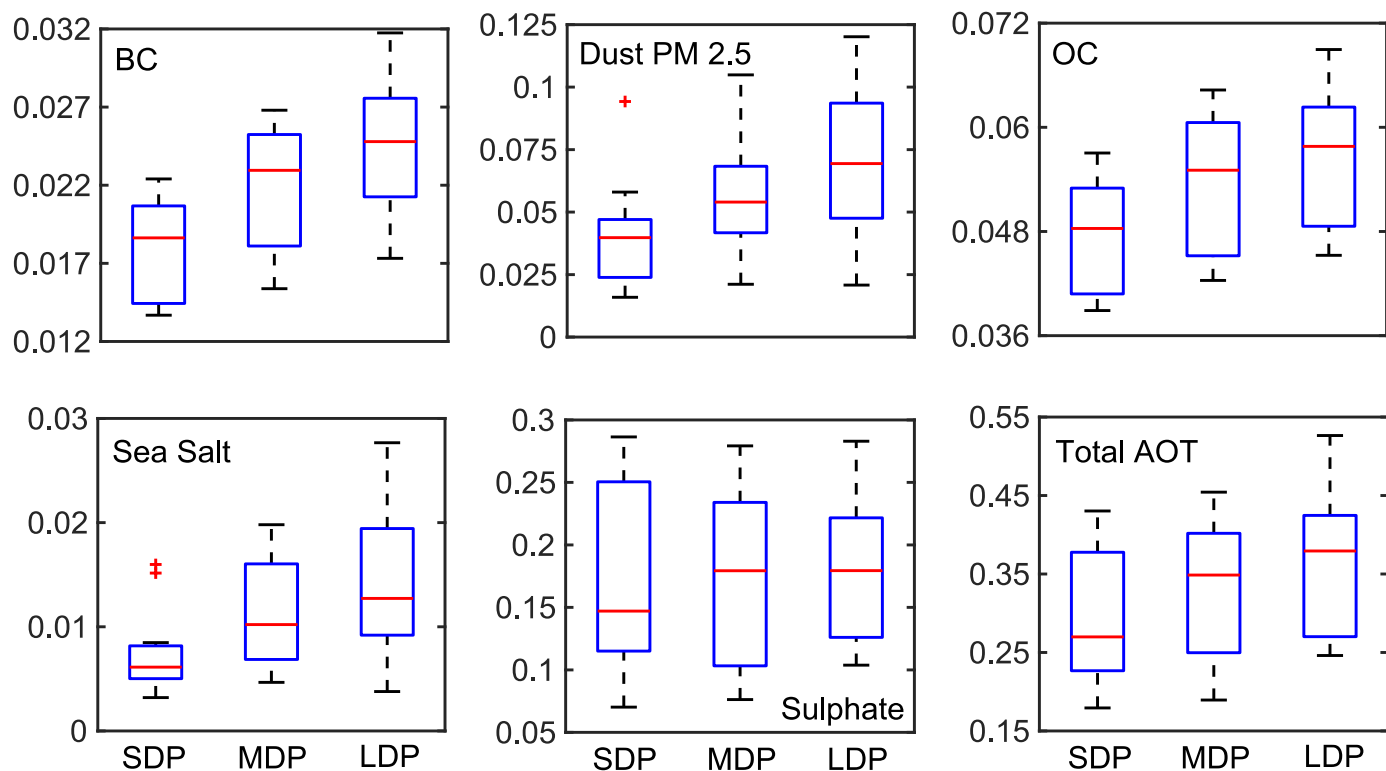
Figure 4. Frequency distribution analysis results of various controlling factors behind DDF evolution for various types of dry phase lengths over Lucknow corresponding to 5 aerosol components such as BC, Dust PM 2.5, OC, Sea Salt and Sulphates.

924

Figure 5. Statistical comparison of the climatology of all parameters during August for region 1, region 1a, Lucknow during various time spans
(a) Dry Day Frequency values between 1956-2015 and (b) Cloud cover parameters (TCC,HCC,MCC,LCC) during 1980-2015 .

925

926

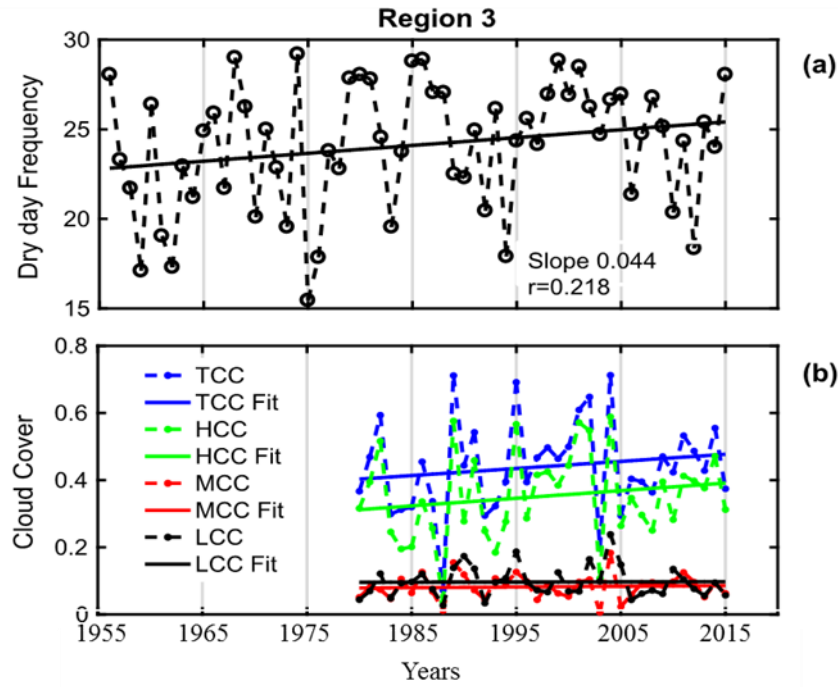927

928

929

931 **Figure 6.** Frequency distribution analysis results of various controlling factors behind DDF evolution
932 for various types of dry phase lengths over region 3, namely: Humidity, SSN, ENSO, Total aerosols,
933 Sulphates, Sea Salt, BC, Dust PM 2.5 and OC.

934

Figure 7. Statistical comparison of the climatology of all parameters during July for region 3 during various time spans (a) Dry Day Frequency values between 1956-2015 and (b) Cloud cover parameters (TCC,HCC,MCC,LCC) during 1980-2015.
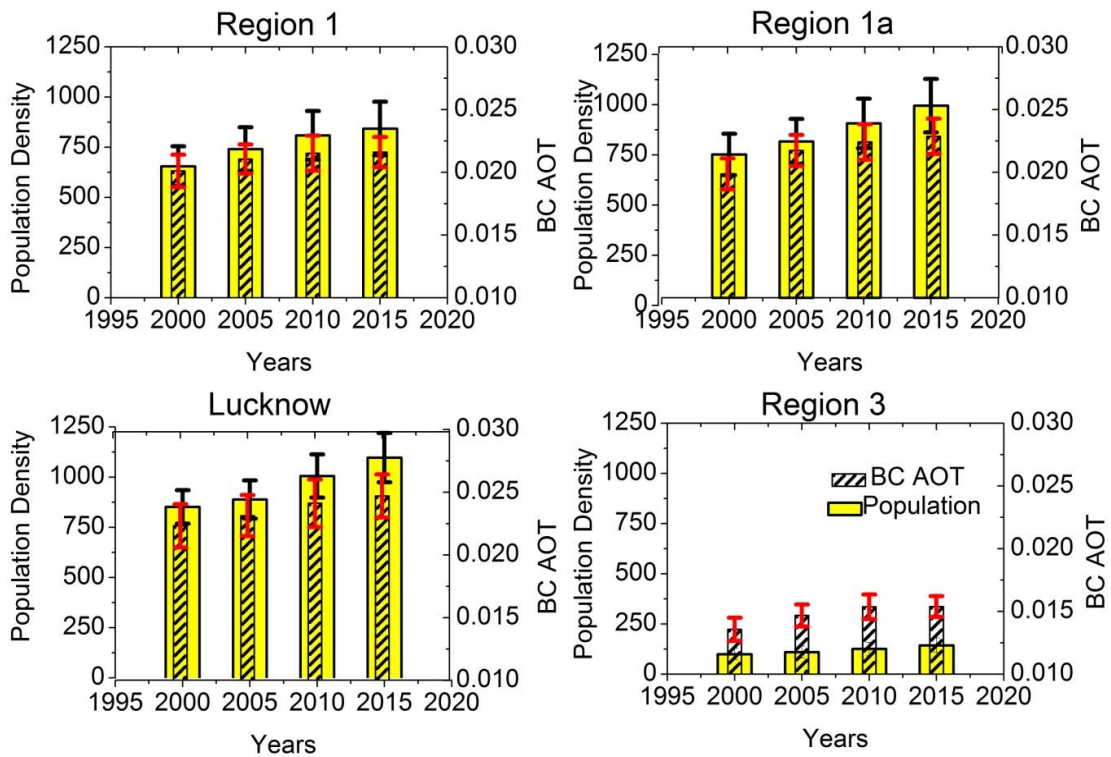
Figure 8. Region-wise population densities and BC AOT values (during August) for Region 1, 1a, Lucknow and Region 3 during (2000-2015), vertical bars represent the corresponding 1 sigma standard deviations values.

944

945

946
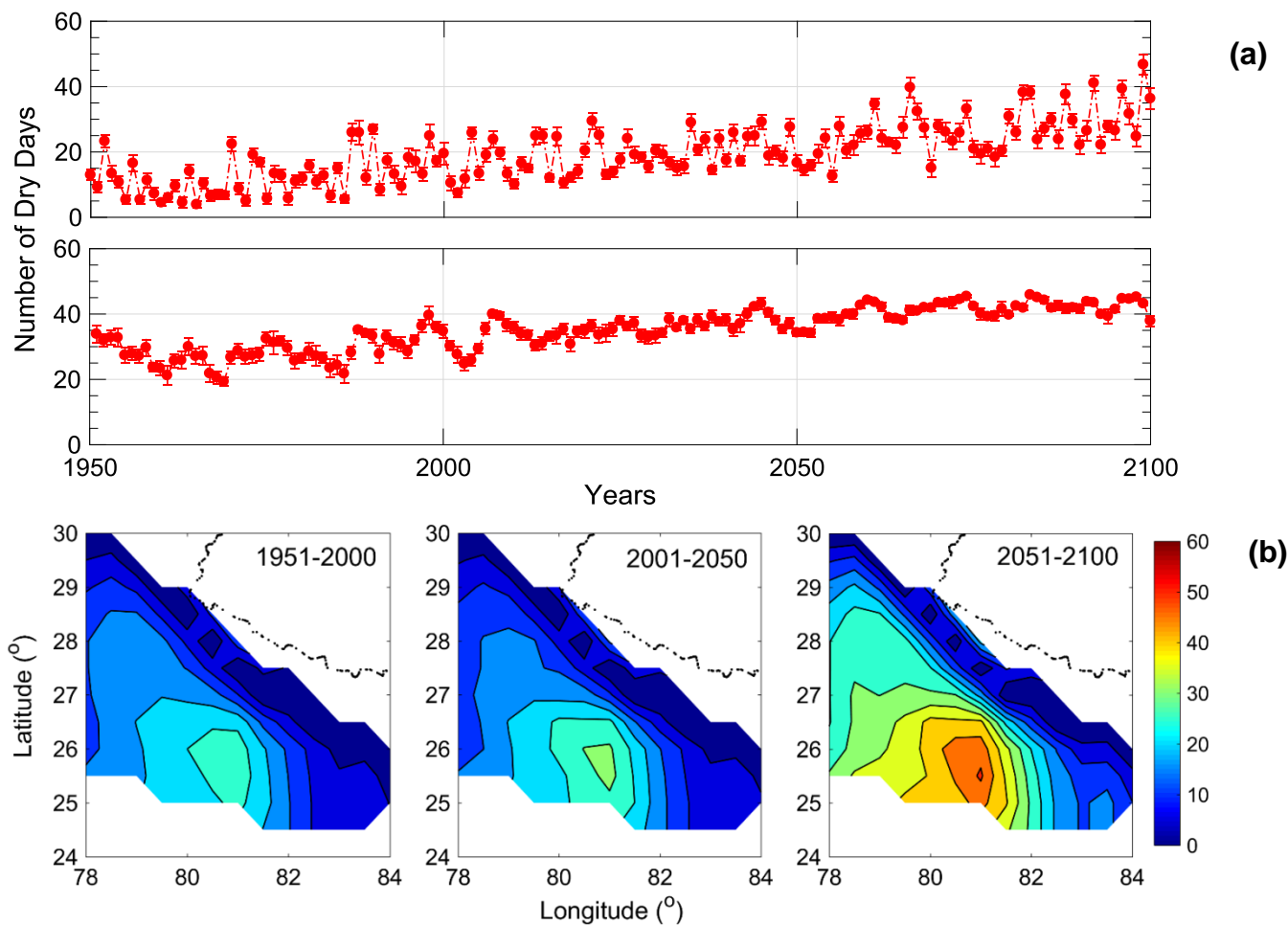
947 **Figure 9.** (a) Climatic variations in dry day frequency over Region 1 and 3 containing both historical data (upto 2005) and RCP8.5 projections
948 (2006-2100) of multi model mean from 3 selected GCMs (b) Projected lat-lon maps of DDF for all three 50 year periods from 1951-2100.